UNIVERSITÄT BERN

Graduate School for Cellular and Biomedical Sciences

University of Bern

Molecular characterization of rare forms of canine neurological diseases as potential models for similar human diseases

PhD Thesis submitted by

Anna Letko

from **the Czech Republic**

for the degree of

PhD in Computational Biology

Supervisor
Prof. Dr. Cord Drögemüller
Institute of Genetics
Vetsuisse Faculty of the University of Bern

Co-advisor
Dr. Rémy Bruggmann
Interfaculty Bioinformatics Unit
Faculty of Science of the University of Bern

Accepted by the Faculty of Medicine, the Faculty of Science and the Vetsuisse Faculty of the University of Bern at the request of the Graduate School for Cellular and Biomedical Sciences

Bern, Dean of the Faculty of Medicine

Bern, Dean of the Faculty of Science

Bern, Dean of the Vetsuisse Faculty Bern

Abstract

Canis lupus familiaris, the domestic dog, possesses a huge variability in traits such as size, conformation, coat color, or character, which reflects the generations of targeted human selection after the dog's domestication thousands of years ago. The phenotypic differences naturally reflect the underlying, often breed-specific, genetic variation. While heterogeneity between breeds is large, at the same time, the individuals within one breed are usually very homogeneous. This specific population structure of modern dog breeds with low genetic diversity favors the propagation of spontaneous occurrences of genetic mutations that might lead to the development of diseases, and thus makes the dog a valuable animal model. Inherited neurological disorders in animals as well as in human patients are incurable, often severe, and result in progressively worsening quality of life; early diagnosis is therefore beneficial for managing the disease development. DNA-based precision medicine using state-of-the-art methods, e.g. whole-genome sequencing (WGS) has been successfully utilized in recent years for routine diagnosis of rare diseases in human as well as in veterinary medicine. Identification of disease-causing variants allows dog breeders to avoid the spread of such variants in the affected dog breed, ultimately improving the health of the whole population through better breeding management, as well as to advance the understanding of the molecular mechanisms involved in corresponding human disease, and may be useful for the development of novel therapeutic strategies.

In this thesis, I took part in the analysis of seven specific canine neurological phenotypes applying different genetic mapping methods, candidate gene analysis, and WGS. I also generated and analyzed extensive genealogical and genomic data on the worldwide Leonberger dog population in regards to its diversity and disease prevalence. Despite its increasing size in recent years, the population lost considerable genetic diversity due to a bottleneck in the last century. The heavy use of popular sires led to high relatedness among the breeding dogs and thus to high inbreeding rates. This facilitated undesirable genetic traits to spread within the gene pool of the Leonberger breed. A private homozygous frameshift variant in the *GJA9* gene was identified in Leonbergers with an adult-onset form of polyneuropathy using genome-wide association study (GWAS) and WGS. The *GJA9* gene encodes a connexin gap junction family protein, which are important components of peripheral myelinated nerve fibers; this discovery for the first time adds *GJA9* to the list of candidate genes for similar human conditions. During a study of additional forms of polyneuropathy and/or laryngeal paralysis, I found a missense variant in the *CNTNAP1* gene in Leonbergers and Saint Bernards showing early signs of laryngeal paralysis. *CNTNAP1* encodes a contactin-associated protein important for the organization of myelinated axons and has been implicated in various forms of human neurological diseases. Interestingly, this variant was seen in several

other unrelated dog breeds and most likely predates modern breed establishment. A similar approach revealed two independent variants in the *NAPEPLD* gene in Leonberger and Rottweiler dogs affected by leukoencephalomyelopathy providing evidence for allelic heterogeneity of this disorder and the first description of *NAPEPLD*-associated inherited defects in the endocannabinoid system associated with myelin disorders. In another study, a form of canine neuroaxonal dystrophy occurred in young adult Rottweiler dogs. WGS data of two cases revealed a homozygous missense variant in the *VPS11* gene, encoding a member of VPS class C complex, a key factor of the endosome-autophagosome-lysosome pathway, previously associated with an infantile-onset neurological syndrome in humans. In a family of Alpine dachsbracke dogs, I used linkage analysis and homozygosity mapping to discover an autosomal recessive variant in the puppies affected by spinocerebellar ataxia, which affects the *SCN8A* gene. The gene encodes a subunit of a channel important for sodium ion transport to neurons in the central nervous system and was previously implicated in human neurogenetic conditions. By in-depth pedigree analysis, I found a common ancestor of two geographically separated families of Saluki dogs in which puppies suffering from succinic semialdehyde dehydrogenase deficiency causing neurological abnormalities were observed. GWAS and subsequent filtering of WGS data of two affected Saluki cases identified a causative variant in the *ALDH5A1* gene encoding an essential enzyme of the gamma-aminobutyric acid neurotransmitter metabolic pathway. Finally, the underlying genetics of a previously described Leigh-like subacute necrotizing encephalopathy in Yorkshire terriers was solved by discovering a perfectly associated loss-of-function indel variant in the *SLC19A3* gene encoding thiamine transporter 2, which is important in brain development, and its disruption was previously seen in similar human neurometabolic disease.

In conclusion, the discovery of the herein described likely pathogenic DNA variants enabled systematic genetic testing of breeding dogs, and selection against the corresponding disorders to improve the health and welfare of the respective populations. This thesis provides molecular descriptions of several canine neurological conditions and presents additional physiologically relevant models of corresponding human diseases. Apparently, species- and site-specific differences in pathological phenotypes for mutations within the same gene exist as seen, e.g. in canine *VPS11*-related neuroaxonal dystrophy. All these spontaneous canine models closely resemble rare human syndromes and provide physiologically relevant models to better understand poorly characterized gene functions, e.g. in defects of the endocannabinoid system related to *NAPEPLD*; and provide potential new candidate genes for corresponding human forms of diseases with yet unsolved genetic etiology, e.g. *GJA9*-associated polyneuropathy. Therefore, this thesis demonstrates that genomic studies of domestic animal species such as the dog improve the understanding of rare complex and heterogeneous groups of neurodegenerative disorders.

Table of contents

Introduction

Hundreds of distinct dog breeds were formed due to human selection and breeding towards specific characteristics since the domestication of dogs. This development ultimately led to the most phenotypically diverse mammal known [1]. During the process, some breed-defining genetic variants underlying particular phenotypic traits such as size, shape, coat type and color, or certain behavior became fixed in the modern breeds. However, artificial selection of purebred mating pairs resulted not only in the desired and often unique phenotypes, but also led to reduced genetic diversity in the closed breeding populations that eventually increased the risk of developing an inherited disease [2]. While historically dogs served predominantly as herders, guardians, or hunters, nowadays they are largely considered assistants, companions, and family members. Therefore, the breeders and owners are increasingly more interested in the health of their dogs including the genetic background of various disorders. Thanks to the public availability of annotated canine genome sequence and the constant development of advanced molecular genetics methods, as well as the active participation of dog owners, all kinds of genotype-phenotype relationships can be studied.

Dog as animal model

Humans first domesticated the dog from wolf more than 15,000 years ago [3]. Archeological and molecular genetic data found that the domestication events happened more than once in different places, and dogs have followed human migration ever since [4,5]. While natural selection happens unintentionally when animals adapt to a changing environment, artificial selection is a process conducted by people during domestication in order to tame wild animals [6]. The popularity and rapid spread of the dog resulted in divergent populations and development of specialized breeds. Over the last few centuries, extreme phenotypic diversification of the species led to the creation of hundreds of dog breeds able to herd, guard, hunt, or guide [1,7–9]. Initially, the selection focused mainly on behavioral characteristics to create the most valuable working dogs. Later, morphological appearance became important and the strong selection pressure for a specific size, coat type, or color resulted in impressive variability among the breeds. In the 19th century, breeding became more restrictive when studbooks for most breeds were closed, and the different populations were reproductively isolated, forming so-called modern dog breeds [1].

Today, the Fédération Cynologique Internationale (the World Canine Organisation) recognizes 353 registered breeds and maintains their standards [10]. This phenotypic variability is accompanied by specific changes in canine genetic architecture. Research of the genetic differences improves our understanding of the canine's diverse history, as well as the underlying molecular mechanisms of different traits and diseases. Restrictive breeding

practices, combined with the use of popular sires (i.e. preferential breeding with only a few individuals, especially males), along with frequent inbreeding have all increased within-breed relatedness, reflecting the decreased genetic diversity as many modern breeds descend from only a limited number of founders. As a consequence, high heterogeneity can be found between breeds, while homogeneity exists within breeds. The short generation times and line- or inbreeding to maintain breed standards lead to a higher rate of emergence of heritable recessive diseases compared to humans. Spontaneous mutants in purebred populations with such unique structure make the dog a valuable model for human disorders [11].

In addition to physiological and clinical similarity of canine and human conditions, pet dogs also share a common environment with their owners, making them a highly valuable large animal model [12]. The freely available database Online Mendelian Inheritance in Man (OMIM) contains comprehensive information on human genes, and genetic traits and disorders with a focus on the genotype-phenotype relationship and currently shows over 25,000 entries [13]. Spontaneously occurring canine common or rare diseases are highly similar to their human counterparts and are often caused by variants in genes that are orthologous to the human genes [12]. The public catalog Online Mendelian Inheritance in Animals (OMIA) currently records a total of 787 canine traits or disorders, of which 474 are classified as potential models for human disease with 245 records having gene associations [14]. Kennel clubs of purebred breeds usually keep detailed pedigree records, and often also additional health information for each privately-owned dog. Active participation of breeding organizations and individual owners and veterinarians in research is essential for discovery of affected dogs, and subsequent identification of deleterious variants associated with their disease. Many purebred dogs appear to have an increased risk of certain disorders suggesting a genetic predisposition [15].

The size of the dog genome is estimated to be 2.4 Gb and is organized in 38 pairs of autosomes and two sex chromosomes. Similarly to the number in human, ~22,000 protein-coding genes are estimated to be present in the canine genome [16]. Since the whole sequence of the dog genome became available in 2005 [17], advanced sequencing technology and improved bioinformatics methods have eased the genetic discovery, allowing identification of more than 400 likely causal variants associated with diseases and traits in the domestic dog [14]. Even despite large population sizes of some popular dog breeds, their genetic diversity remains limited. This fact makes them a powerful tool to study genetic etiology of not only simple Mendelian, but also of more complex heterogeneous disorders, such as cancers or neurodegenerative diseases [18].

Similar to human clinical genetics [19], DNA-based precision medicine using available genomic information has been proven suitable for diagnosing rare diseases in veterinary medicine [20] due to the unprecedentedly high level of medical care standards available for companion animals. Finally, the discovery of causative variants enables the development of genetic tests,

which are used by dog breeders worldwide to stop the spread of inherited diseases and improve the overall health of the respective breeds, while maintaining what genetic diversity exists within the breed. Additionally, the molecular characterization of rare diseases in these affected animals might help in explaining unsolved cases of similar human diseases.

Inherited neurological diseases

Neurodegenerative disorders represent a heterogeneous group of diseases affecting the central (CNS) and/or peripheral nervous system (PNS) with the most prominently affected cell type being the neuron (Figure 1).

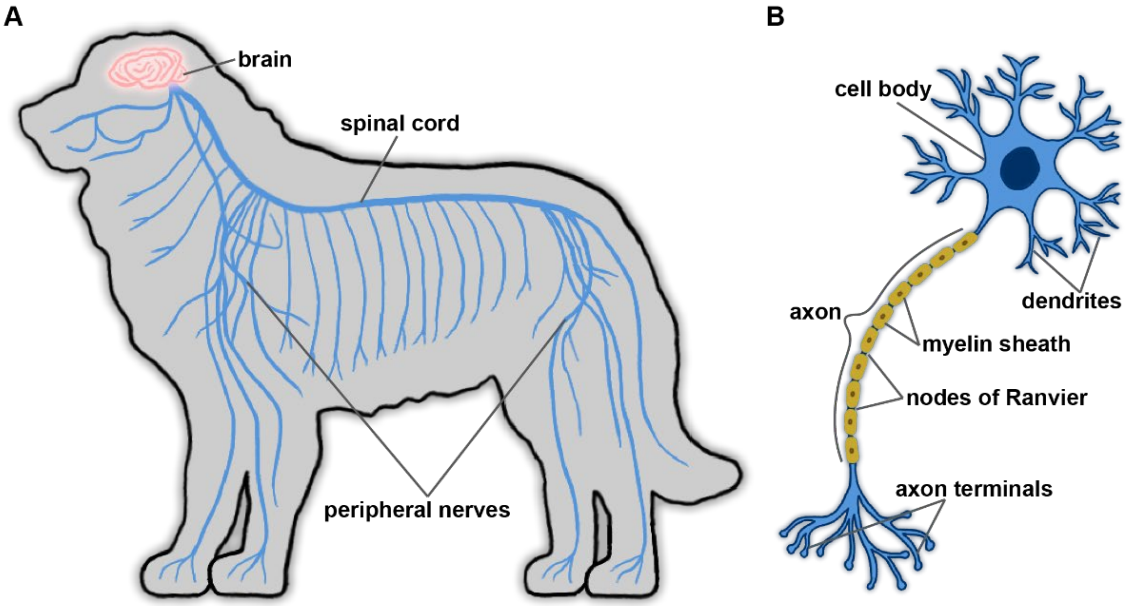

Figure 1. Schematic representation of canine nervous system. **(A)** Central nervous system consists of brain and spinal cord, while peripheral nervous system represents the nerves connecting rest of the body to the CNS. **(B)** Structure of a typical neuron with myelinated axon.

Numerous neurological diseases have been described in humans, as well as in animals, and their molecular characterization improved in recent decades [21]. Diseases of the CNS include myelin disorders, encephalopathies, cerebellar degeneration, neuroaxonal dystrophy, and ataxias [22]. Peripheral neuropathies, disorders of the PNS, are one of the most common neurogenetic diseases, and can be distinguished depending on which part of the nerve is affected [23]. All these disease groups may overlap with the classification of neurometabolic disorder when a metabolic pathway is affected [24].

As in human medicine, neuropathic disorders in dogs often lead to severe phenotypes that negatively impact the quality of life and result in progressive lifelong disabilities, or even premature death because of the lack of effective treatment [25]. While some of these disorders are more common but may present differently (e.g. variable progression of clinical signs and ages of onset) across breeds, others are rare and occur in breeds that appear predisposed to

particular forms of these diseases, with many conditions showing complete breed specificity [15]. This is most likely due to the limited genetic diversity and high relatedness levels and inbreeding within the modern breeds. State-of-the-art genomic tools efficiently characterize the DNA variants and may uncover novel molecular pathways involved in such diseases, highlighting the advantages of the dog model [26].

Neuropathies

Hereditary peripheral neuropathies (OMIM PS118220, PS162400) are a large group of neurological diseases characterized by length-dependent progressive degeneration of the PNS with considerable phenotypic and genetic heterogeneity [26]. Charcot-Marie-Tooth disease (CMT) is globally the most common form of inherited neuropathy in humans with an estimated prevalence of 1:2,500. It is classified as hereditary motor and sensory neuropathy, describing its main symptoms: progressive muscle weakness and sensory loss. Additionally, patients often develop foot deformities, musculoskeletal complications (e.g. hip dysplasia, scoliosis), and some forms also affect cranial nerves and respiratory function [27].

In addition to CMT, hereditary neuropathy with liability to pressure palsies, hereditary motor neuropathies, and hereditary sensory and autonomic neuropathies are recognized and may be complicated by additional neurologic and non-neurologic features. Most of these subgroups show genetic heterogeneity, with a known association of around 100 genes, which is also reflected in a variety of inheritance patterns and involvement of diverse molecular pathomechanisms [28]. More than 30 genes have been implicated in CMT alone. However, only four genes (*PMP22*, *MPZ*, *GJB1*, and *MFN2*) explain over 90% of genetically confirmed cases of CMT [29,30]. The rate of detection of causal variants varies greatly, from as low as 10% in hereditary motor neuropathies to up to 80% in demyelinating forms of CMT. The use of modern sequencing technologies is becoming the standard in precision medicine for the evaluation of known neuropathy genes and the detection of novel variants in patients with unexplained pathology [28].

Many canine inherited peripheral polyneuropathies with or without laryngeal paralysis (LPPN) (OMIA 001292, 001206) show similarities to human CMT [31]. Laryngeal paralysis may develop as part of a polyneuropathy complex and occurs when normally mobile laryngeal cartilages become immobile and dysfunctional due to damage to the recurrent laryngeal nerve. This leads to respiratory obstruction, noisy and distressed breathing, as well as reduced exercise and heat tolerance [32]. It is often treated surgically by unilateral arytenoid lateralization surgery to improve breathing and, therefore, quality of life. However, it may increase the risk of lifelong complications, such as aspiration pneumonia [33]. Different forms of motor and sensory neuropathies often have comparable clinical and histopathological features, including gait abnormalities, muscle atrophy, laryngeal paralysis, bark changes,

excessive panting or swallowing difficulties, and severe nerve fiber loss resulting from chronic axonal degeneration and myelin sheath degradation [34]. Sensory and autonomic neuropathies are less frequent but also described in dogs that often exhibit automutilation syndrome with insensitivity to painful stimuli [26]. Currently, variants in nine genes are recorded in OMIA [14] for various forms of neuropathy in several dog breeds (Table 1).

Table 1. A list of breeds representing selected forms of canine neuropathies (OMIA#), for which the genes harboring candidate causative variants are described, and human disorders (OMIM#) caused by variants in the corresponding orthologs, if known.

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Polyneuropathy and/or laryngeal paralysis						
Leonberger Saint Bernard	?	*ARHGEF10*	SS	001917	Slowed nerve conduction velocity	608136
Leonberger Labrador retriever Saint Bernard	AR	*CNTNAP1*	MS	002301	Congenital hypomyelinating neuropathy Lethal congenital contracture syndrome	602346
Leonberger	AD	*GJA9*	FS	002119	unknown	611923
Alaskan malamute Greyhound	AR	*NDRG1*	MS FS	002120	CMT	605262
Alaskan husky Black Russian terrier Rottweiler	AR	*RAB3GAP1*	SS FS FS	001970	Warburg micro syndrome	602536
Miniature bull terrier	MF	*RAPGEF6*	SS	002222	unknown	610499
Miniature schnauzer	AR	*SBF2*	SS	002284	CMT	607697
Sensory neuropathy						
Border collie mixed breed	AR	*FAM134B*	SS	002032	Hereditary sensory and autonomic neuropathy	613114
Golden retriever	MT	*tRNA-Tyr*	FS	001467	Exercise intolerance	590100

[1] MOI = mode of inheritance as follows: AR = autosomal recessive, AD = autosomal dominant, MT = mitochondrial, MF = multifactorial, ? = unclear.
[2] Type of variant as follows: MS = missense, FS = frameshift, SS = splice-site variant. Multiple entries correspond to different variant types found in respective breeds.

Myelin disorders

Myelin sheaths are formed by neuroglia and cover the axon of many neurons in the CNS as well as PNS. They differ in origin, but share similar structure and function, i.e. to help increase the speed of axonal transmission of neural impulses. The nodes of Ranvier form short segments of the axon not covered by myelin (Figure 1), where ion channels are located and are critical to the normal function [35]. Extensive damage and loss of myelin are followed by degeneration of the axon and often also the cell body; this interrupts nerve transmission and results in forms of polyneuropathy in PNS or white matter diseases in CNS [36]. Primary myelin

disorders of CNS are caused by defects in myelin formation and/or its maintenance and may be grouped into dysmyelinating, hypomyelinating, and spongiform disorders [37].

In human patients, heritable white matter diseases, or leukodystrophies (OMIM PS312080), are classified based on neuroimaging and exhibit a wide range of progressive clinical signs including ataxia, spasticity, dysmetria, or developmental delay. More than 30 distinct disorders meet the defined criteria of the consensus definition of leukodystrophy, while more than 60 other disorders are classified as genetic leukoencephalopathies as a broader group with heterogeneous clinical and pathological manifestations [38]. In some instances, because of the close connection between neuroglia and neurons/axons, it may be difficult to differentiate between a process that primarily affects the myelin (strict leukodystrophy) and a process that primarily affects the grey matter (genetic leukoencephalopathy), especially when the pathology is unclear and performed at the advanced stage of the disease [39]. Recent studies show that almost half of all patients whose clinical manifestations and neuroimaging indicate an inherited white matter disease do not receive a specific diagnosis with pathogenic variant, even when genetic screening of more than 100 leukoencephalopathy-related genes is performed [40,41].

Similar myelin disorders are also known in animal species, including dogs [22,42,43]. Canine forms of this rare group of disorders include several types of myelopathy with lysis of the white matter (Table 2). The age of onset, progression, and severity varies greatly between the different myelin disorders [36,44]. For example, degenerative myelopathy is most common in older dogs who show axonal and myelin degeneration in the spinal cord, as well as hypomyelination and secondary demyelination in peripheral nerves. The clinical signs include hyporeflexia, slowly progressive paresis, and spasticity in the pelvic limbs, with general proprioceptive ataxia [45]. Leukodystrophies and leukoencephalomyelopathies are examples of infantile- or juvenile-onset neurodegenerative disorders characterized by progressive neurological worsening, gait abnormalities, especially in the thoracic limbs, and generalized weakness, spasticity, ataxia, and loss of conscious proprioception. Affected dogs usually have characteristic lesions in the spinal cord white matter, while peripheral nerve and muscle biopsies remain generally unremarkable [46,47]. The historical reports of individual cases and lack of precise classification reflect the heterogeneity of such disorders termed as leukodystrophy, leukoencephalomyelopathy, necrotizing myelopathy, hypomyelination of the CNS, or degenerative myelopathy. The mode of inheritance is mostly autosomal recessive, but a mitochondrial form has also been described (Table 2).

Table 2. A list of breeds representing selected forms of canine myelin disorders (OMIA#), for which the genes harboring candidate causative variants are described, and human disorders (OMIM#) caused by variants in the corresponding orthologs, if known.

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Leukodystrophy						
Standard schnauzer	AR	*TSEN54*	MS	002215	Pontocerebellar hypoplasia	608755
Cairn terrier Irish setter West Highland white terrier	AR	*GALC*	MS IFI MS	000578	Krabbe disease	606890
Hypomyelination of the central nervous system						
Weimaraner	AR	*FNIP2*	FS	000526	unknown	612768
Leukoencephalomyelopathy						
Australian cattle dog Shetland sheepdog	MT	*CYTB*	MS	001130	Encephalopathy Leber hereditary optic neuropathy	516020
Great Dane Leonberger Rottweiler	AR	*NAPEPLD*	FS MS FS	001788	unknown	612334
Degenerative myelopathy						
Bernese mountain dog Boxer Chesapeake Bay retriever German shepherd Hovawart Pembroke Welsh corgi Rhodesian ridgeback	AR	*SOD1* (*SP110*)	MS	000263	Amyotrophic lateral sclerosis Spastic tetraplegia and axial hypotonia	147450

[1] MOI = mode of inheritance as follows: AR = autosomal recessive, MT = mitochondrial.

[2] Type of variant as follows: MS = missense, FS = frameshift, IFI = in-frame insertion. Multiple entries correspond to different variant types found in respective breeds.

Neuroaxonal dystrophies

Neuroaxonal dystrophy (NAD) represents a clinically and genetically heterogeneous group of rare diseases of CNS development, with motor neuron degeneration characterized by localized axonal swelling and secondary myelin changes [48]. NAD in humans (OMIM PS234200) are mostly infantile or juvenile diseases diagnosed histopathologically by the formation of dystrophic axons seen as large spheroids mostly in the grey matter in the brain and spinal cord [49]. The clinical signs include progressive ataxia, paraplegia, hypermetria, proprioceptive deficits, head incoordination, and tremors. Due to neuromuscular dysfunction leading to joint contracture and respiratory failure, this condition is often lethal [50]. Nine forms of NAD are described in humans and termed neurodegeneration with brain iron accumulation (NBIA) or infantile neuroaxonal dystrophy (INAD) following mostly autosomal recessive, but also autosomal dominant and X-linked dominant inheritance. The genetic basis of NBIA/INAD is

known in about 85% of patients and involves variants in eight genes (*PANK2*, *PLA2G6*, *FTL*, *C19orf12*, *WDR45*, *COASY*, *REPS1*, and *CRAT*) [48].

In veterinary medicine, NAD has been described previously in numerous species (OMIA 000715). However, the underlying genetic etiology is currently known only in sheep [51] and dogs [52–54]. The affected dogs present typical early onset of signs such as hypermetria of thoracic limbs, nystagmus, ataxia, behavioral deficits, and progressive limb paralysis. Although some species differences are described, such as lack of brain iron accumulation in dogs compared to humans, common pathological pathways are involved in the disease [22]. Three of the four canine breed-specific forms are associated with genes in which variants are found associated with different human disorders such as neuropathies, paraplegia, or leukodystrophy (Table 3).

Table 3. A list of breeds representing selected forms of canine neuroaxonal dystrophies (OMIA#), for which the genes harboring candidate causative variants are described, and human disorders (OMIM#) caused by variants in the corresponding orthologs, if known.

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Neuroaxonal dystrophy						
Schnauzer-Beagle cross	AR	*MFN2*	IFD	002153	CMT Hereditary motor and sensory neuropathy	608507
Papillon	AR	*PLA2G6*	MS	002105	INAD NBIA Parkinson disease	603604
Spanish water dog	AR	*TECPR2*	MS	001975	Spastic paraplegia	615000
Rottweiler	AR	*VPS11*	MS	002152	Hypomyelinating leukodystrophy	608549

[1] MOI = mode of inheritance as follows: AR = autosomal recessive.
[2] Type of variant as follows: MS = missense, IFD = in-frame deletion.

Spinocerebellar ataxias

Ataxia is a relatively non-specific clinical sign of uncoordinated movements. This might lead to confusion and therefore, standardized terminology has been proposed to distinguish the complex spinocerebellar degenerative disorders in human medicine. This group includes hereditary cerebellar ataxias (HCA) and hereditary spastic ataxias (SPAX) mostly subdivided according to the mode of inheritance and the mutated gene, although some conditions may overlap [21]. More than 70 forms of autosomal dominant (OMIM PS164400) and autosomal recessive (OMIM PS213200) spinocerebellar ataxias are described to date, with more than 60 implicated genes. The definitions are even more complicated by the obvious pleiotropy when a mutation in the same gene may induce various phenotypes [55]. HCA are associated with cerebellar degeneration, and the disease signs include imbalance, poor hand-eye coordination, visual loss, seizures, behavioral symptoms, and peripheral neuropathy [56].

In dogs, phenotypically similar neurodegenerative disorders are described as either cerebellar or spinocerebellar ataxia in several breeds. The classification is currently based mainly on clinicopathological features and lacks clear consensus, which might lead to misdiagnosis. Due to the cerebellar dysfunction, the affected dogs typically show rapidly progressing generalized ataxia, tremors, and general failure to thrive [57]. So far, causative variants in twelve genes have been described, clearly showing the locus heterogeneity of this phenotype (Table 4).

Table 4. A list of breeds representing selected forms of canine hereditary ataxias (OMIA#), for which the genes harboring candidate causative variants are described, and human disorders (OMIM#) caused by variants in the corresponding orthologs, if known.

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Spinocerebellar ataxia						
Parson Russell terrier	AR	*CAPN1*	MS	001820	Spastic paraplegia	114220
Italian spinone	AR	*ITPR1*	IRE	002097	Gillespie syndrome Spinocerebellar ataxia	147265
Alpine dachsbracke	AR	*SCN8A*	MS	002194	Myoclonus Cognitive impairment with or without cerebellar ataxia Developmental and epileptic encephalopathy Benign seizures	600702
Belgian shepherd	AR	*SLC12A6*	FS	002279	Agenesis of the corpus callosum with peripheral neuropathy	604878
Beagle	AR	*SPTBN2*	FS	002092	Spinocerebellar ataxia	604985
Cerebellar ataxia						
Belgian shepherd	AR	*ATP1B2*	SS	002110	unknown	182331
Coton de Tulear	AR	*GRM1*	SS	000078	Spinocerebellar ataxia	604473
Norwegian buhund	AR	*KCNIP4*	MS	002240	unknown	608182
Belgian shepherd Jack Russell terrier Parson Russell terrier Fox terrier (smooth)	AR	*KCNJ10*	MS MS/FS MS MS	002089	SESAME syndrome, enlarged vestibular aqueduct	602208
Gordon setter Old English sheepdog	AR	*RAB24*	MS	001913	unknown	612415
Finnish hound	AR	*SEL1L*	MS	001692	unknown	602329
Hungarian vizsla	AR	*SNX14*	SS	002034	Spinocerebellar ataxia	616105

[1] MOI = mode of inheritance as follows: AR = autosomal recessive.
[2] Type of variant as follows: MS = missense, FS = frameshift, SS = splice-site variant, IRE = intronic repeat expansion. Multiple entries correspond to different variant types found in respective breeds.

Neurometabolic disorders

Neurometabolic disorders, including various forms of storage diseases, are inborn errors of metabolism (IEM), a group of diseases caused by an enzyme deficiency in a metabolic pathway. Individually these diseases are rare, but as a group, they are relatively common, with more than 500 IEMs reported in human medicine. Accumulation of toxic compounds due to

defects in intermediary metabolic pathways, energy deficiency related to mitochondrial respiratory chain defects, and defects in the synthesis of complex molecules in certain cellular organelles are the main disease-causing mechanisms [58]. About a quarter of IEMs manifests in the first few weeks of age, but the clinical signs may also present later in childhood or even much later in adulthood, depending on the enzyme deficiency and involved pathway. IEMs are challenging to diagnose because the signs are usually non-specific, and may include decreased activity, movement abnormalities, mental retardation, developmental delay, respiratory distress, lethargy, or seizures [59]. Genes involved in IEMs encode enzymes, cofactors, or transmembrane transporters. The absence or abnormal function of these proteins results in deficiency or accumulation of specific metabolites, which may lead to lysosomal storage disorders. Most neurometabolic disorders are autosomal recessive, but autosomal dominant and X-linked forms are also known. Furthermore, different mutations in the same gene may result in a wide phenotypic spectrum of diseases [24]. Due to the progressive neurologic deterioration involved in many IEM, the classification of some disorders overlaps with other grouping systems, such as several neuropathies, encephalopathies, spinocerebellar ataxias, NAD, as well as myelin disorders [60].

Various canine neurodegenerative conditions resemble the human neurometabolic disorders, e.g. thiamine-responsive dysfunction syndrome (OMIM PS249270), succinic semialdehyde dehydrogenase deficiency (OMIM 271980), mucopolysaccharidosis (OMIM PS607014), or neuronal ceroid lipofuscinosis (OMIM PS256730). These disorders generally show considerable phenotypic and genetic heterogeneity with similar health problems as in human patients (Table 5). Therefore, studying these naturally occurring dog models furthers the understanding of such a complex group of disorders and the biochemical and molecular pathogenesis [61].

Table 5. A list of breeds representing selected forms of canine neurometabolic disorders (OMIA#), for which the genes harboring candidate causative variants are described, and human disorders (OMIM#) caused by variants in the corresponding orthologs, if known.

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Leigh-like subacute necrotizing encephalopathy						
Alaskan husky Yorkshire terrier	AR	*SLC19A3*	FS	001097	Thiamine metabolism dysfunction syndrome	606152
Succinic semialdehyde dehydrogenase deficiency						
Saluki	AR	*ALDH5A1*	MS	002250	Succinic semialdehyde dehydrogenase deficiency	610045
Neurodegenerative vacuolar storage disease						
Lagotto Romagnolo	AR	*ATG4D*	MS	001954	unknown	611340

continued

Breed	MOI[1]	Associated gene	Variant[2]	OMIA# [14]	Human diseases	OMIM# [13]
Mucopolysaccharidosis						
Great Dane Miniature pinscher Miniature poodle Miniature schnauzer	AR	*ARSB*	NS MS FS SL	000666	Mucopolysaccharidosis	611542
Brazilian terrier German shepherd	AR	*GUSB*	MS	000667	Mucopolysaccharidosis	611499
Boston terrier Golden retriever Plott hound	AR	*IDUA*	FS IFD SS	000664	Mucopolysaccharidosis	252800
Schipperke	AR	*NAGLU*	FS	001342	CMT Mucopolysaccharidosis	609701
Dachshund Huntaway dog	AR	*SGSH*	IFD FS	001309	Mucopolysaccharidosis	605270
Neuronal ceroid lipofuscinosis (NCL)						
Australian cattle dog Tibetan terrier	AR	*ATP13A2*	MS SS	001552	Kufor-Rakeb syndrome Spastic paraplegia	610513
Australian cattle dog Border collie Golden retriever	AR	*CLN5*	NS NS FS	001482	NCL	608102
Australian shepherd	AR	*CLN6*	MS	001443	NCL	606725
Alpine dachsbracke Australian shepherd English setter German shorthaired pointer Saluki	AR	*CLN8*	EGD NS MS NS NS	001506	NCL	607837
American bulldog	AR	*CTSD*	MS	001505	NCL	116840
Chihuahua Chinese crested dog	AR	*MFSD8*	FS	001962	NCL Macular dystrophy with central cone involvement	611124
Dachshund Italian cane corso	AR	*PPT1*	FS SS	001504	NCL	600722
Dachshund	AR	*TPP1*	FS	001472	NCL Spinocerebellar ataxia	607998

[1] MOI = mode of inheritance as follows AR = autosomal recessive.

[2] Type of variant as follows MS = missense, NS = nonsense, FS = frameshift, SS = splice-site variant, SL = loss of the canonical start codon, EGD = deletion of the entire gene, IFD = in-frame deletion. Multiple entries correspond to different variant types found in respective breeds.

Selected methods in genomic analyses

In the late 20th century, the applications of DNA-based diagnostics in animals included cytogenetic and chromosomal techniques for individual identification, parentage control, and discovery of deleterious alleles in inherited disorders [62]. Biological knowledge of the specific disorder and its underlying etiology was a prerequisite for identifying the causal variant in candidate gene approaches. In the 1990s, the first canine genetic linkage map was developed by typing 150 microsatellite markers as a hypothesis-free approach for mapping trait-associated loci [63].

More recently, species-specific single nucleotide polymorphism (SNP) arrays based on simultaneously genotyping thousands of markers with known chromosomal positions have become a routine tool in the field of genetics [64]. This enabled rapid mapping of various traits by methods of linkage [65], homozygosity [66], and genome-wide association studies (GWAS) [67,68]. Lastly, major improvements in molecular biology and computer science in the 21st century allowed the whole-genome sequencing (WGS) technologies to become a powerful routine tool in personalized and precision medicine for humans, as well as companion animals, such as dogs [69,70].

The types of canine variants implicated in disease range from single-nucleotide variants (i.e. substitutions, insertions, or deletions of 1 bp) [71–74] through small indels (i.e. insertions and deletions of <50 bp) [75,76] to complex large structural variants (i.e. genomic rearrangements and copy number variation) [77–79], and may be efficiently identified with SNP array genotyping and/or WGS of individuals, families, or large populations.

In this thesis, a combination of relevant methods has been utilized for mapping, detection, and validation of putative disease-causing genetic variants based on each project's hypothesis that slightly differed due to observed family structure and available samples.

Single nucleotide polymorphism (SNP) genotyping arrays

Canine geneticists were among the first to drive the production of high-quality SNP genotyping arrays for companion animals. The polymorphic SNPs were selected from a 2.5 million SNP map generated as part of the dog genome project [17] and allowed several large-scale SNP genotyping platforms to emerge. These platforms enabled the study of the population structure within and across different dog breeds, as well as genetic variation associated with common and rare diseases relevant to humans. The earliest array versions established around 2007 contained as low as 1,500 SNP markers, and later extended to 22k, 27k, or 49k by different manufactures. In 2011, a high-density SNP array with 172k SNPs was developed in collaboration with the LUPA consortium [80]. The most common currently used versions have 220k (CanineHD BeadChip, Illumina, San Diego, CA, USA) or 710k markers (Axiom CanineHD

Array, Thermo Fisher Scientific, Waltham, MA, USA). In addition, high-quality imputation from a lower to a higher level of SNP array density or even to a whole-genome sequencing density is possible and has been shown successfully in dogs [81,82].

Linkage analysis

Genetic linkage analysis is a mapping approach used to examine the segregation of a particular trait in a given pedigree and to identify chromosomal regions harboring the gene responsible for the studied trait. Linkage and recombination are key concepts in the parameter-based and parameter-free linkage analysis [83]. Genetic linkage occurs when two loci are inherited together due to their proximity more often than expected by chance. Two loci separated by one centimorgan (corresponding to approximately one million base pairs) have a chance of 1% to be separated during a recombination event in one meiosis. With each separation, non-random association of alleles of two loci (the linkage disequilibrium) decreases and only persists over generations if the two loci are physically close to each other [84]. In parameter-based linkage analysis, the co-segregation of genetic markers and a trait is studied under a specific model based on known information about mode of inheritance and penetrance. Parameter-free linkage analysis studies the probability of alleles being identical by descent and does not require a mode of inheritance to be specified [83].

The power of linkage between the trait and marker loci is commonly assessed by the LOD (logarithm of the odds) score. Generally, a LOD score of higher than 3 is understood as genome-wide significant evidence for linkage, while a LOD score of lower than -2 excludes linkage. However, the genome-wide significance threshold is of less concern when using whole-genome sequencing data as it allows for relatively easy study of even multiple regions with only suggestive LOD scores obtained in small pedigrees [84].

Runs of homozygosity

Homozygosity, also known as autozygosity, mapping is a powerful method based on the hypothesis that an individual inherited two copies of an allele from a common ancestor. In dogs, inbreeding and high relatedness within breeds contribute to a high number of long genomic regions that are homozygous [85]. Subsequently, this method is most useful for detecting the disease-associated alleles in monogenic autosomal recessive disorders. The putative disease locus is typically flanked by a few kilobases (kb) to several megabases (Mb) of identical-by-descent markers because chromosomal regions tend to be transmitted together. Both SNP array genotyping or WGS data can be investigated for such extended regions of homozygosity, also known as runs of homozygosity (ROH) [86]. Rarely, a disease may be caused by compound heterozygous variants in a consanguineous pedigree, in which case the homozygosity mapping approach will not yield the expected results [84]. In addition

to the detection of deleterious variants shared by affected individuals, the characterization of ROHs is utilized in population genetics. ROH's length decreases with genetic distance due to recombination events and thus may be used as an indication of recent or ancient inbreeding in the breed's population history. Genomic inbreeding coefficients estimated from ROHs have been shown to be more reliable and precise than pedigree-based coefficients, especially in case of incomplete, unreliable, or missing pedigree information [87].

Genome-wide association study (GWAS)

In case-control studies, GWAS is the common method for identifying candidate genomic regions statistically associated with the studied trait or disease by comparison of allele frequency between unrelated affected and unaffected individuals. Most often this approach uses the SNP array genotyping data with thousands of polymorphic markers. Therefore, the SNP significantly associated with the investigated trait is not causative but considered to be in linkage disequilibrium with the true causal variant [88]. The results of GWAS are often displayed as a Manhattan plot with chromosome numbers on the X-axis and the negative logarithm of the association p-value for each SNP on the Y-axis (Figure 2) [89].

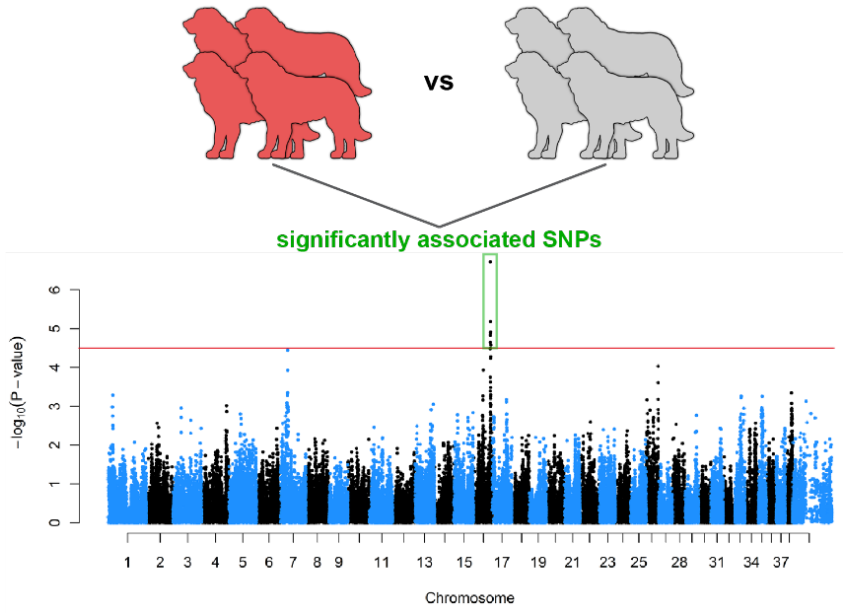

Figure 2. Illustration of GWAS used to detect disease-associated SNPs by comparison of two groups of dogs: cases (red) and controls (grey). The red line represents the Bonferroni correction level used to identify the genome-wide significantly associated SNPs (highlighted in green box). The Manhattan plot of GWAS in *ARHGEF10*-related polyneuropathy was reproduced with permission from [89].

The power of GWAS to detect a true association depends on various factors, such as well-defined phenotype, sample size, or data structure. Mode of inheritance determines the minimum number of cases and controls needed for the GWAS. About 10-20 dogs per group should be sufficient for mapping a simple recessive trait with high penetrance, while for a dominant trait at least 50 cases and 50 controls are needed. Much larger sample sizes are

required for complex or polygenic traits [64]. Spurious associations might arise due to population stratification, family structure, and cryptic relatedness. These are commonly corrected for by the use of principal component analyses, multidimensional scaling methods, or mixed models including pair-wise genetic relationship matrices [90]. Additionally, the results have to be adjusted for multiple testing in order to avoid false-positive significant associations because of the thousands, potentially millions, of markers used. To set a meaningful significance threshold, Bonferroni correction, permutation testing, or false discovery rate methods are commonly implemented [88].

Whole-genome sequencing (WGS)

Sequencing of an individual's whole genome has become affordable and the main genetic technique in human, as well as veterinary, precision medicine [19,91]. Since the development of the Sanger sequencing method (the first generation) in the 1970s, the technology for whole-genome sequencing has evolved and been perfected in all aspects. Around the year 2000, the second-generation, also known as next-generation sequencing (NGS) methods emerged allowing the entire genome to be sequenced at once. Later around 2008, the third-generation sequencing methods further significantly improved the throughput, accuracy, time and cost efficiency of WGS [69]. A major challenge that these technologies bring is the production of a large amount of data that requires considerable computational resources and advanced bioinformatics tools for storage, handling, and downstream analyses [92].

The different NGS platforms vary in methodology, throughput, speed, and read lengths. Illumina's platform is the current leader in short-read sequencing using massively parallel sequencing by synthesis with optical base calling, which generates single or paired-end reads of ~150 bp on average. A major disadvantage of this technology is the inability to reliably resolve repetitive regions and detect larger structural variants in the genome. Long-read technologies, such as single-molecule real-time sequencing from Pacific Biosciences or nanopore sequencing from Oxford Nanopore Technologies, strive to overcome these challenges. Both platforms produce long to ultra-long reads (>1 kb to 2 Mb) enabling not only the analysis of structural variants but also high-quality *de novo* genome assemblies [93].

Although these techniques revolutionized genomic research and enabled the use of WGS for diagnostic purposes, Sanger sequencing is still used for targeted sequencing of candidate genes and validation of NGS output due to its highly accurate results of up to one kb long fragments for a low price [69].

Discovery of private, shared, rare, or breed-specific variants

Using the WGS data, it is possible to identify putative causal variants even without previous mapping strategies in cases where only a single or a limited number of affected animals are available. Short-read sequencing data may be analyzed to reveal a large portion of the genetic variation comprising single nucleotide variants (SNV) and small indels. The number of identified SNVs largely exceeds the number of genotyping array markers [20,94]. This extensive SNV dataset represents a part of an individual's normal variation. However, it may also contain the causative genetic variant underlying the inherited disease the individual suffers from. The pathogenic variants may be identified by comparing different individual datasets with each other. Especially for rare monogenic diseases, this approach is highly effective and has led to a rapid increase in the number of identified causative variants [94].

In order to find such variants, the reads obtained from WGS have to be filtered, deduplicated, and aligned to a reference genome sequence. Subsequently, the SNVs and small indels are called, filtered for quality, and annotated for prediction of each variant's impact on the protein level [95,96]. This approach is usually applied to a set of case and control animals and generates millions of variants across the genome. Therefore, hierarchical filtering strategies are used to reduce the number of variants. Figure 3 illustrates the most commonly used filtering options when using a large set of variants detected from WGS data.

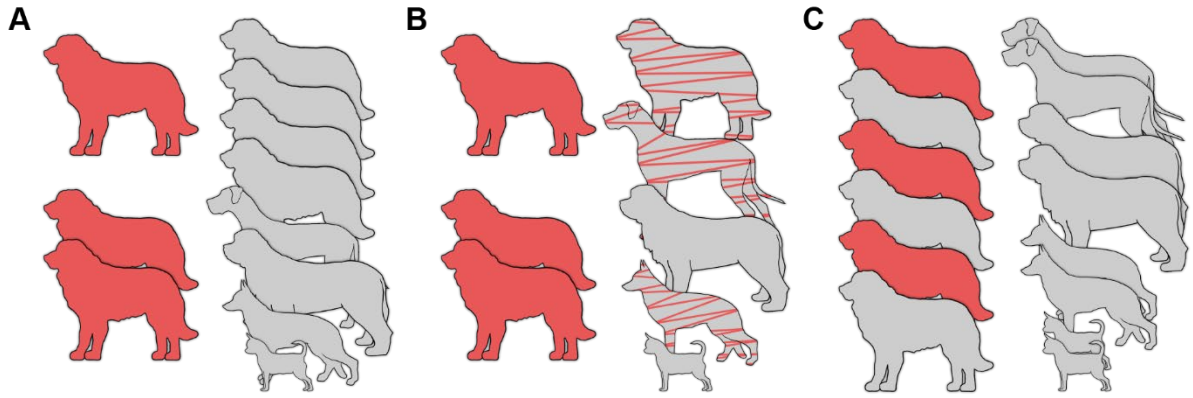

Figure 3. Illustration of different strategies used when filtering variants detected from a large set of whole-genome sequences. Red- and grey-filled dogs represent cases and controls, respectively. The striped dogs represent control dogs that might also carry the variants. **(A)** Detection of private variants unique to the case or shared by multiple cases, and absent from controls. **(B)** Detection of rare or enriched variants present in one or more cases as well as a subset of controls. **(C)** Detection of breed-specific variants present in all animals of one breed and absent from all other breeds.

Usually, one or more cases (affected animals) are compared to controls, which may be represented either by known unaffected animals for the given phenotype or by population controls even with unknown health status of the same breed, or by all other genomes including different breeds (Figure 3A). In the case of rare and enriched variants, the alternative allele frequency in all genomes can be calculated to detect disease-associated variants that are

potentially shared by several breed controls due to e.g. reduced penetrance or carrier status in recessive traits, or animals from other related breeds due to their shared history (Figure 3B). For detecting fixed breed-specific variants, all animals from one breed regardless of their phenotype are compared to all animals of various other breeds (Figure 3C). To further narrow down the list of trait-associated variants, filtering based on the mode of inheritance, especially in trio-based datasets, on a known list of candidate genes, in the region of interest from previous mapping approaches, or based on estimated functional impacts, is often performed.

Estimation of functional impact remains challenging as genetic variants may have different effects from protein-changing variants altering gene function to non-coding variants with no clear impact. Apart from larger structural variants such as deletions of several Mb encompassing the entire gene, the SNVs easiest to detect and assess include start/stop codon alterations and frameshifts, followed by non-synonymous amino acid changes. However, the majority of the variant dataset consists of non-coding variants. The evaluation of the potential impact of synonymous or non-coding SNVs requires an advanced experimental setup [97].

In dogs, hundreds of whole genomes are already publically available and further efforts are underway to characterize and catalog the existing variation across canine populations, including ancient and modern dog breeds, village dogs, wolves, and other wild canids [20,98,99]. Despite the several described challenges, this kind of easily accessible resource makes the discovery of deleterious variants in rare canine neurogenetic diseases possible.

Dog reference assembly

In the mid-2000s, two consecutive assemblies were released (the initial CanFam1.0 and the CamFam2.0) and represented the first draft of the dog genome sequence published by the Dog Genome Sequencing Consortium at the Broad Institute [100] (Figure 4). This assembly spanned 2.41 Gb representing 99% of a female Boxer genome, but still contained many gaps, and was subsequently improved to the latest version CanFam3.1 [101]. Until recently, this genome build served as the annotated reference sequence for canine genomic studies, and is also used as such throughout my thesis.

In the last years, new developments and better accessibility of NGS methods including long-read technologies, together with ever-decreasing costs, resulted in superior and breed-specific *de novo* genome assemblies. Currently, five different breed reference sequences with annotations are available for the dog in the National Center for Biotechnology Information (NCBI) [102].

This NCBI list includes an improved version of the original Boxer assembly (Dog10K_Boxer_Tasha, GCF_000002285.5) [103], and new assemblies from a Great Dane (UMICH_Zoey_3.1, GCF_005444595.1) [104], a German Shepherd (UU_Cfam_GSD_1.0, GCF_011100685.1) [105], a Basenji (UNSW_CanFamBas_1.0, GCF_013276365.1) [106], and a Labrador retriever (ROS_Cfam_1.0, GCF_014441545.1) [107] (Figure 4).

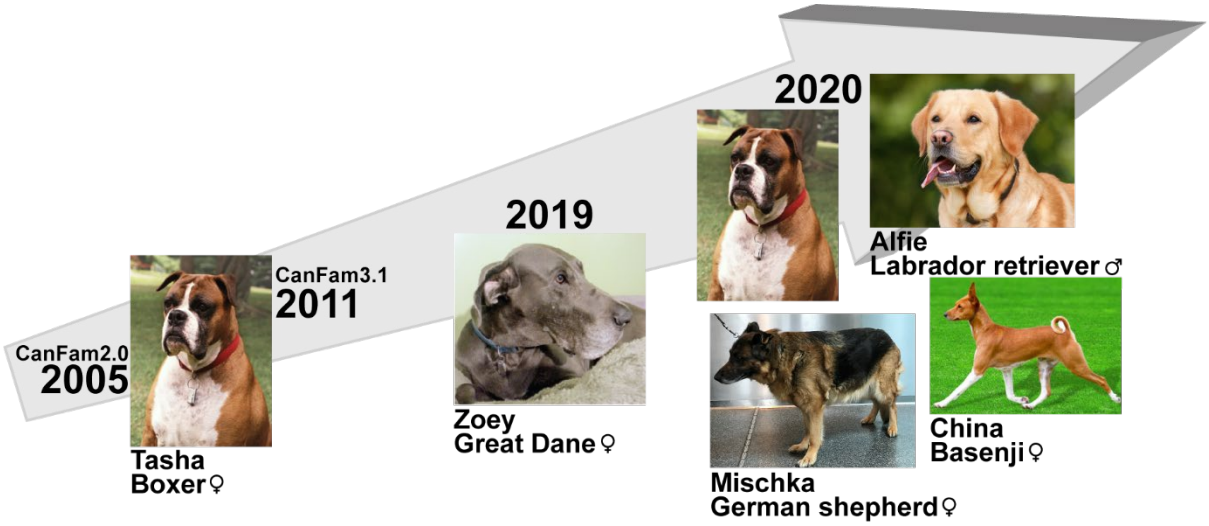

Figure 4. An overview of the development of canine annotated reference genomes over time highlighting the obvious breed differences. Photos adapted with permission from [100,104–107].

The latest and still on-going development of multiple high-quality assemblies and annotations, representing distinct breeds (Figure 4) will enhance the accuracy of future genomic diversity and disease analyses.

Aim and hypothesis of the thesis

The main aim of this thesis was to characterize the genetic etiology of rare forms of canine neurological diseases as potential models for similar human diseases. The general hypothesis assumed simple Mendelian inheritance and a breed-specific underlying variant. The results were expected to improve the health status of respective dog breeds by selective breeding against the unraveled deleterious variants, as well as to provide spontaneous animal models and possible new candidate genes for similar human inherited rare neurological disorders with yet unknown genetic causes.

The following disorders were investigated in detail:

1. peripheral polyneuropathy in Leonbergers,
2. laryngeal paralysis and polyneuropathy in Leonbergers and Saint Bernards,
3. inherited leukoencephalomyelopathy in Leonbergers and Rottweilers,
4. neuroaxonal dystrophy in Rottweilers,
5. spinocerebellar ataxia in a family of Alpine dachsbracke dogs,
6. succinic semialdehyde dehydrogenase deficiency in Saluki dogs,
7. Leigh-like subacute necrotizing encephalopathy in Yorkshire terriers.

In the light of the extensive cohort of available samples and genomic data from purebred Leonbergers and their high prevalence of inherited diseases, an additional aim was to collect current phenotype information about the dogs in order to better define their health status to facilitate the identification of further disease-associated genomic regions. Therefore, an exhaustive genealogical and genomic description of this breed's current population and diversity was performed and is presented in the first chapter of the results section.

Results

Genomic diversity and population structure of the Leonberger dog breed

Journal:	BMC Genetics Selection Evolution
Manuscript status:	published
Contributions:	Conceptualization, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing

Letko *et al. Genet Sel Evol* (2020) 52:61
https://doi.org/10.1186/s12711-020-00581-3

GSE Genetics Selection Evolution

RESEARCH ARTICLE Open Access

Genomic diversity and population structure of the Leonberger dog breed

Anna Letko[1*], Katie M. Minor[2], Vidhya Jagannathan[1], Franz R. Seefried[3], James R. Mickelson[2], Pieter Oliehoek[4] and Cord Drögemüller[1]

Abstract

Background: Leonberger is a giant dog breed formed in the 1850s in Germany. Its post-World War II popularity has resulted in a current global population of ~ 30,000 dogs. The breed has predispositions to neurodegenerative disorders and cancer, which is likely due in large part to limited genetic diversity. However, to date there is no scientific literature on the overall demography and genomic architecture of this breed.

Results: We assessed extensive pedigree records, SNP array genotype data, and whole-genome sequences (WGS) on 142,072, 1203 and 39 Leonberger dogs, respectively. Pedigree analyses identified 22 founder animals and revealed an apparent popular sire effect. The average pedigree-based inbreeding coefficient of 0.29 and average kinship of 0.31 show a dramatic loss of genetic diversity. The observed average life span decreased over time from 9.4 years in 1989 to 7.7 years in 2004. A global health survey confirmed a high prevalence of cancer and neurological disorders. Analysis of SNP-based runs of homozygosity (ROH) identified 125,653 ROH with an average length of 5.88 Mb, and confirmed an average inbreeding coefficient of 0.28. Genome-wide filtering of the WGS data revealed 28 non-protein-changing variants that were present in all Leonberger individuals and a list of 22 potentially pathogenic variants for neurological disorders of which 50% occurred only in Leonbergers and 50% occurred rarely in other breeds. Furthermore, one of the two mtDNA haplogroups detected was present in one dog only.

Conclusions: The increasing size of the Leonberger population has been accompanied by a considerable loss of genetic diversity after the bottleneck that occurred in the 1940s due to the intensive use of popular sires resulting in high levels of inbreeding. This might explain the high prevalence of certain disorders; however, genomic data provide no evidence for fixed coding variants that explain these predispositions. The list of candidate causative variants for polyneuropathy needs to be further evaluated. Preserving the current genetic diversity is possible by increasing the number of individuals for breeding while restricting the number of litters per sire/dam. In addition, outcrossing would help optimize long-term genetic diversity and contribute to the sustainability and health of the population.

Background

Leonberger is a giant dog breed that was formed around the 1850s in Germany as a watch, companion, and family dog [1]. The breed became more popular after World War II, resulting in an estimated current population of ~ 30,000 dogs [2]. Compared to other breeds, Leonberger appears to have a higher predisposition to neurodegenerative disorders and some forms of cancer such as hemangiosarcoma and osteosarcoma [3]. Recent research has shown that variants of the *ARHGEF10* [4] and *GJA9* [5] genes explain about a third of the polyneuropathy-diagnosed Leonberger cases (OMIA 001917–9615; OMIA 002119–9615), and a recessively inherited *NAPEPLD* [6] variant causes juvenile-onset leukoencephalomyelopathy (OMIA 001788–9615).

*Correspondence: anna.letko@vetsuisse.unibe.ch
[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland
Full list of author information is available at the end of the article

Historically, estimation of inbreeding has relied on in-depth pedigrees with inbreeding coefficients being estimated from pedigree-based relationships between ancestors (F_PED). Polymorphic microsatellite marker genotype data have been used to evaluate the genetic diversity in dogs (e.g. [7, 8]), and more recently, genome-wide single nucleotide polymorphism (SNP) genotype data have been used in several breeds [9–13]. Based on SNP data, runs of homozygosity (ROH) can be characterized, which enable quantification of the extent of inbreeding in diploid individuals, especially in the case of incomplete, unreliable, or missing pedigree information [14]. Furthermore, the length of the observed ROH segments can be used to distinguish between recent and ancient inbreeding [15, 16] and to draw conclusions about the population history of breeds [17].

For more than 10 years, livestock breeders have generated massive amounts of genotype data and implemented genomic selection schemes, but for dogs such data is mainly generated for gene-mapping purposes [18, 19]. The first use of SNP data to examine the association of measures of reproductive fitness in dogs was reported in golden retrievers and demonstrated the existence of a statistically significant negative correlation between fecundity and F_ROH [20]. In general, studies are based on SNP genotyping data of variable marker density and generally for less than 100 selected individuals per breed; e.g. Boccardo et al. [12] analysed only 34 individuals of the German shorthaired pointer breed to estimate a genomic inbreeding coefficient F_ROH of 0.17, whereas a mean inbreeding coefficient of only 0.023 was found based on genealogical information, showing the latter was incomplete.

The currently best-studied dog breed using SNP and whole-genome sequence (WGS) data to assess genetic diversity is the Norwegian lundehund, which is known to be at risk for a breed-specific multifactorial life-threatening syndrome [21]. The current lundehund population is highly inbred and optimal contribution selection alone results in no improvement due to the extremely high relatedness of the whole population [22]. SNP genotype data analyses revealed a substantially low genetic diversity in the lundehund and therefore outcrossing with closely-related breeds was recommended to rescue the endangered population [13, 22]. A successful example of replacing a pathogenic allele by outcrossing is the Dalmatian breed, which suffers from hyperuricosuria, and has an extremely high frequency of the deleterious recessive allele of the *SLC2A9* urate transporter gene. In this case, an individual from the pointer breed that was homozygous for the wild type allele of *SLC2A9* was used to produce unaffected heterozygous dogs, which were subsequently backcrossed to produce healthy individuals that are nearly indistinguishable from purebred Dalmatians [23].

WGS has enabled the generation of a large catalogue of genetic variants, which captures much of the variation that exists in modern dogs [19, 24]. A comprehensive set of variants, together with their allele and genotype frequencies within and across breeds, helps distinguish functionally relevant variants from neutral variants, and identify potential breed-specific regions of variation. These data have been shown to be highly useful for the analysis of phenotypic variation [25] and for the identification of loci that contribute to both simple and complex disease susceptibility in dogs [26, 27]. The resulting identification of causative variants contributes to maintaining the sustainability of breeds by reducing the number of inherited health problems. Data on the mitochondrial genome (mtDNA) that is also sequenced during WGS can be used as an additional indicator to assess diversity and to obtain a more comprehensive picture of the origin and history of canid populations [28–30].

To date, there is no report on the genetic characterization of the Leonberger breed using genomic data. In this paper, we used extensive pedigree data on 142,072 dogs, SNP array genotypes of 1203 dogs, and WGS data from 39 dogs to assess the genetic diversity within the current worldwide Leonberger breed population.

Methods

Animals

Samples from 1203 Leonberger individuals that were collected globally were used (see Additional file 1: Table S1). Genomic DNA was isolated from blood using either the Gentra PureGene blood kit (Qiagen) or the Maxwell RSC whole blood DNA kit (Promega). Thirty-nine Leonberger dogs were whole-genome sequenced either during the course of previous studies [5, 6, 31] or as unexplained cases of neurological disorders (see Additional file 1: Table S1). Pedigree records were available for 142,072 dogs through the Worldwide Independent Leonberger Database [2]. Health updates were available for 2726 dogs by owner submission through an online questionnaire.

Pedigree analyses

We assessed extensive pedigree records including 142,072 animals with the oldest recorded date of birth in 1880 until 2016. Pedigree analyses were performed using the open source software EVA v3.0 [32] and an in-house Qualitas pedigree software that is used intensively in routine genetic evaluation analyses. Loss of genetic diversity and the potential to increase diversity were assessed by mean kinship (MK) using the tabular method as described previously by Oliehoek et al. [33]. MK of the current population was calculated for 31,832 selected

dogs that were presumed to be available for breeding and alive at the time of calculation. Males born before 2009 and females born before 2011 that were assumed not to participate in breeding anymore and the known deceased dogs were excluded. Thus, the actual population size might be smaller since the current status for many animals was unknown.

Global health survey

An online questionnaire was sent to members of various European and American Leonberger breeding clubs, and in addition, was also announced periodically via Facebook. A link to an internet-based version of the questionnaire in seven languages was provided to ensure a worldwide reach of owners and breeders [34]. The questionnaire consists of three main parts with questions on general information about the dog and owner, on signs of specific diseases such as neurological disorders and cancer, and on the overall medical history of individual dogs, including the date of death. In total, responses for 2726 dogs collected between 2013 and 2019 were analysed.

SNP array genotyping and ROH analyses

Among the 1203 Leonberger individuals with SNP array genotype data, 308 were genotyped on the 460 k Axiom Canine Genotyping Array Set A (Thermo Fisher Scientific) and 895 on the 170 k Illumina CanineHD BeadChip (Illumina). The PLINK v1.9 software [35] was used to merge the genotype data and to perform quality control pruning and analyses of the runs of homozygosity (ROH). Merging of the different datasets was restricted to overlapping SNPs between both arrays and to SNPs with a genotype call missingness lower than 0.1. The final dataset consisted of 137,476 SNPs. In addition, only biallelic SNPs on the 38 canine autosomes were retained for the ROH analyses and SNPs that deviated from the Hardy Weinberg equilibrium ($p < 0.0001$) were excluded. Finally, 132,711 SNPs were available as the final dataset with one SNP per 16.52 kb. Based on a previously described method [36], we calculated that a minimum number of 70 SNPs and at least one SNP per 20 kb were required to identify a ROH and produce less than 5% randomly generated ROH. For the other parameters used to define a ROH, the default values as defined in PLINK [35] were set. An inbreeding coefficient based on ROH (F_ROH) was calculated as the total length of all ROH for one dog divided by the total length of the autosomal genome covered by SNPs. Multidimensional scaling (MDS) of pairwise genetic distances in PLINK [35] was performed to analyze the population structure. All genome positions refer to the CanFam3.1 reference sequence assembly. Figures were plotted in the R environment v3.6.0 [37] using packages CMplot v3.5.1 [38], plotrix v 3.7–6 [39], and qqman v0.1.4 [40].

Whole-genome sequencing and variant calling

WGS data of 39 Leonberger individuals were obtained by preparing a PCR-free fragment library to generate an average 18.1 × coverage (ranging from 8.3 to 33.9 ×) (see Additional file 1: Table S1) as described previously [24]. Fastq-files were mapped to the dog reference genome assembly CanFam3.1. Variant calling of single nucleotide and small indel variants (SNVs) was performed, and their functional effects were predicted based on the NCBI annotation release 105 as described in [24]. Private variants shared by all 39 Leonberger individuals and the rare variants that were present in at least one Leonberger individual were identified by comparison with the catalogue of variants in 605 publicly available control dogs from 128 various breeds and nine wolves provided by the Dog Biomedical Variant Database Consortium [24]. A list of 113 neuropathy- and Charcot-Marie-Tooth disease-associated genes was extracted from the OMIM [41] and OMIA [42] databases (see Additional file 2: Table S2). These genes were used to filter all 653 available genomes for enriched variants in the sequenced Leonberger dogs and the alternative allele frequency was calculated to detect uncommon variants associated with disease and possibly shared by other breeds related to the Leonberger. Several *in silico* prediction tools, PROVEAN [43], MutPred2 [44], MutPred-Indel [45], MutPred-LOF [46], and PredictSNP [47], were used to predict the biological consequences of the discovered variants on the corresponding proteins. Variability of the mitochondrial genome was explored to determine the diversity in the haplogroups of the studied dogs based on the nomenclature described by both Pereira et al. [48] and Duleba et al. [29]. Integrative Genomics Viewer was used to visually inspect and confirm the detected SNVs [49].

Results

Pedigree analyses

Pedigree information from 142,072 Leonberger individuals was available with a pedigree completeness index across five generations higher than 99% for the animals from the latest cohorts and exceeded 80% in 1935. More than 4000 dogs were born worldwide each year since 2000 (see Additional file 3: Figure S1). We observed an average litter size across cohorts of 6.5 puppies and a constant generation interval of 4 years. In summary, 22 founder animals were identified (10 males, 12 females) among which three (2 males and 1 female) contributed almost 46% to the last cohort (year of birth 2016). The Leonberger breed underwent a severe bottleneck during the 1940s with only 17 inbred dogs (F_PED ranging from

0.14 to 0.35) registered in 1946 (see Additional file 3: Figure S1).

The average pedigree-based inbreeding coefficient (F_PED) and the average kinship showed a steady increase over the years analysed, with an estimated F_PED of 0.29 in the last cohort (year of birth 2016) (Fig. 1a). A popular sire was defined as a male dog that sired at least 33 puppies, corresponding to five litters based on the observed average litter size. Information on the number of offspring was available for 5456 sires born between 1894 and 2016. A popular sire effect was evident since a quarter of all sires produced two thirds of all offspring (Fig. 1b). In addition, although the top breeding male who sired 434 dogs was born in 1985, the second and third top breeding males who sired 394 and 337 registered puppies, respectively, were both born in 2003. This trend is visible throughout the breeding history of Leonbergers (see Additional file 3: Figure S1): the top 27 sires, which were born between 1976 and 2010, produced more than 200 puppies each (corresponding to at least 30 litters per sire). In addition, the date of death was known for 4783 dogs (2453 females and 2330 males) born between 1959 and 2019. Life expectancy was further evaluated for the period between 1989 and 2004 for which more than 100 records were available and we assumed that there were no more living dogs. This cohort included 3044 dogs (1460 males and 1584 females) and revealed a mean of 8.2 (median 8.5) with the male longevity (mean = 7.9) being on average lower than the female longevity (mean = 8.5). In addition, a slow continuous decrease across time was observed, from an average of 9.4 years in 1989 to 7.7 years in 2004 (median: 10.1 to 8.2 years) (Fig. 1c).

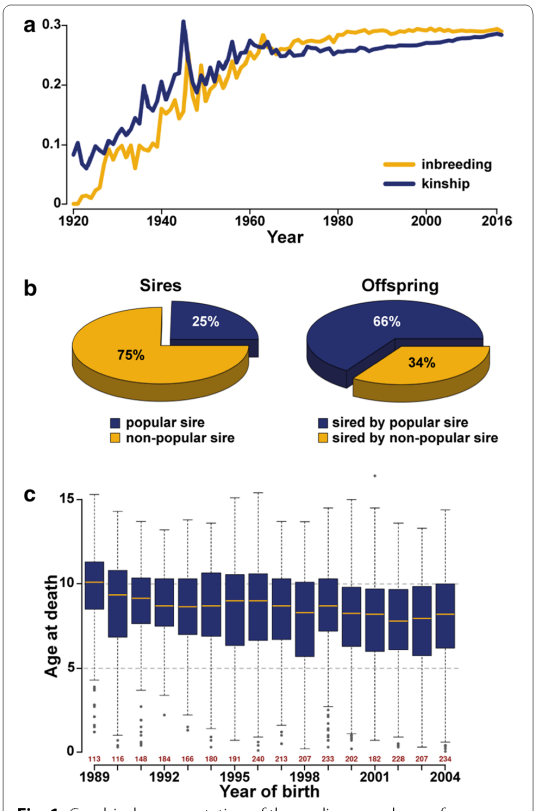

Fig. 1 Graphical representation of the pedigree analyses of Leonberger dogs. **a** Average pedigree-based inbreeding coefficient and kinship estimated per year from 1920 to 2016. **b** Pie charts of the proportion of sires used in breeding clearly showing the popular sire syndrome. Popular sires are males that produced at least 33 puppies and account for 25% of all breeding males. These sires are responsible for 66% of all offspring. **c** Boxplot of age at death per year of birth showing the life expectancy of Leonberger dogs over years 1989–2004. The exact number of records available per year is indicated in red above each year

Mean kinship was analysed on a set of 31,832 Leonberger dogs that were presumed to be available for breeding and the resulting average kinship was 0.31. The difference between minimum (0.29) and maximum (0.33) values of MK was small, and no families that would be highly unrelated to the overall population were found. Three groups color-coded as green (MK < 0.31), yellow (MK = [0.31–0.32]), and orange (MK > 0.32) were created for better visualization. The proportion of dogs with higher or lower MK differed between countries (see Additional file 4: Figure S2). The color-coded MK coefficients have been incorporated into the Worldwide Independent Leonberger Database [2].

Leonberger dogs are at risk for several health disorders

We collected medical data throughout the life of 2726 Leonberger dogs from owner-submitted health updates of their dogs. Of those 2726 Leonberger dogs, 1334 (48.9%) suffered from at least one health condition (see Additional file 5: Table S3). However, individual dogs (n = 544) often suffered from multiple disorders. In total, 586 (21.5%) dogs have reported tumor or cancer, with osteosarcoma (42.5%) and hemangiosarcoma (22.5%) being the two most frequent types. Other frequently seen groups of health issues included orthopaedic (15.8%), neurological (14.8%), endocrine (5.6%), digestive (4.2%), and cardiac (4.1%) problems. Respectively, the most frequent specific disorder in each system category was arthritis (222 cases), polyneuropathy (362 cases), hypothyroidism (142 cases), gastric torsion (54 cases), and dilated cardiomyopathy (45 cases). All reported disorders and the number of dogs suffering from them are listed in Additional file 5: Table S3. Most

of the polyneuropathy-affected dogs (71.0%) showed both breathing and gait abnormalities, whereas for 14.6% only laryngeal paralysis or breathing problems, and for 14.4% only gait abnormalities were reported (see Additional file 5: Table S3).

Multidimensional scaling analysis of worldwide Leonberger population structure

The 1203 dogs genotyped on SNP arrays represented well the Leonberger population because they were sampled throughout time and from various countries (see Additional file 1: Table S1). Subpopulations were expected due to the large geographic distances between the dogs: 579 samples (48.1%) came from the USA and Canada, 602 samples (50.1%) came from European countries, and 22 samples (1.8%) came from Australia, New Zealand, or Japan. However, multidimensional scaling (MDS) of pairwise genetic distances revealed no obvious clustering (Fig. 2a) and (see Additional file 6: Figure S3), although a small group of North American dogs did not overlap completely with the rest of the population. In light of Additional file 6: Figure S3b, the non-overlapping dogs were mostly born before 2000, whereas the more recently born dogs seem to be closer to the European population. In addition, the obvious difference in the number of genotyped dogs born in or before 2000 from North America (199 of 579) and the number of those from Europe (88 of 602) has to be taken into account, since this disproportion in sampling might contribute to the incomplete overlap.

Analysis of runs of homozygosity

The ROH-based genomic inbreeding coefficient (F_ROH) was on average 0.28 (ranging from 0.05 to 0.47). In addition, the comparison between F_PED and F_ROH revealed a few dogs with obviously incorrect pedigree records and enabled the estimation of inbreeding in individuals with missing or unknown pedigrees (Fig. 2b) and (see Additional file 1: Table S1). In total, 125,653 ROH with an average length of 5.88 Mb (ranging from 1.00 to 90.17) were identified. ROH could be divided into five groups based on their size (Fig. 2c). Most of the detected ROH (53.6%) were shorter than 4 Mb and 16.1% were longer than 10 Mb (Fig. 2c). The number of ROH

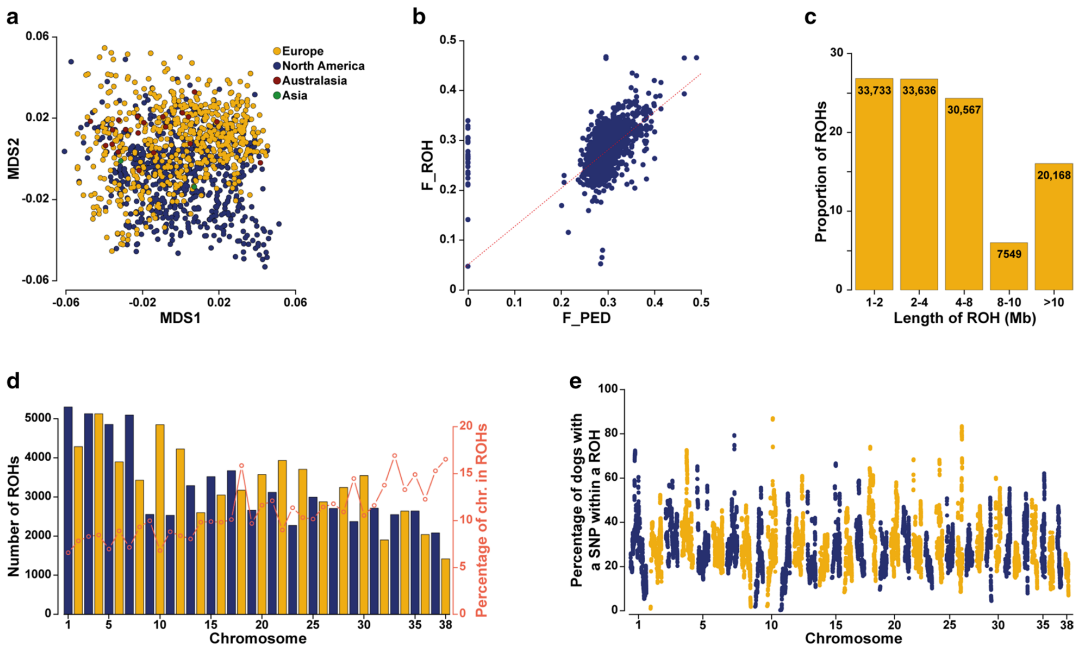

Fig. 2 Graphical representation of the genomic analyses of Leonberger dogs. **a** MDS plot of 1203 dogs showing no obvious subpopulations regardless of the geographical region of residence. **b** Correlation between inbreeding coefficients estimated from the pedigree data (F_PED) and the ROH analysis (F_ROH). Note that the dogs with unknown pedigrees (F_PED = 0) on the left side mostly show the same level of inbreeding as the dogs with known registered records. **c** Frequencies and counts of the different classes of ROH based on their length indicating both old and recent inbreeding. **d** Number of detected ROH per autosome (left y axis) and the proportion of chromosomes in a ROH (right y axis). **e** Manhattan plot showing the proportion of dogs sharing a SNP within a ROH across the 38 autosomes. Most dogs share a region on chromosomes 10 and 26

per chromosome ranged from 1419 on chromosome 38 to 5303 on chromosome 1; the proportion of overall chromosomal regions involved in homozygous regions relative to chromosomal size was highest for chromosome 33 (16.9%) and lowest for chromosome 1 (6.6%) (Fig. 2d). On average, 104.4 ROH were detected per dog with a range from 37 to 177. The total length of ROH across the genome per individual was on average 614 Mb and ranged from 104.6 to 1028.2 Mb. The proportion of dogs with a SNP within a ROH varied across the 38 autosomes as shown in Figure S4 [see Additional file 7: Figure S4]. For chromosome 10 (chr10: 44,529,255–44,831,343), 1047 of the 1203 dogs (87.0%) shared a ROH across ~ 302 kb. Only one protein-coding gene is annotated in this region, *TMEM131*, which encodes a transmembrane protein not known to be associated with any disease. This protein is assumed to play a role in the development and differentiation of T cells and in collagen recruitment and secretion [50, 51]. A ~ 236 kb region (chr26:35,428,129–35,663,964) at the end of chromosome 26 was shared by 1000 (83.1%) of the 1203 analysed dogs (Fig. 2e). Only two pseudogenes and one uncharacterized ncRNA are annotated in this region. BLAST-based comparisons of these loci revealed no significant sequence similarity to human genes.

Whole-genome sequence analysis identified no protein-changing variants fixed in Leonberger dogs

The WGS data of 39 Leonberger dogs including 38 individuals with variable neurological signs (see Additional file 1: Table S1) were compared to 614 publicly available canid WGS [24]. These sequenced Leonbergers included four individuals that were sequenced to identify the pathogenic variants in the *GJA9* and *NAPEPLD* genes described previously [5, 6]. First, we searched for Leonberger breed-specific variants that were absent from the control genomes and identified only 27 non-coding and one coding private variants that were shared by all 39 Leonberger dogs at least in the heterozygous state (see Additional file 8: Table S4). Nine of these variants were located close to the detected ROH intergenic region on chromosome 26. The only coding variant was a synonymous SNV on chromosome 1 in the *ZNF510* gene that encodes a relatively little-studied zinc finger protein.

Leonberger-specific protein-changing variants as new candidates for neuronal disorders

Next, we filtered for rare variants (i.e. alleles that occurred in the genome of at least one Leonberger dog and absent from the 614 control genomes), which led to the discovery of 62,531 SNVs, of which 1379 were predicted to be coding (see Additional file 9: Table S5). Of those, 855 were predicted to be protein-changing with the alternative allele frequency ranging from 0.01 to 0.63 (see Additional file 9: Table S5). Considering the genes that are associated with inherited neuronal disorders in other species, we found 11 potentially pathogenic variants (Table 1). This short list included two previously known pathogenic variants such as the leukoencephalomyelopathy-associated missense variant in *NAPEPLD* [6]

Table 1 Rare protein-changing private variants in Leonberger dogs

Gene	OMIM/ OMIA number	Variant designation[a]			Alternative allele frequency
		Genomic position	Coding DNA change	Protein change	
ATXN7	607640	chr20:27,234,549	c.2285 T > C	p.Val762Ala	0.1154
CEP55	610000	chr28:7,777,489	c.1334A > C	p.Tyr446Ser	0.0256
CNTF	118945	chr18:37,758,771	c.401_402insA	p.Asn137fs	0.0128
ELOVL4	605512	chr12:40,850,011	c.424_433delGGAGCACAGC	p.Gly142fs	0.0128
FDX1L	614585	chr20:50,798,661	c.124_125insGGCCGCCATCACCAC GGCGGTGAGCACCGCCGCCAGCAG CACCAGCCCGTCAGCGTTGAGCCG	p.Thr41_Ala42insGlyProProSerProArgArgTer	0.0128
GJA9[b]	611923/2119–9615	chr15:3,863,519	c.1108_1109delGA	p.Glu370fs	0.0513
MCM3AP	603294	chr31:39,553,337	c.1785_1793delCTCTGAAGG	p.Ser596_Gly598del	0.0128
NAPEPLD[c]	612334/1788–9615	chr18:16,987,520	c.559G > C	p.Ala187Pro	0.1154
PLEKHG5	611101	chr5:60,325,903	c.1585C > T	p.Gln529*	0.0128
SPTBN4	606214/2232–9823	chr1:113,215,064	c.1247_1248insGGTAGCCCATGCGGT	p.Ala416_Ala417insValAlaHisAlaVal	0.0128
SYNE1	608441	chr1:42,549,994	c.17359C > T	p.Arg5788Trp	0.0256

[a] Additional details including the protein prediction effects are described in Additional file 10 Table S6. All positions refer to the CanFam3.1 reference sequence assembly

[b] Previously described polyneuropathy-associated variant [5]

[c] Previously described leukoencephalomyelopathy-associated variant [6]

at a frequency of 0.12 and the polyneuropathy-associated frameshift variant in *GJA9* [5] at a frequency of 0.05. In addition to a single missense variant predicted to be neutral, eight of the other nine variants were predicted as probably deleterious including two missense, two frameshift, two in-frame-insertions, one in-frame-deletion, and one nonsense variant (Table 1) and (see Additional file 9: Table S5 and Additional file 10: Table S6).

Protein-changing variants in candidate genes for neuronal disorders enriched in Leonberger dogs

Because 33 of the 39 sequenced Leonberger individuals suffered from an unexplained form of polyneuropathy (see Additional file 1: Table S1), additional filtering was done for variants that were present in 113 putative candidate genes (see Additional file 2: Table S2) and could be involved in the development of the disease. Focusing on coding variants that occurred in at least one sequenced individual resulted in 649 variants that affected 100 genes, among which 232 were predicted to be protein-changing (see Additional file 11: Table S7). All these variants were also present at variable frequencies in other dog breeds and wolves.

To search for variants that are enriched in Leonberger dogs, we considered those with an alternative allele frequency higher than 0% and lower than 10% in the 614 control canids and that were more than twice as common in the Leonberger dogs as in the controls. This approach yielded a short list of 22 SNVs that affect 17 genes including 21 missense and one in-frame-deletion variant (Table 2). In addition to the 11 variants predicted to be neutral, 11 other variants were predicted as most likely deleterious and represented potentially pathogenic variants causing polyneuropathy (Table 2) and (see Additional file 10: Table S6). Interestingly, a new potentially deleterious missense variant was found in a different codon in the *GJA9* gene in which a variant has already been reported to be associated with polyneuropathy in Leonbergers [5]. This newly discovered *GJA9* variant occurred at a quite high frequency (0.55) within the sequenced cohort of Leonberger dogs but was also comparatively common (frequency of 0.09) among the control

Table 2 Protein-changing variants in polyneuropathy-associated candidate genes enriched in Leonberger dogs compared to 614 controls from various breeds

Gene	OMIM/ OMIA number	Variant designation[a]			Alternative allele frequency	
		Genomic position	Coding DNA change	Protein change	Leonbergers	Controls
CNTNAP1	602346	chr9:20,294,320	c.3863G > C	p.Arg1288Pro	0.3974	0.0366
CNTNAP1	602346	chr9:20,298,261	c.2585G > A	p.Gly862Glu	0.1538	0.0067
DHTKD1	614984	chr2:24,316,951	c.1820C > T	p.Ala607Val	0.0897	0.0049
DIAPH3	614567	chr22:15,880,854	c.2000G > A	p.Cys667Tyr	0.1538	0.0689
DST	113810	chr12:23,771,235	c.18224C > T	p.Thr6075Met	0.0769	0.0008
DST	113810	chr12:23,782,332	c.17800C > T	p.Arg5934Trp	0.4359	0.0330
DST	113810	chr12:23,846,624	c.10661C > T	p.Ser3554Leu	0.0769	0.0369
DYNC1H1	600112	chr8:70,064,306	c.13757C > T	p.Pro4586Leu	0.1282	0.0051
GARS	600287	chr14:43,322,005	c.622G > A	p.Val208Ile	0.1154	0.0025
GJA9	611923/2119–9615	chr15:3,862,761	c.344G > C	p.Arg115Thr	0.5513	0.0868
JPH1	605266	chr29:22,710,800	c.1531A > G	p.Ile511Val	0.1923	0.0299
LOC477508	602072	chr26:16,348,822	c.298G > A	p.Val100Met	0.1795	0.0066
MME	120520	chr23:49,045,461	c.1700A > T	p.Gln567Leu	0.0769	0.0157
NEFH	162230	chr26:22,729,485	c.1046G > A	p.Arg349His	0.2821	0.0116
NEFH	162230	chr26:22,732,974	c.1901_1924delTGAAGG AGGAGGCCAAGTCCCCAG	p.Val640_Pro647del	0.3974	0.0849
NEFH	162230	chr26:22,733,415	c.2326C > G	p.Pro768Ala	0.0513	0.0059
OTOF	603681	chr17:20,534,866	c.3451G > A	p.Ala1151Thr	0.7051	0.0094
PDXK	179020	chr31:37,707,445	c.774G > C	p.Arg258Ser	0.1923	0.0898
PRX	605725	chr1:113,290,407	c.784G > A	p.Ala262Thr	0.1282	0.0257
SBF2	607697	chr21:33,036,988	c.4114C > T	p.Pro1372Ser	0.1667	0.0361
SLC12A6	604878	chr30:853,540	c.2498C > T	p.Ala833Val	0.0513	0.0250
WNK1	605232	chr27:42,911,057	c.7448C > G	p.Thr2483Arg	0.0128	0.0058

[a] Additional details including the protein prediction effects are described in Additional file 10: Table S6. All positions refer to the CanFam3.1 reference sequence assembly

dogs from 35 breeds (see Additional file 9: Table S5). The other variants occurred at different frequencies in several breeds (see Additional file 9: Table S5). Only one missense variant in the *DST* gene was quite rare, i.e. it was present in the heterozygous state in only one Howavart dog (see Additional file 9: Table S5).

Mitochondrial genome diversity is limited in Leonberger dogs

Finally, we evaluated the diversity of the mitochondrial genome of the 39 studied Leonberger dogs and detected only two haplotypes. Based on the nomenclature standard using only the mtDNA D-loop [48], 38 dogs showed the A2 haplotype and one dog the A17 haplotype. The same results were obtained when using the complete mitogenome database standardized haplogroup nomenclature [29] with 38 dogs belonging to the A1b2a1a1 haplogroup and the one dog to the A1b1a1a haplogroup (see Additional file 12: Table S8). In addition, 10 SNVs were detected in the coding part of the mtDNA, among which seven were private, occurring rarely in individual Leonbergers and were not found in the 614 controls (see Additional file 12: Table S8). This list included two variants that affect tRNA genes and a single missense variant in the *ND2* gene observed in one dog only (Table 3).

Discussion

This is the first study to comprehensively describe the genomic diversity and population structure of the Leonberger breed by exploring the distribution of ROH and the SNV based on SNP array and WGS data. In addition, the genealogical pedigree information was used to compare the estimated inbreeding coefficients and mean kinship in the current population. The new genomic regions reported here might underlie breed-specific characteristics and provide information about the genetic structure of the breed, such as selection pressure on specific traits, inbreeding levels, or genetic bottlenecks.

Pedigree analyses based on extensive data revealed average litter size and life expectancy for the Leonberger dogs included in our study that are in agreement with previously reported results [52]. Long ROH segments (~10 Mb) occur as a result of recent inbreeding, whereas short ROH (<2 Mb) indicate genomic regions that are identical by descent from older ancestors [12]. In the Leonberger dogs studied here, the 1- to 2-Mb segments account for approximately 25% of the detected ROH and those that are longer than 8 Mb account for about 20% (Fig. 2c), which indicates both old and recent inbreeding. This is in accordance with the known history of the breed and the observed current breeding practices. Smaller-sized ROH are challenging to detect because of the limited number of markers and their varying density across the genome. For example, the highly shared ROH on chromosome 26 for which 83% of the dogs were homozygous was confirmed by the identification of nine variants that were shared by all 39 sequenced dogs.

The pedigree analysis revealed the occurrence of a clear bottleneck during the 1940s and the extensive use of popular sires and line-breeding that resulted in the recent high kinship. A limited number of ancestors with high genetic contribution that had a big impact on within-breed relatedness, resulted in inbreeding due to this relatedness, and in turn inbreeding depression. It is likely that some disease-causing genetic variants that these ancestors carried are now spread throughout the entire population. Moreover, the MDS analysis did not show any clear differentiation based on geographical or purpose lineages in the Leonberger population, which contrasts with previous results in other breeds, e.g. the English greyhounds and Labrador retrievers that cluster based on specific working and show lineages or the Italian greyhounds and Shetland sheepdogs that form discrete clusters separating the European and North American populations [53].

Table 3 Private variants in mitochondrial DNA of Leonberger dogs

Gene	Variant designation[a]	Proportion of variant[b]	OMIM/OMIA number	Type
tRNA-Phe	m.49A>G	96%	590070	Unknown
tRNA-Val	m.1069 T>C	99%	590105	Unknown
ND2	m.4764 T>C; p.Ile284Thr	100%	516001	Missense
COX1	m.5681A>G	66%, 89%	516030	Synonymous
ND4	m.11211C>T	72%, 66%	516003	Synonymous
ND5	m.13118 T>C	62%	516005	Synonymous
ND5	m.13544A>G	100%, 100%, 100%, 100%	516005	Synonymous

[a] All positions refer to the CanFam3.1 reference sequence assembly

[b] Percentage of variant represents heteroplasmy, 100% represents homoplasmy, for each dog that carries the variant calculated from the coverage of the affected mtDNA region

Our worldwide health survey of almost 3000 Leonberger individuals confirmed a high prevalence of cancer, particularly osteosarcoma and hemangiosarcoma, neuromuscular disorders, and hypothyroidism. In other dog breeds, genetic association studies that aimed at identifying genetic risk factors [27] have revealed a strong association with specific genome regions, e.g. for osteosarcoma [54, 55]. Similar studies could be performed in the Leonberger breed to unravel such associations and identify markers for selection against the elevated risk of developing certain forms of cancer. In addition, hypothyroidism was recorded quite frequently in our survey, confirming earlier reports for this breed [56, 57]. However, diseases such as arrhythmia leading to sudden cardiac death [58] that has been reported to be more prevalent in Finnish Leonbergers occurred only sporadically in our survey. Besides the orthopaedic disorders, which are known to be polygenic, the second most prevalent disease category that was reported in our survey was neurological disorders, with predominantly different forms of polyneuropathy as reported earlier [59]. A recent survey conducted by the Leonberger Health Foundation International based on more than 1000 dogs from 24 countries also highlighted the high prevalence of certain forms of cancer such as osteosarcoma and hemangiosarcoma, as well as that of neurological disorders, such as laryngeal paralysis and other forms of polyneuropathy [60]. Although these two surveys were not truly comprehensive and possibly biased towards selected conditions, they were performed independently at different times, and therefore we considered that they provided reasonable estimates of the true frequencies.

We sequenced the genomes of more than 30 polyneuropathy-affected dogs with the aim of unravelling additional disease-causing variants. The WGS data provided evidence for both breed-specific and enriched variants in the genomes of polyneuropathy-affected Leonbergers at different frequencies. The short list of 21 potentially pathogenic variants in candidate genes for neurological disorders includes two previously reported variants: one in *GJA9* that causes polyneuropathy [5] and one in *NAPE-PLD* that causes leukoencephalomyelopathy [6]; these findings confirm the potential usefulness of the approach chosen. The newly discovered variants are well-known candidates for neurological disorders in various species. Future genotyping of the 19 newly identified potentially deleterious protein-changing variants in cohorts of well-phenotyped Leonberger dogs is needed to establish association of these variants with diseases. The second potentially pathogenic missense variant in *GJA9* is compelling since it is also present at a lower frequency in dogs of unrelated breeds. Two of the three mtDNA SNV that affect tRNA genes will also be interesting to follow-up because a variant in the *tRNA-Tyr* gene causes a familial form of sensory ataxic neuropathy in golden retrievers [61]. Since, to date, we have been unable to identify additional loci that are associated with polyneuropathy disorders inherited in a Mendelian fashion by genome-wide association study, we believe that a subset of these disorders could have a complex genetic nature and that the individual's likelihood of developing the disorder depends on a combination of multiple alleles at multiple loci in addition to environmental factors. Future investigation of the presence of large structural variations throughout the Leonberger genome would also be informative when an improved canine reference assembly, an improved annotation, and long-read sequence data become available. Earlier studies proposed the presence of a sex bias, since more male Leonbergers are affected with polyneuropathy than female Leonbergers [62], but to date no major risk factor has been identified on the X chromosome.

Conclusions

Considerable genetic diversity has been lost in the Leonberger breed due to a bottleneck that occurred during the last century and to an increasing population size thereafter. This situation appears to be due primarily to the use of popular sires resulting in high levels of inbreeding, which has also facilitated the spread of undesirable genetic traits within the gene pool. Maintaining the current level of genetic diversity will only be possible through informed selection decisions, especially by including more dogs in breeding programs, avoiding repeated matings and the use of popular sires, and minimizing coancestry among the selected parents. The established mean kinship groups may also help breeders to select the most suitable mating pairs in addition to the already used pedigree-based inbreeding coefficients. In addition, careful outcrossing might help optimize long-term genetic diversity, increase heterozygosity, and reduce the frequency of disease-causing alleles in order to lower the incidence of various health problems. The list of putative pathogenic variants in candidate genes for neurologic and neuromuscular diseases provided in this paper might enable the identification of new disease-causing mutations in the future. Our results also illustrate the most likely heterogeneous genetic architecture of this group of diseases that was thought to be Mendelian, but resembles more and more a polygenic complex disorder.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12711-020-00581-3.

Additional file 1: Table S1. List of the 1203 dogs analysed in this study with available genomic information. The table includes supplementary descriptive details of the studied Leonberger dogs including accession numbers and known health history of the WGSed dogs.

Additional file 2: Table S2. List of 113 selected genes used for filtering variants in the 653 available WGS data. Putative candidate genes and their phenotype association in humans or animals based on OMIM [41] and OMIA [42] databases showing the number of variants detected in each of these genes in the Leonberger genomes.

Additional file 3: Figure S1. Population size of the Leonberger breed between the years 1920–2016. The blue, yellow, and green lines represent the number of breeding males, breeding females, and puppies they produced per year, respectively. The red dashed line represents the number of expected parent pairs for the given number of born puppies per year, assuming the average litter size of 6.5 puppies. Panel (**a**) shows the increasing population size from 1920 to 1978 and the apparent bottleneck around 1946. In the 1970s, the number of dogs born started to increase rapidly whereas the number of breeding males used started to decrease. Panel (**b**) shows the continuously increasing population from 1979 to 2016. Note that the number of dams is more or less as expected but the number of sires constantly decreases to about half of that of the dams in recent years, illustrating the popular sire syndrome.

Additional file 4: Figure S2. Mean kinship of the estimated current population of 31,832 Leonberger dogs per country. The proportion of dogs belonging to the three groups that indicate the increasing relatedness to the whole population, is shown as green (MK < 0.31), yellow (MK = [0.31–0.32]), and orange (MK > 0.32) by country. The total number of dogs recorded in each country is shown in red at the top of the corresponding columns.

Additional file 5: Table S3. Distribution of disease phenotypes reported in 2726 Leonberger dogs. The table shows the number of dogs for which the owners reported particular disorders. Each more general group is divided in specific subgroups. Note that many dogs have multiple health issues and therefore the sum of the individual counts will be larger than the total of 1334 dogs with at least one health condition reported.

Additional file 6: Figure S3. MDS plots of the 1203 Leonberger dogs highlighted by different groups. Description: Panel (**a**) shows the distribution of the dogs by country. Panel (**b**) shows the dogs divided into five groups by their year of birth. This is also reflected in panel (**c**), where the dogs are coded by their calculated MK coefficient, which was not determined for the older dogs.

Additional file 7: Figure S4. Proportion of dogs with a SNP within a ROH on each of the 38 autosomes. Individual Manhattan plots showing the percentage of 1203 dogs genotyped on the SNP array that share a SNP within a ROH for each canine autosome.

Additional file 8: Table S4. List of private whole-genome sequence variants in 39 Leonberger dogs. Breed-specific SNVs present at least in the heterozygous state in all Leonberger dogs and absent from the 614 control genomes.

Additional file 9: Table S5. List of rare whole-genome sequence variants in 39 Leonberger dogs. Breed-specific SNVs present in at least one Leonberger dog and absent from the 614 control genomes.

Additional file 10: Table S6. List of the potentially pathogenic variants for neurological disorders detected in 39 Leonberger dogs including the predictions of the effect of the variants on the proteins. The table includes detailed description and predictions of the biological consequences of the discovered variants on the corresponding proteins using different in silico prediction tools: PROVEAN [43], the MutPred suite (includes MutPred2 [44], MutPred-Indel [45], MutPred-LOF [46]), and PredictSNP [47].

Additional file 11: Table S7. List of whole-genome sequence variants present in at least one Leonberger dog in the 113 selected genes showing genotypes of all 653 available genomes. SNVs present in polyneuropathy- and Charcot-Marie-Tooth disease-associated candidate genes (listed in Additional file 2: Table S2) and their alternative allele frequency in the Leonberger dogs and 128 other dog breeds and wolves.

Additional file 12: Table S8. Diversity of the mitochondrial genome of 39 Leonberger dogs. The table shows the low diversity of Leonberger mtDNA based on two nomenclature standards [46, 48]: most dogs belong to the A1b2a1a1 or A2 haplogroup whereas only one dog belongs to the A1b1a1a or A17 haplogroup. In addition, seven private SNVs detected in the mtDNA of the Leonberger dogs are also highlighted.

Acknowledgements
The authors are grateful to all Leonberger owners and breeders, the Leonberger Health Foundation International, the Schweizerischer Leonberger Club, and the International Leonberger Union for their support during the entire project. Special thanks go to Wilma and Ben Kroon for sharing the Worldwide Independent Leonberger Database. We thank Nathalie Besuchet-Schmutz and Heidi Signer-Hasler for invaluable technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the WGS and providing high-performance computational infrastructure.

Authors' contributions
AL carried out the analyses, visualized the results, and drafted the manuscript. KMM contributed data and critically revised the manuscript. VJ performed bioinformatics. FRS contributed to the pedigree analysis and interpretation of data. PO performed the mean kinship analysis and helped with the interpretation of data. JRM and CD designed the study and revised the manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The WGS dataset generated and analysed during the current study is available in the European Nucleotide Archive (https://www.ebi.ac.uk/ena). The 39 Leonberger dog genomes have been made freely available under study accession number PRJEB16012 and all sample accession numbers are listed in Table S1 (see Additional file 1: Table S1). The SNP genotyping data is available on reasonable request.

Ethics approval and consent to participate
Dog blood or cheek swab samples were obtained via elective owner submission for diagnostic purposes or were submitted for genotyping of the previously reported polyneuropathy-associated variants in Leonbergers. Written consent was obtained from all dog owners.

Consent for publication
Not applicable.

Competing interests
Both the University of Minnesota and the University of Bern offer genotyping tests for polyneuropathy- and leukoencephalomyelopathy-associated variants in their respective laboratories and all proceeds from these tests go towards ongoing canine genetic research. The authors declare that they have no other competing interests.

Author details
[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland. [2] Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, St. Paul, Minnesota 55108, USA. [3] Qualitas AG, 6300 Zug, Switzerland. [4] Dogs Global, 6871 EK Renkum, the Netherlands.

Received: 18 June 2020 Accepted: 2 October 2020
Published online: 14 October 2020

References

1. Bliss-Isberg C. The Leonberger: a comprehensive guide to the lion king of breeds. Sea Cliff: Revodana Publishing; 2016.
2. The Worldwide Independent Leonberger Database. 2005. https://www.leonberger-database.com/. Accessed 1 Jan 2019.
3. Gough A, Thomas A, O'Neill D. Breed predispositions to disease in dogs and cats. 3rd ed. Oxford: Wiley; 2018.
4. 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.
5. Becker D, Minor KM, Letko A, Ekenstedt KJ, Jagannathan V, Leeb T, et al. A *GJA9* frameshift variant is associated with polyneuropathy in Leonberger dogs. BMC Genomics. 2017;18:662.
6. Minor KM, Letko A, Becker D, Drögemüller M, Mandigers PJJ, Bellekom SR, et al. Canine NAPEPLD-associated models of human myelin disorders. Sci Rep. 2018;8:5818.
7. Melis C, Borg ÅA, Espelien IS, Jensen H. Low neutral genetic variability in a specialist puffin hunter: the Norwegian Lundehund. Anim Genet. 2013;44:348–51.
8. Ontiveros ES, Hughes S, Penedo MCT, Grahn RA, Stern JA. Genetic heterogeneity and diversity of North American golden retrievers using a low density STR marker panel. PLoS One. 2019;14:e0212171.
9. Yang Q, Chen H, Ye J, Liu C, Wei R, Chen C, et al. Genetic diversity and signatures of selection in 15 Chinese indigenous dog breeds revealed by genome-wide SNPs. Front Genet. 2019;10:1174.
10. Mortlock SA, Williamson P, Khatkar MS. Copy number variation and variant discovery in Bullmastiff dogs. Anim Genet. 2019;50:177–81.
11. Gajaweera C, Kang JM, Lee DH, Lee SH, Kim YK, Wijayananda HI, et al. Genetic diversity and population structure of the Sapsaree, a native Korean dog breed. BMC Genet. 2019;20:66.
12. Boccardo A, Marelli SP, Pravettoni D, Bagnato A, Busca GA, Strillacci MG. The German Shorthair pointer dog breed (*Canis lupus familiaris*): Genomic inbreeding and variability. Animals. 2020;10:498.
13. Stronen AV, Salmela E, Baldursdottir BK, Berg P, Espelien IS, Jarvi K, et al. Genetic rescue of an endangered domestic animal through outcrossing with closely related breeds: a case study of the Norwegian Lundehund. PLoS One. 2017;12:0177429.
14. Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: a possible solution to an old problem. Livest Sci. 2014;166:26–34.
15. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of homozygosity in European populations. Am J Hum Genet. 2008;83:359–72.
16. Keller MC, Visscher PM, Goddard ME. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics. 2011;189:237–49.
17. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. BMC Genet. 2012;13:70.
18. Bai B, Zhao WM, Tang BX, Wang YQ, Wang L, Zhang Z, et al. DoGSD: the dog and wolf genome SNP database. Nucleic Acids Res. 2015;43:D777–D78383.
19. Ostrander EA, Wayne RK, Freedman AH, Davis BW. Demographic history, selection and functional diversity of the canine genome. Nat Rev Genet. 2017;18:705–20.
20. Chu ET, Simpson MJ, Diehl K, Page RL, Sams AJ, Boyko AR. Inbreeding depression causes reduced fecundity in Golden Retrievers. Mamm Genome. 2019;30:166–72.
21. Metzger J, Pfahler S, Distl O. Variant detection and runs of homozygosity in next generation sequencing data elucidate the genetic background of Lundehund syndrome. BMC Genomics. 2016;17:535.
22. Kettunen A, Daverdin M, Helfjord T, Berg P. Cross-breeding is inevitable to conserve the highly inbred population of puffin hunter: the Norwegian Lundehund. PLoS One. 2017;12:e0170039.
23. Bannasch D, Safra N, Young A, Karmi N, Schaible RS, Ling GV. Mutations in the SLC2A9 gene cause hyperuricosuria and hyperuricemia in the dog. PLoS Genet. 2008;4:e1000246.
24. Jagannathan V, Drögemüller C, Leeb T. Dog Biomedical Variant Database Consortium (DBVDC). A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. Anim Genet. 2019;50:695–704.
25. Schoenebeck JJ, Ostrander EA. Insights into morphology and disease from the dog genome project. Annu Rev Cell Dev Biol. 2014;30:535–60.
26. Lindblad-Toh K. What animals can teach us about evolution, the human genome, and human disease. Ups J Med Sci. 2020;125:1–9.
27. Ostrander EA, Dreger DL, Evans JM. Canine cancer genomics: Lessons for canine and human health. Annu Rev Anim Biosci. 2019;7:449–72.
28. Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E, Lundeberg J. A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. Proc Natl Acad Sci USA. 2004;101:12387–90.
29. Duleba A, Skonieczna K, Bogdanowicz W, Malyarchuk B, Grzybowski T. Complete mitochondrial genome database and standardized classification system for *Canis lupus familiaris*. Forensic Sci Int Genet. 2015;19:123–9.
30. 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.
31. Hedan B, Cadieu E, Botherel N, Dufaure C, Letko A, Rimbault M, et al. Identification of a missense variant in *MFSD12* involved in dilution of phaeomelanin leading to white or cream coat color in dogs. Genes. 2019;10:386.
32. Berg P, Nielsen J, Sørensen MK. EVA: realized and predicted optimal genetic contributions. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production: 13–18 August 2006; Belo Horizonte. 2006. p. 27–09.
33. Oliehoek PA, Bijma P, van der Meijden A. History and structure of the closed pedigreed population of Icelandic Sheepdogs. Genet Sel Evol. 2009;41:39.
34. Leonberger Health Questionnaire. University of Bern and University of Minnesota. https://www.genetics.unibe.ch/services/dog/european_leonberger_health_questionnaire/index_eng.html. Accessed 1 Jan 2020.
35. Chang CC, Chow CC, Tellier CAML, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
36. Lencz T, Lambert C, DeRosse P, Burdick KE, Vance MT, Kane JM, et al. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. Proc Natl Acad Sci USA. 2007;104:19942–7.
37. R Core Team. A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
38. Yin L. CMplot: Circle Manhattan Plot. 2019. https://github.com/YinLiLin/R-CMplot. Accessed 28 Feb 2020.
39. Lemon J. Plotrix: a package in the red light district of R. R-News. 2006;6:8–12.
40. Turner SD. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. BioRxiv. 2014. https://doi.org/10.1101/005165.
41. Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University. https://omim.org/. Accessed 28 Feb 2020.
42. Online Mendelian Inheritance in Animals, OMIA. Sydney School of Veterinary Science. https://omia.org/. Accessed 28 Feb 2020.
43. Choi Y, Chan AP. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics. 2015;31:2745–7.
44. Pejaver V, Urresti J, Lugo-Martinez J, Pagel KA, Lin GN, Nam H-J, et al. MutPred2: inferring the molecular and phenotypic impact of amino acid variants. biorxiv. 2017. https://doi.org/10.1101/134981.
45. Pagel KA, Antaki D, Lian A, Mort M, Cooper DN, Sebat J, et al. Pathogenicity and functional impact of non-frameshifting insertion/deletion variation in the human genome. PLoS Comput Biol. 2019;15:e1007112.
46. Pagel KA, Pejaver V, Lin GN, Nam H-J, Mort M, Cooper DN, et al. When loss-of-function is loss of function: assessing mutational signatures and impact of loss-of-function genetic variants. Bioinformatics. 2017;33:i389–i398398.
47. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput Biol. 2014;10:e1003440.
48. Pereira L, Van Asch B, Amorim A. Standardisation of nomenclature for dog mtDNA D-loop: a prerequisite for launching a *Canis familiaris* database. Forensic Sci Int. 2004;141:99–108.
49. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.

50. Kerkel K, Schupf N, Hatta K, Pang D, Salas M, Kratz A, et al. Altered DNA methylation in leukocytes with trisomy 21. PLoS Genet. 2010;6:e1001212.
51. Zhang Z, Bai M, Barbosa GO, Chen A, Wei Y, Luo S, et al. Broadly conserved roles of TMEM131 family proteins in intracellular collagen assembly and secretory cargo trafficking. Sci Adv. 2020;6:eaay7667.
52. Leroy G, Phocas F, Hedan B, Verrier E, Rognon X. Inbreeding impact on litter size and survival in selected canine breeds. Vet J. 2015;203:74–8.
53. Lampi S, Donner J, Anderson H, Pohjoismäki J. Variation in breeding practices and geographic isolation drive subpopulation differentiation, contributing to the loss of genetic diversity within dog breed lineages. Canine Med Genet. 2020;7:5.
54. 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.
55. Sakthikumar S, Elvers I, Kim J, Arendt ML, Thomas R, Turner-Maier J, et al. *SETD2* is recurrently mutated in whole-exome sequenced canine osteosarcoma. Cancer Res. 2018;78:3421–31.
56. Smallwood LJ, Barsanti JA. Hypoadrenocorticism in a family of leonbergers. J Am Anim Hosp Assoc. 1995;31:301–5.
57. Segalini V, Hericher T, Grellet A, Rosenberg D, Garnier F, Fontbonne A. Thyroid function and infertility in the dog: a survey in five breeds. Reprod Domest Anim. 2009;44(Suppl 2):211–3.
58. Wiberg M, Niskanen JE, Hytönen M, Dillard K, Hagner K, Anttila M, et al. Ventricular arrhythmia and sudden cardiac death in young Leonbergers. J Vet Cardiol. 2020;27:10–22.
59. Shelton GD, Podell M, Poncelet L, Schatzberg S, Patterson E, Powell HC, et al. Inherited polyneuropathy in Leonberger dogs: A mixed or intermediate form of Charcot-Marie-Tooth disease? Muscle Nerve. 2003;27:471–7.
60. Leonberger Health Survey 2019. Leonberger Health Foundation International. https://www.leohealth.org/images/surveys-research/LHFI_HealthSurvey_PresentationMay2019.pdf. Accessed 1 Jan 2020.
61. Baranowska I, Jäderlund KH, Nennesmo I, Holmqvist E, Heidrich N, Larsson NG, et al. Sensory ataxic neuropathy in golden retriever dogs is caused by a deletion in the mitochondrial *tRNATyr* gene. PLoS Genet. 2009;5:e1000499.
62. Hultin Jäderlund K, Baranowska Körberg I, Nødtvedt A. Inherited polyneuropathy in Leonberger dogs. J Vet Intern Med. 2011;25:997–1002.

Publisher's Note

A *GJA9* frameshift variant is associated with polyneuropathy in Leonberger dogs

Journal:	BMC Genomics
Manuscript status:	published
Contributions:	Investigation, Visualization, Writing – original draft, Writing – review & editing

Becker et al. BMC Genomics (2017) 18:662
DOI 10.1186/s12864-017-4081-z

BMC Genomics

RESEARCH ARTICLE Open Access

A *GJA9* frameshift variant is associated with polyneuropathy in Leonberger dogs

Doreen Becker[1†], Katie M. Minor[2†], Anna Letko[1†], Kari J. Ekenstedt[2,3], Vidhya Jagannathan[1], Tosso Leeb[1], G. Diane Shelton[4], James R. Mickelson[2†] and Cord Drögemüller[1*†]

Abstract

Background: Many inherited polyneuropathies (PN) observed in dogs have clinical similarities to the genetically heterogeneous group of Charcot-Marie-Tooth (CMT) peripheral neuropathies in humans. The canine disorders collectively show a variable expression of progressive clinical signs and ages of onset, and different breed prevalences. Previously in the Leonberger breed, a variant highly associated with a juvenile-onset PN was identified in the canine orthologue of a CMT-associated gene. As this deletion in *ARHGEF10* (termed LPN1) does not explain all cases, PN in this breed may encompass variants in several genes with similar clinical and histopathological features.

Results: A genome-wide comparison of 173 k SNP genotypes of 176 cases, excluding dogs homozygous for the *ARHGEF10* variant, and 138 controls, was carried out to detect further PN-associated variants. A single suggestive significant association signal on CFA15 was found. The genome of a PN-affected Leonberger homozygous for the associated haplotype was sequenced and variants in the 7.7 Mb sized critical interval were identified. These variants were filtered against a database of variants observed in 202 genomes of various dog breeds and 3 wolves, and 6 private variants in protein-coding genes, all in complete linkage disequilibrium, plus 92 non-coding variants were revealed. Five of the coding variants were predicted to have low or moderate effect on the encoded protein, whereas a 2 bp deletion in *GJA9* results in a frameshift of high impact. *GJA9* encodes connexin 59, a connexin gap junction family protein, and belongs to a group of CMT-associated genes that have emerged as important components of peripheral myelinated nerve fibers. The association between the *GJA9* variant and PN was confirmed in an independent cohort of 296 cases and 312 controls. Population studies showed a dominant mode of inheritance, an average age of onset of approximately 6 years, and incomplete penetrance.

Conclusions: This *GJA9* variant represents a highly probable candidate variant for another form of PN in Leonberger dogs, which we have designated LPN2, and a new candidate gene for CMT disease. To date, approximately every third PN-diagnosed Leonberger dog can be explained by the *ARHGEF10* or *GJA9* variants, and we assume that additional genetic heterogeneity in this condition exists in the breed.

Keywords: Dog, Rare disease, Neurological disorder, Peripheral nerve, Polyneuropathy, Charcot-Marie-tooth, Genome wide association, Whole-genome resequencing, Gene test, Connexons, Connexin genes, Gap junctions

Background

Several studies have reported the occurrence of peripheral neuropathies (PN) in many breeds of dogs [1–4]. Affected dogs present with abnormal motor signs including weakness, hypotonia and muscle atrophy secondary to denervation [1]. Focal signs such as laryngeal paralysis, a change in the pitch of the bark, inspiratory stridor, and dyspnea, can be the only clinical abnormalities early in the progressive course of the disease and are considered as strong indicators of the presence of an underlying neuropathy that may later progress [1–3]. An observational study across breeds revealed that approximately every second dog with laryngeal paralysis had evidence of diffuse PN [3].

The histopathological phenotype seen in many canine PNs is similar to the most common group of inherited polyneuropathies in humans, known as Charcot-Marie-

* Correspondence: cord.droegemueller@vetsuisse.unibe.ch
†Equal contributors
[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland
Full list of author information is available at the end of the article

Tooth (CMT) disease [5]. Nearly 30 years ago it was already speculated that some forms of canine PN represent inherited diseases [2]. This was recently confirmed by unraveling causative recessive variants in the canine orthologue of a human CMT-associated gene (*NDRG1*) in Greyhounds [6] and Alaskan Malamute dogs [7] with early-onset PN (OMIA 001292–9615). More recently, *RAB3GAP1*-associated forms of syndromic PN accompanied by ocular abnormalities and neuronal vacuolation (OMIA 001970–9615) were described independently in Black Russian Terriers [8], Rottweilers [9] and Alaskan Husky dogs [10].

The Leonberger breed was formed by the crossing of several large-bodied breeds, including Saint Bernards, Newfoundlands, and Great Pyrenees. Chronic nerve fiber loss associated with axonal degeneration, decreased myelinated fiber density and a shift of the axonal size-frequency distribution toward smaller fibers is the predominant finding in peripheral nerves of PN-affected Leonberger dogs [1, 5]. PN in Leonbergers is twice as frequent in males as females, has a wide age range of onset, and due to the absence of pain and gradual onset of clinical signs, may go undetected for a number of months to years. An *ARHGEF10* loss-of-function 10 bp deletion, affecting another human CMT-associated candidate gene, was shown to be highly associated with a severe juvenile-onset form of PN in Leonberger, as well as Saint Bernard dogs (OMIA 001917–9615) [11]. In a cohort of more than 200 affected Leonberger dogs approximately 80% of all PN cases were found to be negative for the mutant *ARHGEF10* allele [11], suggesting that PN in this breed may be a result of several genetically distinct variants having similar clinical and histopathological features. The goal of the present study was to identify additional disease-associated variants for PN in Leonberger dogs using a positional cloning strategy.

Methods

Animals

Genomic DNA was isolated from blood using either the Gentra PureGene blood kit (Qiagen) or the Maxwell RSC whole blood DNA kit (Promega). All dogs were genotyped for an *ARHGEF10* deletion (i.e., the LPN1 allele) as previously described [11], and only dogs that were homozygous wild type or heterozygous for the *ARHGEF10* allele were selected for genotyping on the SNP arrays. The phenotypic characterization of PN in Leonberger dogs has been described elsewhere [5] and the previously established criteria to select cases and controls were applied [11]. Samples from a total of 7455 Leonbergers, including 922 dogs with detailed phenotype records (Additional file 1) were used during this study. Initially, a cohort of 314 dogs was used for the GWAS; later a follow-up cohort of 608 dogs with detailed phenotype information was collected, in addition to 6533 dogs that were submitted without known phenotype status. A total of 56 PN-affected Leonberger dogs homozygous for the deletion in *ARHGEF10* were excluded from the mapping studies.

Histopathology

As part of a larger study, diagnostic peroneal nerve biopsies collected under general anesthesia or peroneal nerve samples collected at necropsy were immersed into 10% neutral buffered formalin, resin embedded and cut into thick (1 μm) sections for evaluation by a single pathologist (GDS).

Genome analyses

A total of 314 dogs (176 cases and 138 controls; Additional file 1), were genotyped on the Illumina CanineHD BeadChip (Illumina, San Diego, CA, USA) that contains 173,677 SNP markers. PLINK software [12] was used to prune the genotyping data: (1) to remove SNPs with more than 10% missing genotype calls per marker; (2) to exclude uninformative SNPs with a minor allele frequency below 5%; and (3) to exclude SNPs which exceed the Hardy-Weinberg disequilibrium p-value of 0.0001. Population stratification, resulting from close familial relationships, was confirmed by the genomic inflation factor of 1.55 calculated during an association test. Therefore, a mixed model approach utilizing the GenABEL package [13] was applied for the association analysis to correct for this population stratification and any cryptic relatedness. This correction resulted in a genomic inflation factor of 1.04. Haplotypes around the significantly associated locus were constructed using fastPHASE [14]. All genome positions refer to the CanFam3.1 reference sequence assembly.

Whole genome sequencing and variant calling

Whole genome sequence data at 12× coverage was obtained from a PN-affected Leonberger dog after preparation of a fragment library with a 250 bp insert size and collection of 293,912,640 paired-end reads (2 × 100 bp) using a HiSeq2500 instrument (Illumina). The sequence data analysis and variant calling was performed as described before [15]. The annotation version CanFam3.1.75 (http://www.ensembl.org) was used to predict the functional effects of detected variants as described previously [15]. The genome sequencing data were deposited in the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena) under sample accession number SAMEA47266168 within study accession PRJEB16012. The sequence data from the Leonberger case were compared to the Boxer reference genome (CanFam3) and 202 publically available control dogs of 66 various different breeds (Airedale Terrier, Alaskan

Husky, Alaskan Malamute, Alpine Dachsbracke, American Staffordshire Terrier, Australian Cattle dog, Australian Shepherd, Australian Terrier, Basenji, Basset, Bavarian Hound, Beagle, Bearded Collie, Berger Blanc Suisse, Border Collie, Bull Terrier, Bullmastiff, Cairn Terrier, Cavalier King Charles Spaniel, Chihuahua, Chinese indigenous dog, Chow Chow, Cocker Spaniel, Curly Coated Retriever, Dachshund, Doberman Pinscher, Elo, Entlebucher Sennenhund, Eurasier, French Bulldog, German Shepherd, German Spitz, German Wirehaired Pointer, Golden Retriever, Great Dane, Greater Swiss Mountain Dog, Greyhound, Heideterrier, Hovawart, Irish Terrier, Jack Russell Terrier, Jagdterrier, Kromfohrländer, Kunming Dog, Labrador Retriever, Lagotto Romagnolo, Landseer, Leonberger, Malinois, Miniature Bull Terrier, Norwich Terrier, Old English Sheepdog, Perro de Agua Español, Pomeranian, Poodle, Rhodesian Ridgeback, Rottweiler, Saluki, Shetland Sheepdog, Siberian Husky, Sloughi, Tibetan Mastiff, Weimaraner, West Highland White Terrier, Whippet, and Yorkshire Terrier) and 3 wolves; deposited in the European Nucleotide Archive (http://www.ebi.ac.uk/ena) under study accession numbers PRJEB10823, PRJEB13139, PRJEB13468, PRJEB13723, PRJEB14110, PRJEB14840, PRJEB16012, PRJEB4544, PRJEB5500, PRJEB5874, PRJEB5875, PRJEB6076, PRJEB6079, PRJEB7734, PRJEB7735, PRJEB7736, PRJEB7903, PRJEB9437, PRJEB9590, PRJEB9591, and PRJNA266585.

Additional genotyping

Sanger sequencing was used to confirm the Illumina sequencing results and to perform targeted genotyping for 6 variants identified from whole genome sequencing. For these experiments we amplified PCR products using AmpliTaqGold360Mastermix (Life Technologies) and purified PCR products were directly sequenced on an ABI3730 capillary sequencer (Life Technologies). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes).

The *GJA9* deletion was genotyped by fragment size analysis (primers: CCTGACAACCACAGTGGAAA (forward) and AGAGCAGTGGTTCCTTTTGC (reverse)) on an ABI3730 capillary sequencer and analyzed with the GeneMapper 4.0 software (Life Technologies).

Comparison of *GJA9* allele and genotype frequencies in PN cases and controls was performed in the original GWAS cohort, as well as an independent cohort, by standard chi-square tests.

Results

GWAS mapping of a PN associated locus on CFA15

After pruning for low genotyping rate, low minor allele frequency, and failure to meet Hardy-Weinberg equilibrium, 108,801 SNPs remained for the association analysis. A suggestive significant association signal on CFA15 was identified, with the two most significant SNPs (BICF2G630442965, at bp position 12,345,641; and BICF2G630442606 at bp position 12,765,329) achieving a corrected p-value of 1.97×10^{-6} (Fig. 1a; Additional file 2). The quantile-quantile plot of observed versus expected p-values of this mixed model analysis also supports the effectiveness of the correction for population structure and the significance of the CFA15 locus (Fig. 1b).

Haplotype reconstruction revealed evidence for an approximately 23 Mb sized disease-associated haplotype at the centromeric end of acrocentric CFA15 (Additional file 3). Five PN-affected dogs were homozygous for this haplotype spanning from bp position 0 to 22,723,949 and 37 other cases were identified as heterozygous carriers of a single copy of the identical 22.7 Mb sized haplotype. Further, 8 PN-affected dogs carried shorter versions of this haplotype due to recombination events, which enabled narrowing of the critical region to approximately 7.7 Mb (bp position 1,762,076 to 9,441,628) (Additional file 3). This region contains 135 annotated protein-encoding genes and 72 other loci.

Whole genome sequence analysis identifies a frameshift variant in the GJA9 gene

Whole genome sequence was generated from a PN-affected Leonberger homozygous for the associated CFA15 haplotype and variants in the critical interval were identified. We hypothesized that the causative variant should be absent from breeds unrelated to the Leonbergers, as pedigree analysis and the large size of the associated haplotype clearly indicated a relatively young origin of the mutation. A total of 98 variants in the interval unique to the sequenced case remained after filtering against 202 control genomes of 66 different dog breeds and 3 wolves (Additional file 4). This represented 92 homozygous private non-coding variants as well as 6 coding variants, namely: synonymous variants in *AGO3*, *CSMD2* and *GRIK3;* missense variants in *MYCL* and *MACF1* that correspond to the reference alleles in other species, and a 2 bp deletion causing a frameshift in *GJA9* (Additional file 4). All 6 coding variants were genotyped by Sanger sequencing and found to be in complete linkage disequilibrium in a cohort of 96 Leonberger dogs, of which 5 dogs were homozygous and 91 heterozygous for the *GJA9* variant.

The *GJA9* gene was pursued as the most plausible positional candidate gene due to its membership in the connexin gene family known to be causative for human CMT. The identified 2 bp deletion results in a premature stop codon at amino acid residue 382 of the 514 residue full length protein (Fig. 2). Specifically, the *GJA9* variant (CanFam3.1: chr15.3863,524_3863,525delAG) results

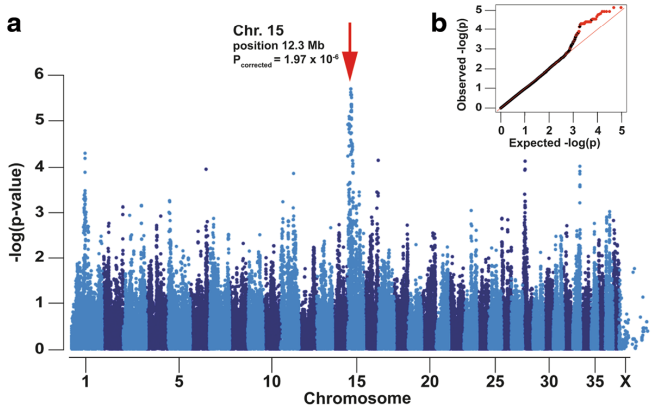

Fig. 1 Genome-wide association study for polyneuropathy in Leonberger dogs. **a** Note the association signal on chromosome 15. **b** The quantile–quantile (QQ) plot shows the observed vs. expected log *P*-values. The straight line in the QQ plot indicates the distribution of SNP markers under the null hypothesis, and the skew at the right edge indicates those markers that are more strongly associated with the trait than would be expected by chance. Markers from chromosome 15 are shown in red

in a frameshift (ENSCAFT00000038555: c.1107_1108delAG) and premature stop codon (F1PSG8_CANLF: p.Glu370AsnfsTer12) that is predicted to truncate almost half of the intracellular C-terminus of the encoded connexin. Thus we concluded that the *GJA9* deletion is much more likely to cause PN than the other 5 variants.

LPN2: A dominant inherited form of PN with incomplete penetrance

To investigate the segregation and frequency of the *GJA9* variant, we genotyped in total 7455 Leonberger dogs (Table 1). This includes all 314 Leonberger dogs from the original GWAS cohort, a follow-up cohort of 608 dogs with detailed phenotype records available (Additional file 1), and 6533 mostly young Leonbergers with unknown phenotype status that were submitted for diagnostic purposes, which represent the global population of the breed. There was a highly significant difference in *GJA9* allele frequencies between PN cases and controls from the original GWAS cohort ($p < 5.5 \times 10^{-8}$), reflected in a difference in genotype frequencies under a dominant model ($p < 1.2 \times 10^{-7}$), but not a recessive model ($p < 0.05$) (Table 1). These figures give only borderline support for a recessive hypothesis and make it much less well supported than a dominant hypothesis. This result was replicated in the follow-up cohort (genotype frequency different at $p < 1.9 \times 10^{-7}$, genotype frequency different in a dominant model at $p < 1.7 \times 10^{-7}$, but not a recessive model at $p < 0.15$) (Table 1). The allele frequency difference in these combined cohorts was significant at $p < 2.8 \times 10^{-14}$ and the genotype frequency difference under the dominant model was significant at $p < 6.8 \times 10^{-14}$, clearly indicating dominant inheritance.

We then analyzed 137 dogs that carried at least one copy of the *GJA9* mutant allele to determine age of onset and penetrance of the mutant allele (Additional file 1). The owner-reported mean average age of onset in 100 dogs that showed clinical signs was 6 years, and varied from 1 to 10 years (Fig. 3). Due to the limited number of *GJA9* homozygous mutant PN-affected dogs it was not possible to assess whether heterozygous or homozygous PN cases developed clinical signs at different stages in life. Both homozygous mutant and heterozygous dogs of both sexes developed PN, again consistent with an autosomal dominant pattern of inheritance. By 8 years of age 84 of the 137 (61%) dogs carrying the mutant allele showed signs of PN, and within their lifetime 110 of the 137 (80%) showed signs, indicative of a reduced penetrance (Additional file 1). Lastly, the frequency of the *GJA9* variant in the global population of Leonbergers was estimated as 0.032 (Table 1).

A total of 528 Leonbergers in our sample database were diagnosed with PN when the 56 additional PN-affected dogs homozygous for the deletion in *ARHGEF10* were included. Thus, the disease phenotype of approximately 11% (56/528) of all PN-affected Leonbergers could be explained by homozygosity for the *ARHGEF10* (LPN1) variant, whereas homozygosity or heterozygosity for the herein presented PN-associated variant in the *GJA9* gene, which we now term LPN2, explains 21% (110/528) of all cases. Of the remaining 362 PN cases, 85 are LPN1 heterozygous and 277 PN-affected dogs (52%) do not carry either known variant (Additional file 1). Because it is still unclear whether

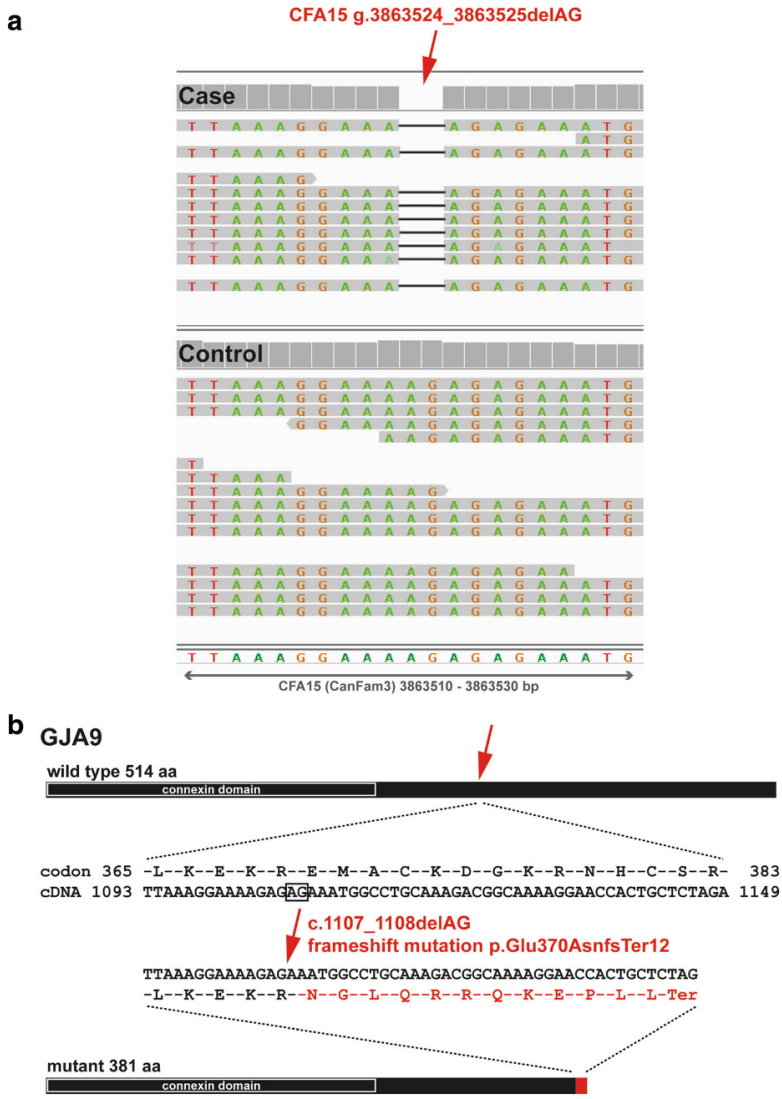

Fig. 2 A *GJA9* frameshift variant in polyneuropathy affected Leonberger dogs. **a** IGV screenshots showing the homozygous presence of the 2 bp deletion in the *GJA9* gene on CFA15 in reads from the sequenced case (marked by a red arrow) compared to a control dog genome. **b** The schematic representation indicates that the 2 bp deletion leads to a frameshift that is predicted to produce a significantly truncated intracellular C-terminus of canine GJA9. Note that the transmembranal N-terminal connexin domain that is important for gap junction formation is predicted to be unaffected

having one copy of the LPN1 allele is sufficient to cause disease, we examined the LPN1 allele frequency in remaining cases (0.12) compared to controls (0.07). This provides evidence of some additional genotypic risk associated with the possession of a single copy of the *ARHGEF10* variant, albeit at a level below 2. Altogether, these allele frequencies continue to present ambiguity ($p = 6.7 \times 10^{-4}$) in interpreting the correct mode of inheritance of the LPN1 allele (Additional file 1).

Indistinguishable histopathology in PN affected Leonberger dogs

Resin sections from the peroneal nerve were qualitatively evaluated from 5 Leonberger dogs with PN and

Table 1 *GJA9* c.1103_1104delAG genotype frequencies in three Leonberger cohorts

Polyneuropathy status	Total	wt/wt (homozygous normal; N/N)	wt/del (heterozygous; D/N)	del/del (homozygous mutant; D/D)
GWAS cohort	314	0.82 (258)	0.16 (51)	0.02 (5)
Affected[a]	176	0.72 (127)	0.25 (44)	0.03 (5)
Non-affected[a,b]	138	0.95 (131)	0.05 (7)	0
Follow-up cohort	608	0.87 (527)	0.13 (79)	0.003 (2)
Affected[a]	296	0.79 (235)	0.20 (59)	0.006 (2)
Non-affected[a,b]	312	0.94 (292)	0.06 (20)	0
Total	7455	0.94 (6995)	0.06 (443)	0.002 (17)
Affected[a]	474	0.76 (362)	0.22 (105)	0.02 (7)
Non-affected[a,b]	450	0.94 (423)	0.06 (27)	0
Unknown[c]	6533	0.95 (6210)	0.05 (313)	0.002 (10)

[a]Only dogs that were homozygous wild type or heterozygous *ARHGEF10* carriers
[b]Only dogs ≥8 years old that showed no signs of PN
[c]Dogs without known phenotype status that were submitted for diagnostic purposes

the *GJA9* variant (LPN2 heterozygous) and were compared to Leonberger dogs with PN and the *ARHGEF10* variant (LPN1 homozygous). None of the sections evaluated were from dogs positive for both the *GJA9* and *ARHGEF10* variants. The prominent pathologic abnormality was variably severe nerve fiber loss resulting from chronic axonal degeneration (Fig. 4). Large nerve fiber loss was most prominent with an increased population of small caliber nerve fibers. Similar changes were present in all nerves regardless of the genetic variant.

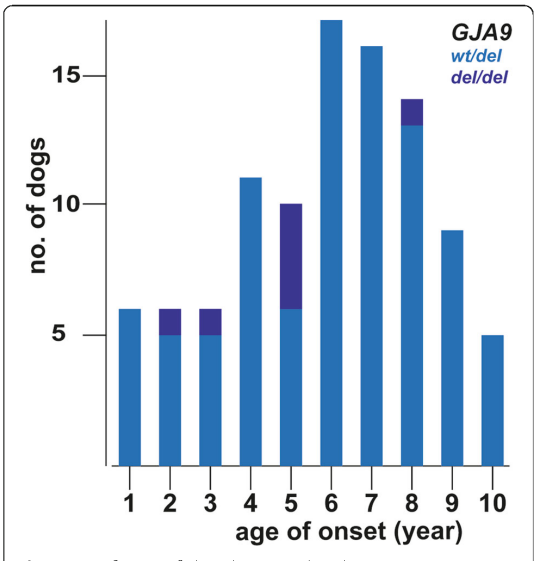

Fig. 3 Age of onset of clinical signs in dogs homozygous or heterozygous for the *GJA9* mutation. Age-of-onset of clinical signs for the 100 PN cases in which one or two mutant *GJA9* alleles were present

Discussion

Inherited canine PN has clinical similarities to the genetically heterogeneous CMT peripheral neuropathies in people. PN in the Leonberger breed shows a variable expression of clinical signs and different ages of onset from <1 year up to 10 years of age. The previously identified *ARHGEF10* frame shift deletion (termed LPN1) causes a severe early-onset disease [11]. Initially, the LPN1 variant explained over 20% of cases; however, due to widespread adoption of genetic testing by breeders, we have not observed a LPN1 case born since 2011. Consequently, the proportion of PN cases explained by the LPN1 mutation has dropped precipitously; currently it explains only approximately 11% of PN cases. Here, our search to identify additional PN genes employed a GWAS with 176 PN affected Leonbergers, which were not homozygous carriers of the recessive *ARHGEF10* loss-of-function variant, and 138 controls. Our present finding of the *GJA9* variant strongly suggests a causative allele for a second inherited form of PN (LPN2), with a wider range of ages of onset and clinical severity, which explains an additional 21% of PN in this breed. These results further confirm genetic similarities between Mendelian forms of PN in dogs and humans. The characteristic axonal degeneration and nerve fiber loss affecting long, large diameter nerves in PN-affected Leonberger dogs resembled mixed forms of human CMT [5]. While qualitative evaluation of peripheral nerve histopathology can confirm the presence of chronic axonal degeneration, nerves can only respond in a limited number of ways to insult of any type and may not be of value in distinguishing LPN variants in Leonberger dogs. Quantitative studies are in progress. The identification of the *ARHGEF10* variant finally confirmed the previously assumed

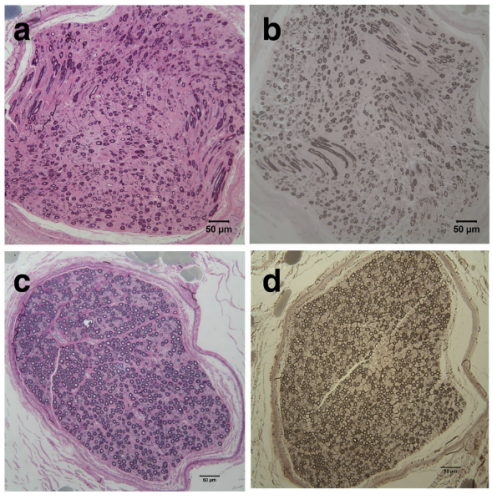

Fig. 4 Representative histopathology of the peroneal nerve from male Leonberger dog with polyneuropathy and heterozygous *GJA9* genotype. Resin embedded peroneal nerve sections collected post-mortem from a 5 year old male intact Leonberger dog with clinical signs of polyneuropathy. **a** Toluidine blue-acid fuchsin stain showing loss of predominantly large myelinated fibers and qualitatively increased populations of small calibre nerve fibers. **b** Paraphenelenediamine stain for myelin showing thin but intact myelin sheaths in many small nerve fibers. For comparison, images (**c**, toluidine blue-acid fuchsin stain) and (**d**, paraphenelenediamine stain) are from the peroneal nerve of a 6 year old female neutered Leonberger dog that was negative for both the *ARHGEF10* and *GJA9* variants and without clinical signs of polyneuropathy

similar pathogenesis of PN in Leonberger dogs to CMT disease in people.

The herein identified 7.7 Mb-sized disease-associated haplotype contained a total of 98 private variants. All 6 private coding variants were experimentally confirmed to occur in complete linkage disequilibrium, the additional 92 non-coding variants were not analyzed further, because the frameshift variant in *GJA9* affects a putative candidate gene, and the variants in *MYCL* and *MACF1* can most likely be ruled out by non-conservation. *GJA9* encodes connexin 59, a connexin gap junction family protein, whose members have emerged as important components of peripheral myelinated nerve fibers. A total of 21 human connexin genes are identified (20 in the mouse), which can be classified into 5 groups based on sequence homology [16]. Six connexin subunits oligomerize to form a hemichannel (also known as connexon) and hemichannels from two adjacent cells dock together at their extracellular domains to form a functional gap junction channel that mediates direct intercellular communication in many physiological processes [16]. Cells and tissues commonly express more than one type of connexin, enabling the formation of homomeric and heteromeric gap junctions [17], and junctional communication between different cells often requires such heterotypic gap junctions [16].

Diverse genetic diseases in humans, including disorders of the nervous system, are caused by variants in connexin genes [18]. The *GJB1* gene, with more than 400 known variants, is causally implicated in one of the five most common CMT subtypes in people [OMIM 304040; 19,20]. The encoded connexin 32 forms gap junctions between the folds of Schwann cell membranes. These Schwann cells wrap around the axons of peripheral nerves and form a layer of myelin that is critical for the conduction of nerve impulses. Mutations in *GJB1* appear to hinder diffusion across the concentric layers of myelin, resulting in disruption of myelin and hence axonal degeneration [19, 20]. Due to the male sex bias, similar phenotype, and the slowly progressive nature of the PN disease in Leonberger dogs, it was initially speculated, that variants in the canine ortholog of *GJB1* encoding connexin 32 could be responsible [21]. Our identification of a variant in the related gap junction protein GJA9 nicely confirms these earlier assumptions. Interestingly, as we have observed in our canine study, there is a high phenotypic variability within CMT genotypes and variant-specific manifestations between different human CMT types [22]. Recently, the human *GJA9* gene, among others, was reported as novel candidate disease gene for neurogenetic disorders after the first results from clinical exome sequencing of 149 patients with various neurocognitive phenotypes [23]. But, to our knowledge, there is as yet no non-synonymous *GJA9* variant known to be associated with CMT in people. Therefore this canine *GJA9* associated form of PN provides an interesting animal model and adds this gene to the list of candidates for human CMT. This is especially of value as there is no *GJA9* ortholog in the mouse genome [24]. Finally, in the light of the current status of the dog reference genome and its incomplete annotation, we acknowledge that the list of private variants of the disease-associated haplotype is most likely not complete. In addition to the fact that we were not able to exclude the reported 97 private variants other than the *GJA9* frameshift variant, one could assume that other undetected variants located in existing reference sequence gaps or unrecognized structural variants exist on the disease-associated haplotype.

The age-of-onset analysis indicates that LPN2 is a dominantly inherited trait with incomplete penetrance. Several heterozygous dogs >8 years of age still have not shown clinical signs of weakness. This might suggest that nerve cells can still function reasonably well with a single copy of the normal allele, or even the absence of the GJA9 protein entirely. The predicted protein from

the mutant allele contains <75% of the amino acids present in the wild type protein. It is possible that since the four N-terminal transmembrane domains forming the connexin channel domain are unchanged, the identified variant with its significantly shortened intracellular C-terminus may not be a complete loss-of-function allele. Other mechanisms to explain the somewhat complex genotype-age of onset-phenotype relationship include the possibilities that other connexins partially serve the function of GJA9, a dominant-negative effect of the mutant protein, or simple haploinsufficiency affected by genetic background. We further suggest that the incomplete association between *GJA9* genotype and PN in our large population is partly due to contributions from phenocopies, misdiagnoses, as well as genetic heterogeneity.

These findings have serious implications for breeders, as genetically susceptible dogs may not develop clinical disease until later in life, if at all, and often well after a dog may have bred multiple times. Since the offering of the LPN2 test to the Leonberger community nearly 7500 dogs have been genotyped for the *GJA9* deletion. Although this is not a random sample of the breed, the finding of a mutant allele frequency of approximately 0.03, and its dominant nature, indicates it should be considered a serious problem for the health and welfare of breed.

Conclusions

Highly probable causative variants in the *ARHGEF10* and *GJA9* genes have now been identified for two forms of PN in Leonberger dogs. PN in this breed exhibits remarkable variation in phenotypic severity and age of onset among affected dogs, suggesting the influence of modifiers of both the *GJA9* form and additional forms not explained by either known variant. It is likely that many more Leonberger dogs with PN will be needed to improve the power to detect loci that contribute to the non-*ARHGEF10* /*GJA9* form(s) of PN in Leonbergers, which might represent a multigenic complex trait. Characterization of genomic variation underlying PN phenotypic variation in this breed with small effective population size will present a powerful resource for understanding the molecular causes behind variable canine polyneuropathy phenotypes. Finally, this study adds *GJA9* for the first time to the list of candidate genes for human CMT.

Additional files

Additional file 1: Phenotype records and GJA9 genotypes of 922 Leonberger dogs with detailed phenotype records. (XLSX 469 kb)

Additional file 2: GWAS results comparing 176 cases vs. 138 controls. (XLSX 11371 kb)

Additional file 3: Haplotype reconstruction for the centromeric region of acrocentric CFA15 of 314 dogs. (XLSX 3145 kb)

Additional file 4: List of 98 private sequence variants of the sequenced PN-affected Leonberger in the critical region on CFA15. (XLSX 31 kb)

Abbreviations

BAM: Binary version of a sequence alignment/map (SAM) file; bp: Base pairs; CMT: Charcot-Marie-Tooth; GJA9: Gap junction protein alpha 9; GWAS: Genome wide association study; IGV: Integrative Genomics Viewer; LPN1: Leonberger polyneuropathy type 1; LPN2: Leonberger polyneuropathy type 2; OMIA: Online Mendelian Inheritance in Animals; PCR: Polymerase chain reaction; SNP: Single nucleotide polymorphism

Acknowledgements

The authors gratefully acknowledge Paul Mandigers and Kaspar Matiasek for additional Leonberger samples, and all Leonberger breeders and the Health Committee of the International Union for Leonberger Dogs for their tremendous support during the entire project. The authors wish to thank Brigitta Colomb, Nathalie Besuchet-Schmutz, Muriel Fragnière, Monica Burgers, and Jill Pesayco for their invaluable technical assistance. The Next Generation Sequencing Platform of the University of Bern is acknowledged for performing the whole genome re-sequencing experiment and the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure. We thank the Dog Biomedical Variant Database Consortium (Gus Aguirre, Catherine André, Danika Bannasch, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Hannes Lohi, Cathryn Mellersh, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Kim Summers, Frank van Steenbeck, Claire Wade) for sharing whole genome sequencing data from control dogs. We also acknowledge all canine researchers who deposited dog whole genome sequencing data into public databases.

Funding

Not applicable.

Availability of data and materials

The genome data has been made freely available under sample accession number SAMEA47266168 at the European Nucleotide Archive [http://www.ebi.ac.uk/ena/data/view/PRJEB16012].

Authors' contributions

DB, KMM, AL, KJE performed phenotype assessment, GWAS mapping and genotyping. GDS evaluated biopsies of peripheral nerves. VJ performed bioinformatics. TL, JRM and CD designed the study. CD and JRM drafted the paper. All authors participated in writing the manuscript and have read and approved the final version.

Ethics approval and consent to participate

This study was performed using protocols approved by the Institutional Animal Care and Use Committees (IACUC) of the University of Minnesota (UM), by the "Cantonal Committee For Animal Experiments" (Canton of Bern; permits 23/10, 48/13 and 75/16) for the University of Bern (UB), and by the Institutional Animal Care and Use Committees (IACUC) of the University of California San Diego (UCSD). Written consent was obtained from all dogs' owners. Leonberger samples were obtained primarily via elective owner submission or were submitted for genotyping of the previously reported *ARHGEF10* variant. Most of these dogs do not have complete medical information and were used only for a population study. Samples submitted to UCSD additionally included nerve and/or muscle biopsies for diagnostic purposes.

Consent for publication

Not applicable.

Competing interests

Both the University of Minnesota and the University of Bern are offering a genotyping test for LPN2 in their respective laboratories and proceeds from these tests go toward ongoing canine genetic research. The authors declare that no other competing interests exist.

Publisher's Note

Author details
[1]Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland. [2]Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, St. Paul, MN 55108, USA. [3]Department of Basic Medical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, Indiana 47907, USA. [4]Department of Pathology, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA.

Received: 20 March 2017 Accepted: 21 August 2017
Published online: 25 August 2017

References

1. Granger N. Canine inherited motor and sensory neuropathies: an updated classification in 22 breeds and comparison to Charcot-Marie-tooth disease. Vet J. 2011;188:274–85.
2. Braund KG, Steinberg HS, Shores A, Steiss JE, Mehta JR, Toivio-Kinnucan M, Amling KA. Laryngeal paralysis in immature and mature dogs as one sign of a more diffuse polyneuropathy. J Am Vet Med Assoc. 1989;194:1735–40.
3. Bookbinder LC, Flanders J, Bookbinder PF, Harvey HJ, Barry JS, Cheetham J. Idiopathic canine laryngeal paralysis as one sign of a diffuse polyneuropathy: an observational study of 90 cases (2007-2013). Vet Surg. 2016;45:254–60.
4. Hultin Jäderlund K, Baranowska Körberg I, Nødtvedt A. Inherited polyneuropathy in Leonberger dogs. J Vet Intern Med. 2011;25:997–1002.
5. Shelton GD, Podell M, Poncelet L, Schatzberg S, Patterson E, Powell HC, Mizisin AP. Inherited polyneuropathy in Leonberger dogs: a mixed or intermediate form of Charcot-Marie-tooth disease? Muscle Nerve. 2003;27:471–7.
6. Drögemüller C, Becker D, Kessler B, Kemter E, Tetens J, Jurina K, Jäderlund KH, Flagstad A, Perloski M, Lindblad-Toh K, Matiasek K. A deletion in the N-myc downstream regulated gene 1 (NDRG1) gene in greyhounds with polyneuropathy. PLoS One. 2010;5:e11258.
7. Bruun CS, Jäderlund KH, Berendt M, Jensen KB, Spodsberg EH, Gredal H, Shelton GD, Mickelson JR, Minor KM, Lohi H, Bjerkås I, Stigen O, Espenes A, Rohdin C, Edlund R, Ohlsson J, Cizinauskas S, Leifsson PS, Drögemüller C, Moe L, Cirera S, Fredholm M. A Gly98Val mutation in the N-Myc downstream regulated gene 1 (NDRG1) in Alaskan malamutes with polyneuropathy. PLoS One. 2013;8:e54547.
8. Mhlanga-Mutangadura T, Johnson GS, Schnabel RD, Taylor JF, Johnson GC, Katz ML, Shelton GD, Lever TE, Giuliano E, Granger N, Shomper J, O'Brien DP. A mutation in the Warburg syndrome gene, RAB3GAP1, causes a similar syndrome with polyneuropathy and neuronal vacuolation in black Russian terrier dogs. Neurobiol Dis. 2016;86:75–85.
9. Mhlanga-Mutangadura T, Johnson GS, Ashwini A, Shelton GD, Wennogle SA, Johnson GC, Kuroki K, O'Brien DP. A homozygous RAB3GAP1:c.743delC mutation in Rottweilers with neuronal Vacuolation and Spinocerebellar degeneration. J Vet Intern Med. 2016;30:813–8.
10. Wiedmer M, Oevermann A, Borer-Germann SE, Gorgas D, Shelton GD, Drögemüller M, Jagannathan V, Henke D, Leeb T. A RAB3GAP1 SINE Insertion in Alaskan Huskies with Polyneuropathy, Ocular Abnormalities, and Neuronal Vacuolation (POANV) Resembling Human Warburg Micro Syndrome 1 (WARBM1). G3 (Bethesda). 2015;6:255–62.
11. Ekenstedt KJ, Becker D, Minor KM, Shelton GD, Patterson EE, Bley T, Oevermann A, Bilzer T, Leeb T, Drögemüller C, Mickelson JR. An ARHGEF10 deletion is highly associated with a juvenile-onset inherited polyneuropathy in Leonberger and Saint Bernard dogs. PLoS Genet. 2014;10:e1004635.
12. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.
13. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM. GenABEL: an R library for genome-wide association analysis. Bioinformatics. 2007;23:1294–6.
14. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum Genet. 2006;78:629–44.
15. Kyöstilä K, Syrjä P, Jagannathan V, Chandrasekar G, Jokinen TS, Seppälä EH, Becker D, Drögemüller M, Dietschi E, Drögemüller C, Lang J, Steffen F, Rohdin C, Jäderlund KH, Lappalainen AK, Hahn K, Wohlsein P, Baumgärtner W, Henke D, Oevermann A, Kere J, Lohi H, Leeb T. A missense change in the ATG4D gene links aberrant autophagy to a neurodegenerative vacuolar storage disease. PLoS Genet. 2015;11:e1005169.
16. Bai D. Structural analysis of key gap junction domains–lessons from genome data and disease-linked mutants. Semin Cell Dev Biol. 2016;50:74–82.
17. Saez JC, Berthoud VM, Branes MC, Martinez AD, Beyer EC. Plasma membrane channels formed by connexins: their regulation and functions. Physiol Rev. 2003;83:1359–400.
18. Abrams CK, Scherer SS. Gap junctions in inherited human disorders of the central nervous system. Biochim Biophys Acta. 1818;2012:2030–47.
19. Kleopa KA, Sargiannidou I. Connexins, gap junctions and peripheral neuropathy. Neurosci Lett. 2015;596:27–32.
20. Fridman V, Bundy B, Reilly MM, Pareyson D, Bacon C, Burns J, Day J, Feely S, Finkel RS, Grider T, Kirk CA, Herrmann DN, Laurá M, Li J, Lloyd T, Sumner CJ, Muntoni F, Piscosquito G, Ramchandren S, Shy R, Siskind CE, Yum SW, Moroni I, Pagliano E, Zuchner S, Scherer SS, Shy ME. Inherited Neuropathies Consortium.CMT subtypes and disease burden in patients enrolled in the Inherited Neuropathies Consortium natural history study: a cross-sectional analysis. J Neurol Neurosurg Psychiatry. 2015;86:873–8.
21. Granger N, Escriou C, Thibaud JL, Riche C, Talbot K, Jeffery ND, Blot S. Polyneuropathy in Leonberger dogs: an emerging pan-Europea polyneuropathy. Birmingham, UK: Proceedings BSAVA Congress; 2007. p. 488–9.
22. Cornett KM, Menezes MP, Bray P, Halaki M, Shy RR, Yum SW, Estilow T, Moroni I, Foscan M, Pagliano E, Pareyson D, Laurá M, Bhandari T, Muntoni F, Reilly MM, Finkel RS, Sowden J, Eichinger KJ, Herrmann DN, Shy ME, Burns J. Inherited Neuropathies Consortium Phenotypic Variability of Childhood Charcot-Marie-Tooth Disease. JAMA Neurol. 2016;73:645–51.
23. Gripp KW, Adam MP, Hudgins L, Carey JC. 36th annual David W. Smith workshop on malformations and morphogenesis: abstracts of the 2015 annual meeting. Am J Med Genet A. 2016;170:1665–726.
24. Söhl G, Nielsen PA, Eiberger J, Willecke K. Expression profiles of the novel human connexin genes hCx30.2, hCx40.1, and hCx62 differ from their putative mouse orthologues. Cell Commun Adhes. 2003;10:27–36.

A *CNTNAP1* missense variant is associated with canine laryngeal paralysis and polyneuropathy

Journal:	Genes
Manuscript status:	published
Contributions:	Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing

genes

MDPI

Article

A *CNTNAP1* Missense Variant Is Associated with Canine Laryngeal Paralysis and Polyneuropathy

Anna Letko [1,*,†], Katie M. Minor [2,†], Steven G. Friedenberg [3], G. Diane Shelton [4], Jill Pesayco Salvador [4], Paul J. J. Mandigers [5], Peter A. J. Leegwater [5], Paige A. Winkler [6], Simon M. Petersen-Jones [6], Bryden J. Stanley [6], Kari J. Ekenstedt [7], Gary S. Johnson [8], Liz Hansen [8], Vidhya Jagannathan [1], James R. Mickelson [2,‡] and Cord Drögemüller [1,‡]

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; vidhya.jagannathan@vetsuisse.unibe.ch (V.J.); cord.droegemueller@vetsuisse.unibe.ch (C.D.)
2 Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, Saint Paul, MN 55108, USA; minork@umn.edu (K.M.M.); micke001@umn.edu (J.R.M.)
3 Department of Veterinary Clinical Sciences, College of Veterinary Medicine, University of Minnesota, Saint Paul, MN 55108, USA; fried255@umn.edu
4 Department of Pathology, School of Medicine, University of California San Diego, La Jolla, CA 92093-0709, USA; gshelton@ucsd.edu (G.D.S.); jpesayco@ucsd.edu (J.P.S.)
5 Department of Clinical Sciences, Utrecht University, 3584 CM Utrecht, The Netherlands; p.j.j.mandigers@veterinair-neuroloog.nl (P.J.J.M.); P.A.J.Leegwater@uu.nl (P.A.J.L.)
6 Department of Small Animal Clinical Sciences, College of Veterinary Medicine, Michigan State University, East Lansing, MI 48824, USA; winkler.paige@gmail.com (P.A.W.); peter315@msu.edu (S.M.P.-J.); stanle32@msu.edu (B.J.S.)
7 Department of Basic Medical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN 47907, USA; kje0003@purdue.edu
8 Department of Veterinary Pathobiology, University of Missouri, Columbia, MO 65211, USA; JohnsonGS@missouri.edu (G.S.J.); HansenL@missouri.edu (L.H.)
* Correspondence: anna.letko@vetsuisse.unibe.ch; Tel.: +41-31-631-25-29
† These authors contributed equally to the work.
‡ James R. Mickelson and Cord Drögemüller shared last author.

Received: 28 October 2020; Accepted: 26 November 2020; Published: 27 November 2020

Abstract: Laryngeal paralysis associated with a generalized polyneuropathy (LPPN) most commonly exists in geriatric dogs from a variety of large and giant breeds. The purpose of this study was to discover the underlying genetic and molecular mechanisms in a younger-onset form of this neurodegenerative disease seen in two closely related giant dog breeds, the Leonberger and Saint Bernard. Neuropathology of an affected dog from each breed showed variable nerve fiber loss and scattered inappropriately thin myelinated fibers. Using across-breed genome-wide association, haplotype analysis, and whole-genome sequencing, we identified a missense variant in the *CNTNAP1* gene (c.2810G>A; p.Gly937Glu) in which homozygotes in both studied breeds are affected. *CNTNAP1* encodes a contactin-associated protein important for organization of myelinated axons. The herein described likely pathogenic *CNTNAP1* variant occurs in unrelated breeds at variable frequencies. Individual homozygous mutant LPPN-affected Labrador retrievers that were on average four years younger than dogs affected by geriatric onset laryngeal paralysis polyneuropathy could be explained by this variant. Pathologic changes in a Labrador retriever nerve biopsy from a homozygous mutant dog were similar to those of the Leonberger and Saint Bernard. The impact of this variant on health in English bulldogs and Irish terriers, two breeds with higher *CNTNAP1* variant allele frequencies, remains unclear. Pathogenic variants in *CNTNAP1* have previously been reported in human patients with lethal congenital contracture syndrome and hypomyelinating neuropathy, including vocal cord palsy and severe respiratory distress. This is the first report of contactin-associated LPPN in dogs characterized by a deleterious variant that most likely predates modern breed establishment.

Keywords: *Canis familiaris*; whole-genome sequencing; rare disease; contactin; neurological disorder; Leonberger; Saint Bernard; Labrador retriever

1. Introduction

Laryngeal paralysis (LP) can result from trauma or neoplasia involving the recurrent laryngeal nerves, peripheral nerve disease, or a primary or secondary disease affecting the muscle or neuromuscular junction. Loss of normal function of the larynx leads to breathing difficulties, reduced exercise and heat tolerance, as well as an increased risk of aspiration pneumonia [1]. Laryngeal nerve disease results in degeneration and atrophy of intrinsic laryngeal muscles followed by decreased or absent movement of the attendant laryngeal cartilages. During breathing, these cartilages control airflow into and out of the trachea. Affected dogs have stridor, may have a change in vocalization, and difficulty breathing due to the flaccid laryngeal vocal folds and corniculate processes of the arytenoid obstructing the lumen of the airway [1]. Normal laryngeal function protects the airway by closing off the lumen to prevent aspiration of food or water. In LP-affected dogs, the vocal folds remain in a paramedian position, causing airway resistance and turbulence, instead of abducting, as they normally would, to open the airway during inspiration. Frequently, affected dogs suffering from LP are treated by crico- or thyro-arytenoid laryngoplasty surgery, to improve breathing and, therefore, quality of life [2]. As the recurrent laryngeal nerve axons are some of the longest in the body [3], LP is often reported as part of a more generalized length-dependent polyneuropathy (PN) complex, which manifests with additional signs including proprioceptive and motor abnormalities, slowly progressing pelvic limb weakness, and loss of limb muscle mass [4].

Various mostly breed-specific canine inherited neuropathies form a heterogeneous group of degenerative diseases affecting motor and/or sensory and autonomic peripheral nerves. This group includes mixed forms of LP and PN [5], i.e., the laryngeal paralysis and polyneuropathy complex (LPPN), which has variable ages of onset among and across several dog breeds (OMIA 001206-9615, OMIA 001292-9615). Late-onset forms, e.g., geriatric onset laryngeal paralysis polyneuropathy (GOLPP), are also observed in various breeds including Labrador retrievers [6]. Leonberger dogs are known to be susceptible to LPPN; recently, a short list of potentially pathogenic variants for neurological disorders in this breed derived from whole-genome sequencing has been presented [7]. To date, variants in *ARHGEF10* [8] and *GJA9* [9] have already been associated with certain forms of the disorder and designated with breed-specific names Leonberger polyneuropathy type 1 (LPN1; OMIA 001917-9615) and Leonberger polyneuropathy type 2 (LPN2; OMIA 002119-9615), respectively. These two variants, however, do not explain all the phenotypically described cases in Leonbergers [7]. The *ARHGEF10* variant has also been reported in the related Saint Bernard breed, but again it did not explain all LPPN cases [8]. Alaskan huskies, black Russian terriers, and Rottweilers with PN including LP and respiratory distress are known to have deleterious variants in the *RAB3GAP1* gene, a member of the RAB3 protein family implicated in regulated exocytosis of neurotransmitters and hormones (OMIA 001970-9615) [10–12]. Another major risk factor for canine LP recently described in miniature bull terriers and bull terriers is a variant in the *RAPGEF6* gene encoding a widely expressed nucleotide exchange factor whose function is not well understood (OMIA 002222-9615) [13].

In general, there are limits to precisely diagnosing neurological diseases in dogs in the clinic. For example, in a previous study [14], we noticed Leonbergers that were initially clinically diagnosed as polyneuropathy-affected, although, in fact, they were suffering from leukoencephalomyelopathy, a juvenile-onset neurodegenerative disorder of the CNS white matter with distinctive pathological features, caused by a recessive variant in the *NAPEPLD* gene (OMIA 001788-9615).

Our aim in this study was to identify additional causative genetic variants associated with younger-onset laryngeal paralysis and polyneuropathy (LPPN), by focusing on two closely related giant dog breeds [15], namely the Leonberger and Saint Bernard.

2. Materials and Methods

2.1. Ethics Statement

All animal experiments were performed according to local regulations, and all animals in this study were examined with the consent of their owners. The study was approved under IACUC protocol 1903-36865A at the University of Minnesota, the Michigan State University Institutional Animal Care and Use Committee (AUF number 01/11-009-00), and by the Cantonal Committee for Animal Experiments (Canton of Bern; permit 71/19) at the University of Bern.

2.2. Animal Selection

Data on 15,378 dogs from 243 breeds, 321 dogs of mixed or unknown heritage, and 62 wild canids were collected in three different sets for this study (Table 1 and Supplementary Table S1). The discovery cohort included 426 Leonbergers either showing signs of LPPN with an age of onset ≤5 years or healthy control dogs at ≥8 years of age, and 91 Saint Bernards either showing signs of LPPN with an age of onset ≤5 years or population control dogs with genome-wide association study (GWAS) data available from unrelated studies [15,16]. All 517 Leonbergers and Saint Bernards were genotyped for the *ARHGEF10* variant [8], and the Leonbergers were also genotyped for the *GJA9* variant [9], in order to include only dogs homozygous wild type for these loci. In addition, the Leonbergers were genotyped for a previously reported leukoencephalopathy-associated *NAPEPLD* variant [14], in order to rule out another known underlying neurological disease with similar clinical phenotype.

Table 1. Number of canids in each studied cohort.

Cohort [1]	Dog Breed/Species	Total	Phenotype		
			LPPN-Affected [2]	LPPN Non-Affected [3]	Unknown
Discovery (n = 517)	Leonberger	426	126	300	0
	Saint Bernard	91	18	14	59
Validation (n = 1070)	Leonberger	859	500	359	0
	Saint Bernard	11	11	0	0
	Labrador retriever [4]	200	150	50	0
Population (n = 14,174)	243 dog breeds	13,798	0	0	13,798
	Unknown/mixed heritage	314	0	0	314
	Wolf	58	0	0	58
	Golden jackal	2	0	0	2
	Andean fox	1	0	0	1
	Dhole	1	0	0	1

[1] Additional details including the source of all data are available in Supplementary Materials Table S1. [2] Includes affected dogs with previously described laryngeal paralysis and polyneuropathy (LPPN)-associated variants in *ARHGEF10* [8] and *GJA9* [9]. [3] Includes dogs homozygous for the previously described leukoencephalomyelopathy-associated variant in *NAPEPLD* [14]. [4] Includes seven mixed-breed dogs enrolled in the Michigan State University's GOLPP study [6].

A validation cohort used for targeted genotyping of the newly discovered variant consisted of 1070 dogs with known LPPN phenotypes (Table 1 and Supplementary Table S1). There was no age of onset restriction for the cases in the validation cohort. Included in the validation cohort were 193 Labrador retrievers and seven mixed-breed dogs from an ongoing geriatric onset laryngeal paralysis polyneuropathy (GOLPP) study; these were used as an independent validation group (Table 1).

Finally, a population cohort consisting of 14,112 dogs, 58 wolves, two golden jackals, one Andean fox, and one dhole (Table 1 and Supplementary Table S1), with no available information about their health status, was used to determine the absence/presence and frequency of the described variant-associated haplotype across canids.

The information about age of onset of clinical signs in the LPPN-affected dogs was available for a subset of 770 dogs from the discovery and validation cohorts with detailed health information from three breeds (596 Leonbergers, 28 Saint Bernards, and 146 Labrador retrievers) The statistical significance of the differences between groups was evaluated with Student's t-test and $p < 0.05$ was considered as significant.

2.3. Sample Preparation

Genomic DNA was isolated from EDTA blood samples, buccal swabs, or archived muscle biopsies by using either the Gentra PureGene kit (Qiagen, Hilden, Germany) or the Maxwell RSC Whole Blood DNA kit (Promega, Dübendorf, Switzerland).

Clinical cases of polyneuropathy and laryngeal paralysis were evaluated in three dogs homozygous for the studied *CNTNAP1* variant with available nerve biopsies; these included a 3-year-old Saint Bernard, a 3-year-old Leonberger, and a 9-year-old Labrador retriever (Supplementary Table S1). Three normal adult dog samples from Labrador retrievers were used as controls. The ages for the control dogs were 8–10 years of age. All the archived nerve specimens were obtained years prior to the identification of the *CNTNAP1* variant. Peroneal nerve specimens, pinned on cork discs to maintain length and orientation, were immersion-fixed in 2.5% glutaraldehyde in 0.1 M phosphate buffer before shipment. Upon receipt, the nerves were postfixed in 1% aqueous osmium tetroxide for 3 to 4 h before dehydration in a graded alcohol series and propylene oxide. After infiltration with a 1:1 mixture of propylene oxide and araldite resin for 4 h, nerves were placed in 100% araldite resin overnight before embedding in fresh araldite resin. Thick sections (1 µm) were cut with glass knives and either stained with toluidine blue prior to light microscopic examination, or stained with paraphenylenediamine prior to morphometry.

2.4. Axonal Size Frequency Distributions and G-Ratios

Axonal size-frequency distributions of myelinated fibers were performed on transverse sections of selected peroneal nerve biopsies determined to be adequately fixed, free from artifact, and with an intact perineurium. Images were obtained from a single section of each nerve biopsy using the Photoshop image analysis system. Profiles containing paranodal regions or Schmidt–Lanterman clefts were not included. Using a ×60 objective, the final magnification of the digitized image was equivalent to 1 pixel = 0.091 µm. Myelinated fibers were individually identified and selected prior to being sorted with an automated process into bins based on axonal area. G-ratios were calculated as the ratio between the diameter of the axon itself and the outer diameter of the myelinated fiber.

2.5. Single Nucleotide Polymorphism Array Genotyping and Imputation

The discovery cohort (426 Leonbergers and 91 Saint Bernards) was genotyped by using either the Axiom Canine Set A or HD arrays (Thermo Fisher Scientific, Waltham, MA, USA) or the Illumina CanineHD BeadChip array (Illumina, San Diego, CA, USA). Samples genotyped on lower single nucleotide polymorphism (SNP) density arrays were imputed with Beagle 4.1 [17,18], using a diverse reference dataset containing 526,045 variants in 49 wolves and 2871 dogs (including 65 Leonbergers and 23 Saint Bernards; Supplementary Table S1). The data were filtered to include only biallelic SNPs with a minor allele frequency ≥0.02, a per-SNP genotyping rate ≥95%, and a per-individual genotyping rate ≥95%. The reference dataset was phased on a per-chromosome basis, using Beagle 4.1 with default parameters of 10 iterations and an effective population size of 200. Next, a target dataset containing approximately 174,000 variants in 402 Leonbergers and 66 Saint Bernards (Supplementary Table S1) was filtered to include only biallelic SNPs with a minor allele frequency ≥0.02, and then checked for concordance with the filtered and phased reference dataset, using the Beagle 4 utility conform-gt [19]. Conforming sites of the target dataset were imputed to the reference dataset on a per-chromosome basis, using Beagle 4.1 with the following settings: window size 50 kb, overlap 3 kb, effective population size 200, and 10 iterations. The per-chromosome imputed data were concatenated

and sorted, using VCFtools 0.1.13 [20]; variants with a Beagle 4.1 dosage R-squared (DR2) ≥0.7 were retained for downstream analysis. In total, we used imputed SNP data for 468 Leonberger and Saint Bernard dogs in this study.

To evaluate haplotypes across breeds, approximately 126,000 SNPs common across genotyping platforms were extracted from non-imputed SNP genotype data for 12,931 canids, which were either generated during this study or publicly available (Supplementary Table S1) and phased with Beagle 4.1 as described above.

2.6. Genome-Wide Association Study and Fine-Mapping

The discovery cohort from the two breeds combined contained 517 dogs (144 LPPN-cases and 373 controls). Quality control filtering steps of the imputed SNP array genotyping data were carried out by using PLINK v1.9 [21]. The dataset was pruned for low minor allele frequency (0.05) and failure to meet Hardy–Weinberg equilibrium (0.0001) and consisted of 289,553 markers. An across-breed genome-wide association study (GWAS) was performed with GEMMA v0.98 [22], using a linear mixed model including an estimated kinship matrix from centered genotypes to correct for the genomic inflation. The significance threshold was estimated by Bonferroni correction. Manhattan and Q–Q plots of the corrected *p*-values were generated in R environment v3.6.0 [23], using the qqman package [24]. Haplotypes around the significantly associated locus obtained from GWAS were constructed by using Beagle 4.1 for all canids with available SNP genotype data (n = 13,399) (Supplementary Table S1). All genome positions refer to the CanFam3.1 reference assembly.

2.7. Whole-Genome Sequencing

Whole-genome sequence (WGS) data of 716 publicly available dogs of 131 different breeds, and nine wolves [25] (Supplementary Table S2) were studied in order to identify the causative variant in the disease-associated region obtained by the GWAS. This set included 34 Leonbergers and two Saint Bernards diagnosed with a form of LPPN unexplained by the previously known variants in *ARHGEF10* [8], *GJA9* [9], *RAB3GAP1* [10–12] and *RAPGEF6* [13], as well as seven Leonbergers and one Saint Bernard used as controls. The sequence data analysis and calling of single nucleotide variants and small indels (SNVs), including the prediction of functional effects, were described previously [25]. The Integrative genomics viewer (IGV) software 2.8.2 [26] was used for visual inspection and screening for structural variants in the region of interest in the affected dogs' WGS.

2.8. Targeted Genotyping

Polymerase chain reaction (PCR) and Sanger sequencing were used to validate and genotype the variant identified from WGS. PCR products from genomic DNA were amplified by using AmpliTaqGold360 MasterMix (Thermo Fisher Scientific), and the purified PCR amplicons were directly sequenced on an ABI3730 capillary sequencer (Thermo Fisher Scientific). The *CNTNAP1* missense variant (XM_548083.6:c.2810G>A) was genotyped, using the following primers: TCCCTTGCCCTCCCTATATC (forward) and AGTCCTAATGCCCTCTGCTG (reverse). The sequence data were analyzed by using Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA).

2.9. Protein Predictions

The MutPred2 [27], PROVEAN [28] and PON-P2 [29] in silico prediction tools were used to predict biological consequences of the discovered variant on the encoded protein. All references to the canine *CNTNAP1* gene correspond to the accessions NC_006591.3 (NCBI accession), XM_548083.6 (mRNA), and XP_548083.3 (protein). The Genome Aggregation Database (gnomAD) [30] was searched for the corresponding variant in the human *CNTNAP1* gene (NP_003623.1).

2.10. Availability of Data and Material

The WGS are freely available at the European Nucleotide Archive (ENA). All accession numbers of the used genomes are available in the Supplementary Table S2. The sources of SNP array genotyping data published before are detailed in the Supplementary Table S1, and the dataset generated for this study is available from the corresponding author on reasonable request. All genome positions are reported with respect to the dog reference genome assembly CanFam3.1 and NCBI annotation release 105.

3. Results

3.1. Phenotype

The herein studied affected dogs showed generic signs of LPPN (Supplementary Table S1) with the key feature across breeds being breathing difficulty, often described as noisy or raspy breathing (Supplementary Video S1). Due to LP, 247 dogs (121 Leonbergers, 10 Saint Bernards, 114 Labrador retrievers, and 2 mixed-breed dogs) underwent an arytenoid lateralization surgery, 25 of which (ten Leonbergers, seven Saint Bernards, six Labrador retrievers, and two mixed-breed dogs) tested homozygous for the studied *CNTNAP1* variant (Supplementary Table S1). Additional clinical signs, which were noted variably among the dogs, included difficulty swallowing, changes in barking frequency and quality, high-stepping and uncoordinated gait, stumbling and tripping, exercise intolerance, and limb muscle atrophy.

3.2. Neuropathological and Morphometric Findings

Peroneal nerve biopsies were evaluated from three archived normal adult Labrador retriever (8–10 years) samples (representative image in Figure 1a), and three LPPN-affected dog samples: a nine-year-old male Labrador retriever (Figure 1b), a three-year-old male Leonberger (Figure 1c), and a three-year-old male Saint Bernard (Figure 1d), all of which tested homozygous for the studied *CNTNAP1* variant. Compared to control nerve, pathological changes were similar among affected dogs of all three breeds and included a subjective decrease in the number of myelinated nerve fibers compared to control nerve (Figure 1b–d) with scattered inappropriately thin myelin sheaths for the axon diameter (best demonstrated in Figure 1b,c). The inappropriately thin myelinated fibers were not found in the nerves of control dogs. Myelin splitting and ballooning, onion-bulb formations, and axonal degeneration were not observed in any of the biopsies.

A histogram of axonal size-frequency distribution of the relative percentage of small (<5 µm) and large (>5 µm) myelinated nerve fibers is shown for the three control Labrador retrievers, the LPPN-affected Leonberger, and the LPPN-affected Labrador retriever (Supplementary Figure S1) described above. As only a partial nerve fascicle was available for the LPPN-affected Saint Bernard, those data were not included. The large and small nerve fibers are determined by the axon diameters and this does not refer to the thickness of the myelin sheath. Compared to the average values for the control Labrador retrievers, the affected Labrador retriever showed an increased population of small caliber nerve fibers and a decreased population of large caliber nerve fibers. In contrast, the affected Leonberger showed a decreased population of small fibers and an increased population of larger fibers. Calculated G-ratio, a quantitative measure of myelin thickness, was 0.586 ± 0.031 (range 0.552–0.609) for the control Labrador retrievers, 0.543 for the affected Leonberger, and 0.575 for the affected Labrador retriever.

3.3. Genome-Wide Association Study and Fine-Mapping

The across-breed GWAS using the discovery cohort of *ARHGEF10*- and *GJA9*- negative Leonbergers and Saint Bernards (144 cases vs. 373 controls) revealed a single genome-wide significantly associated region for LPPN (Figure 2a). The 15 best-associated markers were used to define a 4.6 Mb region of interest between 19.1 and 23.7 Mb on chromosome 9 (Supplementary Table S3). Fine-mapping of

this region, using the available haplotypes from non-imputed SNP array genotyping data, included 87 markers centered on the best associated SNP (chr9:20,271,681). This revealed one homozygous haplotype present most frequently in LPPN-affected dogs (n = 21) of both breeds and not present in homozygosity in any of the controls (Supplementary Table S4). Therefore, the disease-associated region was narrowed to ~0.98 Mb (bp position 19,393,936 to 20,371,611), by a combination of sharing in the 21 homozygous cases from the discovery cohort (14 Leonbergers and 7 Saint Bernards), coupled with recombination events. Based on this analysis, we hypothesized that the causative variant explaining the GWAS hit was localized on this specific haplotype occurring in both breeds from the discovery cohort. Subsequent haplotype analysis of all 13,399 canids with available SNP genotype data provided evidence for the presence of this haplotype in total of 25 dog breeds (Supplementary Table S4).

3.4. Identification of the Candidate Causative Variant

In total, 38 protein coding genes and five lncRNAs were annotated within the disease-associated ~0.98 Mb critical interval on chromosome 9 (Figure 2b). Visual inspection of this region in the WGS of five LPPN-affected dogs homozygous for the associated haplotype (four Leonbergers and one Saint Bernard) revealed no evidence for the presence of structural variants. Filtering variants for homozygous alternative genotypes shared in these five dogs within the critical interval yielded 872 intronic or intergenic, 12 synonymous, and 18 protein-changing variants (Supplementary Table S5). In addition, WGS data were available for three Leonbergers out of the 55 dogs from the discovery cohort (Supplementary Table S1) carrying a single copy of the identified disease-associated haplotype (Supplementary Table S4). Filtering for heterozygous variants in complete linkage disequilibrium with this haplotype in three dogs reduced the number of putative variants to 93 intronic or intergenic, one synonymous, and one protein-changing variant (Supplementary Table S5).

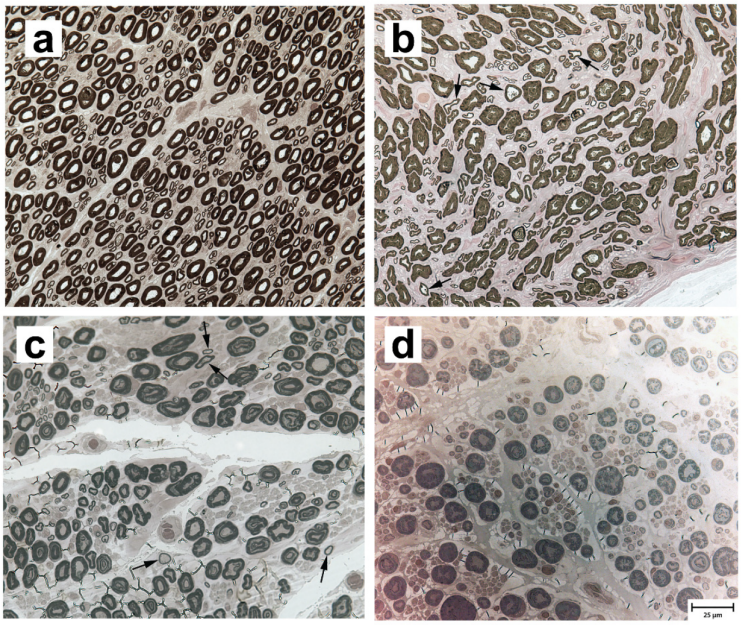

Figure 1. Paraphenylenediamine-stained resin sections from the peroneal nerves of four dogs. (**a**) An adult normal control Labrador retriever, (**b**) a nine-year-old LPPN-affected Labrador retriever, (**c**) a three-year-old LPPN-affected Leonberger, and (**d**) a three-year-old LPPN-affected Saint Bernard. All three LPPN-affected dogs were homozygous for the *CNTNAP1* variant. Arrows in (**b**) and (**c**) point to nerve fibers that are inappropriately thin for the axon diameters. Bar in lower right image indicates 25 µm and is valid for all images.

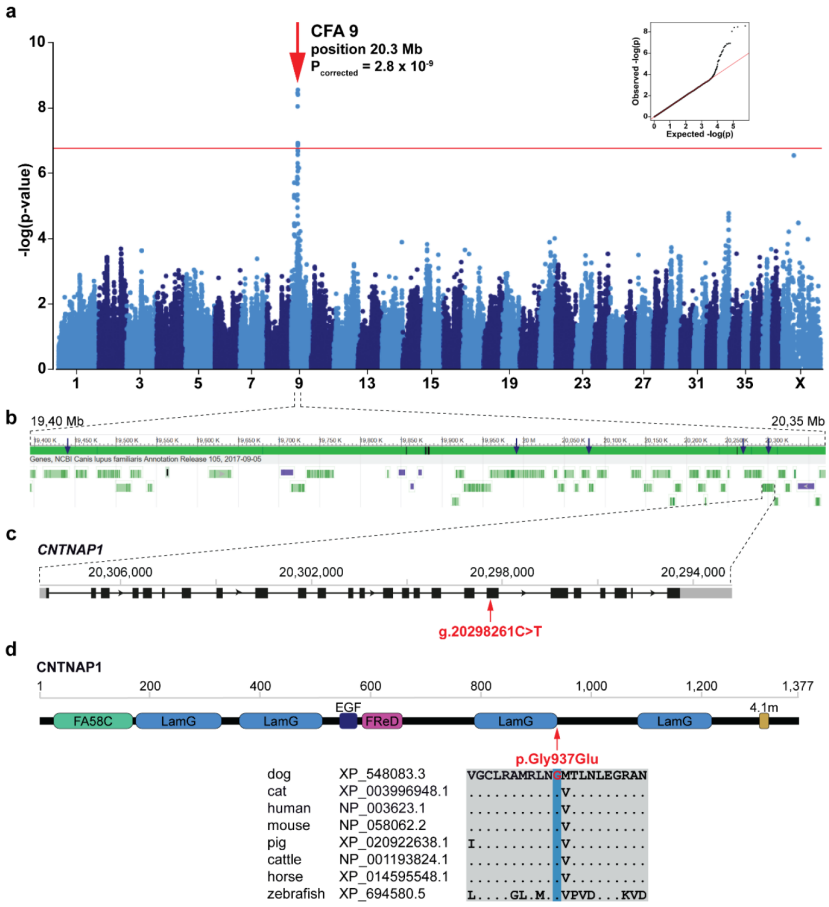

Figure 2. Identification of a new LPPN-associated locus and variant in Leonbergers and Saint Bernards. (**a**) Manhattan plot for the two-breed genome-wide association study (GWAS) using 144 LPPN-affected dogs and 373 normal control dogs indicates a signal with multiple associated single nucleotide polymorphisms (SNPs) on chromosome 9. The -log *p*-values for each SNP are plotted on the *y*-axis versus each canine chromosome on the *x*-axis. The red line represents the Bonferroni corrected significance threshold (−log(*p*-value) = 6.76). Inset: Corrected Q–Q plot confirms that the observed *p*-values of the best-associated markers have stronger association with the trait than expected by chance (null hypothesis, red line). (**b**) Gene content in the ~0.98 Mb region of interest. Blue arrows show the best-associated markers from GWAS. Green bars represent the different genes and violet bars represent lncRNAs. (**c**) Schematic representation of the *CNTNAP1* gene showing the variant (XM_548083.6:c.2810G>A) location in exon 18. (**d**) Schematic representation of the CNTNAP1 protein with its domains: coagulation factor 5/8 C-terminal domain (FA58C), laminin G domains (LamG), calcium-binding EGF-like domain (EGF), fibrinogen-related domain (FReD), and putative band 4.1 homologues' binding motif (4.1m). The CNTNAP1 amino acid substitution (XP_548083.3:p.Gly937Glu) position is shown together with the amino acid multiple alignment indicating the residue is highly conserved across species.

CNTNAP1 represents a functional candidate gene due to its involvement in human congenital hypomyelination, where vocal cord palsy is a common clinical finding [31], so we pursued the lone remaining missense variant in the *CNTNAP1* gene (CFA9:g.20298261C>T; c.2810G>A; p.Gly937Glu) further. This *CNTNAP1* variant was predicted to be deleterious by several prediction tools (MutPred2

score: 0.884, PROVEAN score: −7.667, PON-P2 probability for pathogenicity: 0.848). It is located in exon 18 of the *CNTNAP1* gene (Figure 2c) and affects a highly conserved amino acid residue at the end of the third laminin G domain of the CNTNAP1 protein (Figure 2d). Two missense variants (rs905697967:p.Gly938Arg and rs763033339:p.Gly938Glu) in the human *CNTNAP1* coding region at the corresponding position were found in the gnomAD [30]. Both variants were reported with allele frequency 7.95×10^{-4} and no homozygous individuals were detected [30]. The canine missense variant in *CNTNAP1* was present in 5 of the 688 control canid WGS with a frequency of 0.004 (Supplementary Table S5), including one homozygous (English bulldog) and four heterozygous dogs (golden retriever, Labrador retriever, English bulldog, and Kerry blue terrier).

3.5. The CNTNAP1 Variant Occurs in Several Breeds

Available SNP array genotype data of 13,337 dogs and 62 wild canids (Supplementary Table S1) were inspected for the identified *CNTNAP1*-associated haplotype (Supplementary Table S4). Out of this group, targeted genotyping by PCR and Sanger sequencing was performed in 2469 canids and demonstrated perfect concordance between the *CNTNAP1*-associated haplotype and the *CNTNAP1*:c.2810G>A genotype (Supplementary Table S4). In addition, 2362 dogs without SNP array data were directly genotyped for the variant (Supplementary Table S1); this included 557 dogs from the validation cohort and 1805 from the population cohort.

In total, the variant was found in 25 different breeds and no wild canids. Homozygotes for the missense allele in breeds other than the Leonberger, Saint Bernard, and Labrador retriever were identified in 46 English bulldogs, six Irish terriers, two boxers, one bullmastiff, one Peruvian hairless dog, one Yorkshire terrier, and one golden retriever (Supplementary Table S1), all with unknown precise health history. Additionally, two LPPN-affected mixed-breed dogs enrolled in the Michigan State University's GOLPP study [6] were also homozygous for the missense allele (Supplementary Table S1).

Analysis of the validation cohort that included the three breeds with available health information (Leonberger, Saint Bernard, and Labrador retriever) demonstrated that the *CNTNAP1* variant is not present in a homozygous state in any dog apparently non-affected with LPPN (Table 2). The 18 homozygous LPPN-affected Leonbergers represent 4.1% of all as yet unexplained cases with any age of onset that were not carrying the previously identified disease-causing variants in *ARHGEF10* and *GJA9*. For the Saint Bernard, 10 out of 24 (41.6%) diagnosed dogs were homozygous mutant, whereas only 4.7% of GOLPP-affected Labrador retrievers carried two copies of the *CNTNAP1* variant. The homozygous A/A *CNTNAP1* genotype also occurred rarely in single dogs of the population controls from each of the three studied breeds (Table 2). Altogether, the mutant allele frequency was estimated as 6.6% in the studied Leonbergers (n = 2738 dogs), 13.9% in Saint Bernards (n = 305 dogs), and 5.2% in Labrador retrievers (n = 1524 dogs). Among the 22 other breeds segregating for this variant, the allele frequency was highest in the English bulldogs (n = 193 dogs) and Irish terriers (n = 184 dogs), estimated at 46.6% and 17.1%, respectively (Supplementary Table S1).

Mean age of onset in the LPPN-affected dogs was investigated for a subset of 770 dogs with detailed health information from three breeds (596 Leonbergers, 28 Saint Bernards, and 146 Labrador retrievers) and showed a marked difference between the cases depending on the different underlying genetic causes (Figure 3). The average age of onset of clinical signs in the limited number of dogs homozygous for the herein described *CNTNAP1* variant was 3.4, 2.1, and 7.5 years in Leonbergers, Saint Bernards, and Labrador retrievers, respectively. In comparison, the age of onset of clinical signs in the previously characterized *ARHGEF10*-associated polyneuropathy [8] was seen in Leonbergers and Saint Bernards with average ages of 2.2 and 1.6 years, respectively. Additionally, affected Leonbergers with the *GJA9* frameshift variant [9] had average age of onset of their clinical signs of 6.2 years. Interestingly, the LPPN-affected Labrador retrievers that do not carry the herein identified *CNTNAP1* variant and come from the GOLPP study [6] showed a higher average age of onset of 11.5 years (Figure 3). The difference in age of disease onset between the dogs with the identified *CNTNAP1* variant and the cases without known disease-causing mutation was statistically significant in all

three breeds (Leonberger p-value = 0.000001002, Saint Bernard p-value = 0.01681, Labrador retriever p-value = 0.002662). The difference between dogs with the *CNTNAP1* variant and the *ARHGEF10* variant was significant in Leonbergers (p-value = 0.002538) but not significant in Saint Bernards (p-value = 0.3095). The difference between Leonbergers with the *CNTNAP1* variant and the *GJA9* variant was statistically significant (p-value = 0.00000001797).

Table 2. Segregation of the *CNTNAP1*:c.2810G>A genotypes with laryngeal paralysis and polyneuropathy (LPPN) in three breeds with available health information.

Breed	LPPN Status	*CNTNAP1* Genotypes		
		G/G	G/A	A/A [1]
Leonberger (n = 2738)	Affected (n = 434) [2]	358	58	18
	LPN1/LPN2 (n = 192) [3]	180	11	1
	Non-affected (n = 659) [4]	605	54	0
	Population controls (n = 1453)	1258	192	3
Saint Bernard (n = 305)	Affected (n = 24) [2]	9	5	10
	LPN1 (n = 5) [3]	2	3	0
	Non-affected (n = 14)	9	5	0
	Population controls (n = 262)	213	46	3
Labrador retriever (n = 1524)	Affected (n = 148)	132	9	7
	Non-affected (n = 45)	42	3	0
	Population controls (n = 1331)	1200	128	3

[1] Includes the three herein described histopathologically confirmed cases. [2] LPPN-affected dogs that tested negative for *ARHGEF10* [8] and *GJA9* [9] mutations. [3] LPPN-affected dogs homozygous for the previously described polyneuropathy-associated variant in *ARHGEF10* (LPN1) [8], and homozygous or heterozygous for the variant in *GJA9* (LPN2) [9]. [4] Includes cases homozygous for the previously described leukoencephalomyelopathy-associated variant in *NAPEPLD* [14].

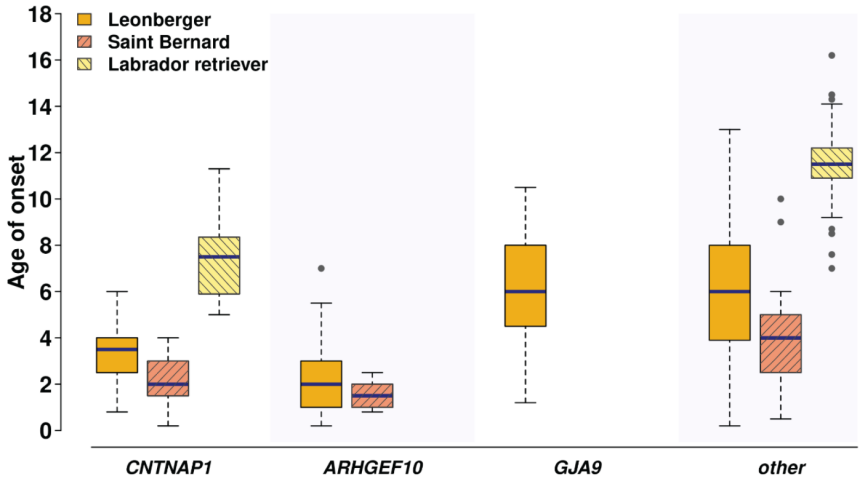

Figure 3. Age of onset of clinical signs for laryngeal paralysis and polyneuropathy (LPPN) differs depending on underlying genetic variants and across the three breeds. Comparison of the age of onset of clinical signs in the LPPN-affected dogs (n = 770), which were genotyped homozygous for the polyneuropathy-associated variants in *CNTNAP1*:c.2810G>A, or *ARHGEF10*:c.1955_1958+6delCACGGTGAGC [8], and homozygous or heterozygous for the variant in *GJA9*:c.1107_1108delAG [9], as well as the age of onset in the yet unexplained cases (other), is shown for Leonbergers (gold bars; n = 596), Saint Bernards (pink bars; n = 28), and Labrador retrievers (yellow bars; n = 146). Note that the *ARHGEF10* variant is only present in Leonbergers and Saint Bernards, and the *GJA9* variant only in Leonbergers.

4. Discussion

This study has revealed strong evidence for a new potentially pathogenic variant associated with laryngeal paralysis and polyneuropathy (LPPN) initially observed in two closely related giant dog breeds, the Leonberger and Saint Bernard. Interestingly, the variant also explains some cases of GOLPP in Labrador retrievers and segregates at different frequencies in 22 other unrelated dog breeds, including English bulldogs and Irish terriers, suggesting that the derived allele predates modern breed formation. Apparently, the variant was not purged by either selection or drift.

The affected *contactin-associated protein 1* (*CNTNAP1*) gene has been previously implicated in human autosomal recessive neurological diseases with a broad spectrum of clinical phenotypes and neonatal and childhood onsets: congenital hypomyelinating neuropathy type 3 (OMIM 618186), lethal congenital contracture syndrome 7 (OMIM 616286), and childhood-onset Charcot–Marie–Tooth disease [32]. *CNTNAP1* is essential in the formation of paranodal axoglial junctions in myelinated axons and is also involved in regulating neural progenitor cells and the development of the cerebral cortex [33]. Pathological variants in the *CNTNAP1* gene may lead to defective or absent proteins critical to development of central or peripheral nervous systems. Even though, based on the current gnomAD [30] database, the corresponding glycine to glutamic acid exchange occurs very rarely in humans, so far, there is no evidence reported for disease association. Human patients with other *CNTNAP1* homozygous frameshift or nonsense variants show a more severe disorder with early-onset neurological disease, including severe respiratory compromise and early lethality, while those carrying missense variants can survive beyond infancy [34]. This suggests that the missense alleles affecting the myelination and development of paranodal junctions may be hypomorphic and have some residual function. Although the precise role of the protein domains in ligand binding is not fully understood, several missense variants were predicted to impact the domain structure and protein folding [34].

Pathological changes in semi-thin transverse resin sections of the peroneal nerves of three affected dogs had similarities to those in published human cases, including reduced myelinated nerves and inappropriately thin myelin sheaths for the axon diameters [35]; however, in the affected dogs, thinly myelinated fibers were scattered and fewer in number. This may reflect the severity of disease in neonatal onset human cases and milder disease with an adult onset in dogs. The increased population of small fibers in the affected Labrador retriever, as compared to the affected Leonberger, may reflect attempts at regeneration in the Labrador retriever or genetic differences in other modifying genes. There are several limitations to the pathological studies in these dogs: The number of affected dogs with available peripheral nerve biopsies was small; detailed study of the paranodal areas of the peroneal nerves was limited by the retrospective nature of the study, the use of archived nerve specimens obtained many years prior to the identification of the variant, and the necessity of preparation of the nerves at the time of original processing for teased fibers and for longitudinal evaluation; and the standard processing for diagnostic specimens in the laboratory of one of the authors (GDS) is in transverse section. Future in-depth prospective studies of the peripheral nerves, including laryngeal nerves in more cases of confirmed LPPN with the *CNTNAP1* gene variant in each breed, are necessary to fully evaluate the observed pathological changes.

The neurological diseases identified in humans associated with variants in *CNTNAP1* support our recent speculation, based on the enrichment of this allele in Leonbergers [7], that the herein-described missense variant predicted in silico to be deleterious represents a promising candidate causative mutation for inherited neurological disorders in dogs. The striking genetic association data implicate that this mutation affects the function of the encoded protein, although we have not studied this further. Homozygosity for the missense variant in the *CNTNAP1* gene is significantly associated with the development of LPPN in large and giant-sized dogs, indicating recessive inheritance in all three studied breeds (Leonberger, Saint Bernard, and Labrador retriever). As yet we do not have convincing evidence for causality in smaller dog breeds segregating for this variant, such as the English bulldog or Irish terrier, and further study with reliably phenotyped populations is needed. However, we hypothesize that the apparently higher allele frequency in English bulldogs may be a result of

underdiagnosed LP due to breathing difficulties related to brachycephalic airway syndrome, including laryngeal collapse [36] obscuring a neurodegenerative LP. The observed higher allele frequency in Irish terriers, although without available health information, suggests that the association between the variant and LPPN phenotype is breed-specific and may not be pathogenic in some breeds although this needs to be evaluated. We also hypothesize that the observed later age of onset in the Labrador retriever group, compared to the Leonbergers and Saint Bernards, might be either due to the different genetic breed background and/or their smaller stature and correspondingly shorter laryngeal nerve length; the latter was previously suggested by correlation between growth (specifically height) and laryngeal neuropathy in horses [37].

5. Conclusions

In conclusion, we identified a potentially causative genetic variant in *CNTNAP1* associated with autosomal recessive younger-onset LPPN in large and giant dogs, specifically Leonbergers, Saint Bernards, and Labrador retrievers. Our results represent the first large animal model for a *CNTNAP1*-related neurodegenerative disease. The developed genetic test enables veterinary diagnostics and selective breeding against this deleterious variant across breeds to reduce the occurrence of LPPN. Therefore, selecting based on this additional disease-associated variant, which we have designated LPPN3, will enable dog breeders to make even greater strides in controlling the propagation of this devastating disorder and maintaining the health of Leonberger, Saint Bernard, and Labrador retriever populations. However, the fact that not all LPPN cases from the three intensively studied breeds carried the described variant, together with the broad range in age of onset of the clinical signs for the as yet unexplained cases, indicates that still unknown genetic heterogeneity of different forms of canine LPPN need to be studied in future.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/12/1426/s1. Figure S1: Histogram showing the axonal size-frequency distribution of myelinated nerve fibers in peroneal nerve sections from control Labrador retrievers (average of three archived adult control dog samples), an LPPN-affected Labrador retriever, and an LPPN-affected Leonberger. Both affected dogs were homozygous for the *CNTNAP1* variant. Axons are distributed into bins determined by axonal diameter based on perimeter. Table S1: Sample designations and detailed information of all dogs used in the study. Table S2: Sample designations and affection status of all whole-genome sequences used for filtering variants. Table S3: The 10,000 most significant markers sorted by *p*-value obtained from the across-breed genome-wide association study. Table S4: Haplotype diversity for the LPPN-associated genome region on chromosome 9. Table S5: Variants in the region of interest obtained from GWAS on chromosome 9. Video S1: Video illustrating the clinical phenotype of two LPPN-affected dogs: a Leonberger and a Saint Bernard.

Author Contributions: Conceptualization, J.R.M. and C.D.; data curation, B.J.S. and V.J.; formal analysis, A.L. and K.M.M.; funding acquisition, J.R.M. and C.D.; investigation, A.L., K.M.M., G.D.S., and J.P.S.; methodology, A.L. and K.M.M.; resources, P.J.J.M., P.A.J.L., P.A.W., S.M.P.-J., K.J.E., G.S.J., and L.H.; software, S.G.F. and V.J.; supervision, J.R.M. and C.D.; visualization, A.L., K.M.M., and G.D.S.; writing—original draft, A.L., K.M.M., G.D.S., J.R.M., and C.D.; writing—review and editing, A.L., K.M.M., S.G.F., G.D.S., K.J.E., J.R.M., and C.D. All authors have read and agreed to the published version of the manuscript.

Funding: The University of Minnesota received a gift from the Leonberger Health Foundation to support research in this breed. Kari J. Ekenstedt (K.J.E.) was supported by the Office of the Director, National Institutes of Health (NIH), under award number K01-OD027051. Other authors have no external funding to report.

Acknowledgments: The authors are grateful to the breeders and owners of all dogs who provided samples and shared valuable information about their dogs. We thank Nathalie Besuchet Schmutz for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the WGS and providing high-performance computational infrastructure.

Conflicts of Interest: Both the University of Minnesota and the University of Bern are offering genotyping tests for polyneuropathy- and leukoencephalomyelopathy-associated variants in their respective laboratories, and proceeds from these tests go toward ongoing canine genetic research. S.G.F is guest editor of this special issue of *Genes*, but has not in any way been involved in or interacted with the journal's review process or editorial decision-making. The authors declare that they have no other competing interest.

References

1. Kitshoff, A.M.; Van Goethem, B.; Stegen, L.; Vandekerckhov, P.; De Rooster, H. Laryngeal paralysis in dogs: An update on recent knowledge. *J. S. Afr. Vet. Assoc.* **2013**, *84*, 1–9. [CrossRef]
2. Monnet, E. Surgical Treatment of Laryngeal Paralysis. *Vet. Clin. N. Am. Small Anim. Pract.* **2016**, *46*, 709–717. [CrossRef]
3. Mackin, G.A. Diagnosis of patients with peripheral nerve disease. *Clin. Podiatr. Med. Surg.* **1994**, *11*, 545–569. [PubMed]
4. Gabriel, A.; Poncelet, L.; Van Ham, L.; Clercx, C.; Braund, K.G.; Bhatti, S.; Detilleux, J.; Peeters, D. Laryngeal paralysis-polyneuropathy complex in young related Pyrenean mountain dogs. *J. Small Anim. Pract.* **2006**, *47*, 144–149. [CrossRef] [PubMed]
5. Granger, N. Canine inherited motor and sensory neuropathies: An updated classification in 22 breeds and comparison to Charcot-Marie-Tooth disease. *Vet. J.* **2011**, *188*, 274–285. [CrossRef] [PubMed]
6. GOLPP|College of Veterinary Medicine at MSU. Available online: https://cvm.msu.edu/scs/research-initiatives/golpp (accessed on 7 September 2020).
7. Letko, A.; Minor, K.M.; Jagannathan, V.; Seefried, F.R.; Mickelson, J.R.; Oliehoek, P.; Drögemüller, C. Genomic diversity and population structure of the Leonberger dog breed. *Genet. Sel. Evol.* **2020**, *52*, 61. [CrossRef]
8. Ekenstedt, K.J.; Becker, D.; Minor, K.M.; Shelton, G.D.; Patterson, E.E.; Bley, T.; Oevermann, A.; Bilzer, T.; Leeb, T.; Drögemüller, C.; 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. [CrossRef]
9. Becker, D.; Minor, K.M.; Letko, A.; Ekenstedt, K.J.; Jagannathan, V.; Leeb, T.; Shelton, G.D.; Mickelson, J.R.; Drögemüller, C. A GJA9 frameshift variant is associated with polyneuropathy in Leonberger dogs. *BMC Genom.* **2017**, *18*, 662. [CrossRef]
10. Wiedmer, M.; Oevermann, A.; Borer-Germann, S.E.; Gorgas, D.; Shelton, G.D.; Drögemüller, M.; Jagannathan, V.; Henke, D.; Leeb, T. A RAB3GAP1 SINE Insertion in Alaskan Huskies with Polyneuropathy, Ocular Abnormalities, and Neuronal Vacuolation (POANV) Resembling Human Warburg Micro Syndrome 1 (WARBM1). *G3 Genes Genomes Genet.* **2016**, *6*, 255–262. [CrossRef]
11. Mhlanga-Mutangadura, T.; Johnson, G.S.; Schnabel, R.D.; Taylor, J.F.; Johnson, G.C.; Katz, M.L.; Shelton, G.D.; Lever, T.E.; Giuliano, E.; Granger, N.; et al. A mutation in the Warburg syndrome gene, RAB3GAP1, causes a similar syndrome with polyneuropathy and neuronal vacuolation in Black Russian Terrier dogs. *Neurobiol. Dis.* **2016**, *86*, 75–85. [CrossRef]
12. Mhlanga-Mutangadura, T.; Johnson, G.S.; Ashwini, A.; Shelton, G.D.; Wennogle, S.A.; Johnson, G.C.; Kuroki, K.; O'Brien, D.P. A Homozygous RAB3GAP1:c.743delC Mutation in Rottweilers with Neuronal Vacuolation and Spinocerebellar Degeneration. *J. Vet. Intern. Med.* **2016**, *30*, 813–818. [CrossRef] [PubMed]
13. Rasouliha, S.H.; Barrientos, L.; Anderegg, L.; Klesty, C.; Lorenz, J.; Chevallier, L.; Jagannathan, V.; Rösch, S.; Leeb, T. A RAPGEF6 variant constitutes a major risk factor for laryngeal paralysis in dogs. *PLoS Genet.* **2019**, *15*, 1–17. [CrossRef] [PubMed]
14. Minor, K.M.; Letko, A.; Becker, D.; Drögemüller, M.; Mandigers, P.J.J.; Bellekom, S.R.; Leegwater, P.A.J.; Stassen, Q.E.M.; Putschbach, K.; Fischer, A.; et al. Canine NAPEPLD-associated models of human myelin disorders. *Sci. Rep.* **2018**, *8*, 5818. [CrossRef] [PubMed]
15. Parker, H.G.; Dreger, D.L.; Rimbault, M.; Davis, B.W.; Mullen, A.B.; Carpintero-Ramirez, G.; Ostrander, E.A. Genomic Analyses Reveal the Influence of Geographic Origin, Migration, and Hybridization on Modern Dog Breed Development. *Cell Rep.* **2017**, *19*, 697–708. [CrossRef]
16. Huang, M.; Hayward, J.J.; Corey, E.; Garrison, S.J.; Wagner, G.R.; Krotscheck, U.; Hayashi, K.; Schweitzer, P.A.; Lust, G.; Boyko, A.R.; et al. A novel iterative mixed model to remap three complex orthopedic traits in dogs. *PLoS ONE* **2017**, *12*, e0176932. [CrossRef]
17. Browning, S.R.; Browning, B.L. Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies by Use of Localized Haplotype Clustering. *Am. J. Hum. Genet.* **2007**, *81*, 1084–1097. [CrossRef]
18. Browning, B.L.; Browning, S.R. Genotype Imputation with Millions of Reference Samples. *Am. J. Hum. Genet.* **2016**, *98*, 116–126. [CrossRef]
19. Browning, B. Conform-gt Program. Available online: https://faculty.washington.edu/browning/conform-gt.html (accessed on 22 September 2020).

20. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
21. Chang, C.C.; Chow, C.C.; Tellier, L.C.A.M.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [CrossRef]
22. Zhou, X.; Stephens, M. Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **2012**, *44*, 821–824. [CrossRef]
23. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
24. Turner, S.D.; Turner, D.S. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *bioRxiv* **2014**, *81*, 559–575. [CrossRef]
25. Jagannathan, V.; Drögemüller, C.; Leeb, T. Dog Biomedical Variant Database Consortium, (DBVDC) A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. *Anim. Genet.* **2019**, *50*, 695–704. [CrossRef] [PubMed]
26. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [CrossRef] [PubMed]
27. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.A.; Lin, G.N.; Nam, H.-J.; Mort, M.; Cooper, D.N.; Sebat, J.; Iakoucheva, L.M.; et al. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv* **2017**, 134981. [CrossRef]
28. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef]
29. Niroula, A.; Urolagin, S.; Vihinen, M. PON-P2: Prediction method for fast and reliable identification of harmful variants. *PLoS ONE* **2015**, *10*, e117380. [CrossRef]
30. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]
31. Low, K.; Stals, K.; Caswell, R.; Clayton-Smith, J.; Donaldson, A.; Foulds, N.; Splitt, M.; Norman, A.; Urankar, K.; Vijayakumar, K.; et al. CNTNAP1: Extending the phenotype of congenital hypomyelinating neuropathy in 6 further patients. *Neuromuscul. Disord.* **2017**, *27*, S148. [CrossRef]
32. Freed, A.S.; Weiss, M.D.; Malouf, E.A.; Hisama, F.M. CNTNAP1 mutations in an adult with Charcot Marie Tooth disease. *Muscle Nerve* **2019**, *60*, E28–E30. [CrossRef]
33. Sabbagh, S.; Antoun, S.; Mégarbané, A. CNTNAP1 Mutations and Their Clinical Presentations: New Case Report and Systematic Review. *Case Rep. Med.* **2020**, *2020*. [CrossRef]
34. Low, K.; Stals, K.; Caswell, R.; Wakeling, M.; Clayton-Smith, J.; Donaldson, A.; Foulds, N.; Norman, A.; Splitt, M.; Urankar, K.; et al. Phenotype of CNTNAP1: A study of patients demonstrating a specific severe congenital hypomyelinating neuropathy with survival beyond infancy. *Eur. J. Hum. Genet.* **2018**, *26*, 796–807. [CrossRef] [PubMed]
35. Vallat, J.-M.; Nizon, M.; Magee, A.; Isidor, B.; Magy, L.; Péréon, Y.; Richard, L.; Ouvrier, R.; Cogné, B.; Devaux, J.; et al. Contactin-Associated Protein 1 (CNTNAP1) Mutations Induce Characteristic Lesions of the Paranodal Region. *J. Neuropathol. Exp. Neurol.* **2016**, *75*, 1155–1159. [CrossRef] [PubMed]
36. Meola, S.D. Brachycephalic Airway Syndrome. *Top. Companion Anim. Med.* **2013**. [CrossRef] [PubMed]
37. Boyko, A.R.; Brooks, S.A.; Behan-Braman, A.; Castelhano, M.; Corey, E.; Oliveira, K.C.; Swinburne, J.E.; Todhunter, R.J.; Zhang, Z.; Ainsworth, D.M.; et al. Genomic analysis establishes correlation between growth and laryngeal neuropathy in Thoroughbreds. *BMC Genom.* **2014**, *15*, 259. [CrossRef] [PubMed]

Canine *NAPEPLD*-associated models of human myelin disorders

Journal:	Scientific Reports
Manuscript status:	published
Contributions:	Investigation, Formal analysis, Visualization, Writing – original draft, Writing – review & editing

SCIENTIFIC REPORTS

OPEN

Canine *NAPEPLD*-associated models of human myelin disorders

K. M. Minor[1], A. Letko[2], D. Becker[2], M. Drögemüller[2], P. J. J. Mandigers[3], S. R. Bellekom[3], P. A. J. Leegwater[3], Q. E. M. Stassen[3], K. Putschbach[4], A. Fischer[4], T. Flegel[5], K. Matiasek[4], K. J. Ekenstedt[6], E. Furrow[1], E. E. Patterson[1], S. R. Platt[7], P. A. Kelly[8], J. P. Cassidy[8], G. D. Shelton[9], K. Lucot[10], D. L. Bannasch[10], H. Martineau[11], C. F. Muir[11], S. L. Priestnall[11], D. Henke[12], A. Oevermann[13], V. Jagannathan[2], J. R. Mickelson[1] & C. Drögemüller[2]

Received: 6 November 2017
Accepted: 20 March 2018
Published online: 11 April 2018

Canine leukoencephalomyelopathy (LEMP) is a juvenile-onset neurodegenerative disorder of the CNS white matter currently described in Rottweiler and Leonberger dogs. Genome-wide association study (GWAS) allowed us to map LEMP in a Leonberger cohort to dog chromosome 18. Subsequent whole genome re-sequencing of a Leonberger case enabled the identification of a single private homozygous non-synonymous missense variant located in the highly conserved metallo-beta-lactamase domain of the *N-acyl phosphatidylethanolamine phospholipase D* (*NAPEPLD*) gene, encoding an enzyme of the endocannabinoid system. We then sequenced this gene in LEMP-affected Rottweilers and identified a different frameshift variant, which is predicted to replace the C-terminal metallo-beta-lactamase domain of the wild type protein. Haplotype analysis of SNP array genotypes revealed that the frameshift variant was present in diverse haplotypes in Rottweilers, and also in Great Danes, indicating an old origin of this second *NAPEPLD* variant. The identification of different *NAPEPLD* variants in dog breeds affected by leukoencephalopathies with heterogeneous pathological features, implicates the NAPEPLD enzyme as important in myelin homeostasis, and suggests a novel candidate gene for myelination disorders in people.

The classification of human leukoencephalopathies was initially based upon pathology and biochemistry and has been applied to disorders caused by toxic, acquired vascular, or infectious insults, as well as inherited disorders[1]. This scheme has recently been updated to a case definition of leukodystrophies that refer to 30 distinct disorders with wasting (dystrophy) of the brain's white matter (leuko) and a consensus definition of heritable white matter disorders based on neuroimaging[1,2]. Interestingly, nearly half of all patients whose neuroimaging studies indicate white matter disease and whose clinical manifestations suggest a genetic etiology do not receive a specific diagnosis[3]. More than 60 distinct types of genetic leukoencephalopathies (gLE), a recently introduced broader term[1], are associated with white matter lesions in the central nervous system (CNS), and in people these represent a heterogeneous group of disorders with both highly variable clinical and pathologic manifestations[1,4]. A recent genetic screening of 118 leukoencephalopathy-related genes in 49 patients diagnosed with gLE showed evidence for pathogenic variants in 40.8% of them[5].

In humans, primary myelin disorders of CNS (so called white matter diseases) are caused by defects in myelin formation and/or maintenance and include dysmyelinating (abnormally formed myelin) diseases,

[1]Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, MN, 55108, USA. [2]Institute of Genetics, University of Bern, Bern, 3001, Switzerland. [3]Department of Clinical Sciences of Companion Animals, Utrecht University, Utrecht, 3508, CM, The Netherlands. [4]Centre for Clinical Veterinary Medicine, Ludwig-Maximilians-University, Munich, 80539, Germany. [5]Department of Small Animal Medicine, University of Leipzig, Leipzig, 04103, Germany. [6]Department of Basic Medical Sciences, Purdue University, West Lafayette, IN, 47907, USA. [7]Small Animal Medicine and Surgery, University of Georgia, Athens, GA, 30602, USA. [8]Veterinary Sciences Centre, University College Dublin, Dublin, D04 V1W8, Ireland. [9]Department of Pathology, University of California, La Jolla, CA, 92093, USA. [10]Department of Population Health and Reproduction, University of California-Davis, Davis, CA, 95616, USA. [11]Pathobiology and Population Sciences, The Royal Veterinary College, North Mymms, AL9 7TA, UK. [12]Division of Clinical Neurology, University of Bern, Bern, 3001, Switzerland. [13]Division of Neurological Sciences, University of Bern, Bern, 3001, Switzerland. K. M. Minor, A. Letko, J. R. Mickelson and C. Drögemüller contributed equally to this work. Correspondence and requests for materials should be addressed to C.D. (email: cord.droegemueller@vetsuisse.unibe.ch)

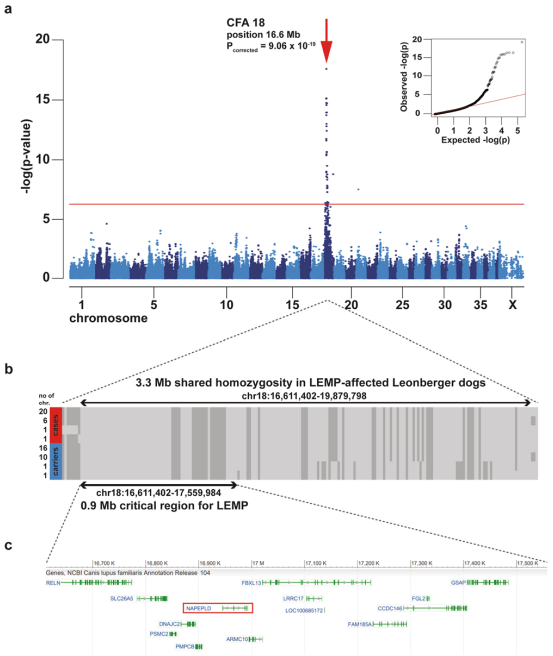

Figure 1. Positional cloning of the LEMP-associated locus in Leonbergers. (**a**) Manhattan plot for the GWAS using 14 LEMP-affected dogs and 186 control dogs is shown and indicates a signal with multiple associated SNPs on chromosome 18. The -log P-values for each SNP are plotted on the y-axis versus each canine autosome and the X chromosome on the x-axis. The red line represents the Bonferroni corrected significance threshold (−log (P) = 6.35). A mixed model analysis corrected for population stratification was carried out as described in the Methods. Inset: Corrected QQ-plot confirms that the actually observed P-values of the best associated markers have stronger association with the trait than expected by chance (null hypothesis, red line). (**b**) Haplotype analysis of SNP array genotypes of 14 cases and 28 carriers allowed fine mapping of the critical region for LEMP to a 0.9 Mb interval. Each line represents a unique haplotype. (**c**) The LEMP-associated region contains 14 loci including *NAPEPLD* gene.

hypomyelinating disorders (decreased myelin production), and spongy vacuolar degeneration of myelin[4]. Myelin disorders have also been reported in miscellaneous domestic animal species including various breeds of dog[6,7]. Although infrequently seen, during the last 40 years several breed-specific forms of myelopathy in which there is lysis of the white matter have been described and termed leukoencephalomyelopathies (OMIA 001788-9615) in Afghan Hounds[8], Rottweilers[9–12] and Leonbergers[13], or as necrotizing myelopathy in Kooiker dogs[14]. Affected dogs present clinically weak and ataxic with loss of conscious proprioception (Supplementary Video S1). Usually these diseases occur in young animals suggesting a hereditary basis.

As similar myelin disorders are known in people, this study aimed to identify the genetic cause of canine leukoencephalomyelopathy (LEMP) in Leonbergers and Rottweilers as complementary models. Herein we report the identification of a causative gene for both these forms of canine LEMP that represents a novel candidate gene for human myelin disorders such as gLE disease.

Results

Leukoencephalomyelopathy (LEMP) in Leonbergers is associated with the region of *N-acyl phosphatidylethanolamine phospholipase D* (*NAPEPLD*) on chromosome 18. For the GWAS we utilized 14 Leonberger neurological cases clinically compatible with LEMP (ages of onset 1.3–4 years), in which seven cases were confirmed by necropsy, and one dog also confirmed via magnetic resonance imaging (Supplementary Table S1). Additionally, we included 186 neurologically healthy Leonberger controls (eight years and older) based on prior 170k SNP array genotyping data (Fig. 1a). A genomic inflation factor (lambda) of 2.29 indicated the presence of population stratification and possible cryptic relatedness. We performed a multidimensional scaling analysis revealing no indication for clustering of cases outside the controls (Supplementary Fig. S1). We therefore performed an association analysis using the mixed model function that resulted in lambda dropping to 1.187. We obtained a highly significant association signal on chromosome 18 (Fig. 1a; $p_{corrected} = 9.06 \times 10^{-19}$). Haplotype analysis of the LEMP-affected Leonbergers showed a 3.3 Mb area of extended shared homozygosity from positions 16.6 to 19.9 Mb (Fig. 1b).

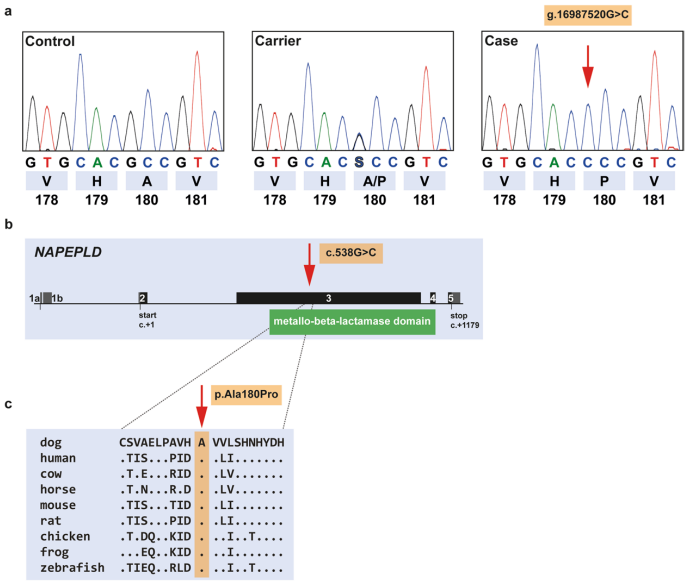

Figure 2. The *NAPEPLD* missense variant detected in LEMP-affected Leonbergers. (**a**) Chromatograms of wild type, carrier, and an affected dog indicate the c.538 G > C variant which changes codon 180 (shown below). (**b**) The variant is located in exon 3 of canine *NAPEPLD* that encodes a functionally important domain of the NAPEPLD protein. (**c**) The predicted p.Ala180Pro exchange affects an evolutionary conserved residue. The multiple sequence amino acid alignment was done using accessions XP_005631036.1 (*Canis lupus familiaris*), NP_001116310.1 (*Homo sapiens*), NP_001015680.1 (*Bos taurus*), XP_014594420.1 (*Equus caballus*), NP_848843.1 (*Mus musculus*), NP_955413.1 (*Rattus norwegicus*), NP_001025901.1 (*Gallus gallus*), XP_002933136.1 (*Xenopus tropicalis*) and NP_001074082.2 (*Danio rerio*).

Breed	Status	Total	G/G (homozygous normal; N/N)	C/G (heterozygous; D/N)	C/C (homozygous mutant; D/D)
Leonberger		7,086	5,956	1,054*	76[1]
	Affected[2]	27			27[1]
	Non-affected#,2	574	486	82	6
	Unknown§	6,485	5,470	972	43
Great Danes§		262	262		
St. Bernard§		47	47		
Newfoundland§		10	10		
Entlebucher Mountain dog§		10	10		
Appenzeller Mountain dog§		8	8		
Bernese Mountain dog§		7	7		
Greater Swiss Mountain dog§		6	6		

Table 1. *NAPEPLD* c.538 G > C genotype frequencies in Leonbergers and 7 other related breeds. *Including 10 obligate carriers (dam/sire of LEMP cases). #Owner reported no signs of LEMP in dogs older than 8 years. §Dogs without known phenotype status that were submitted for diagnostic purposes. [1]Includes the two cases reported before[13]. [2]Allele frequency difference $p < 1.5 \times 10^{-89}$.

A missense variant in the *NAPEPLD* gene is associated with LEMP in Leonbergers. Whole genome re-sequencing (WGS) was performed on a Leonberger LEMP case homozygous for the associated haplotype. Subsequently, sequence variants in the mapped interval were called. The pedigree analysis (Supplementary Fig. S2) and the large size of the homozygous interval indicated a relatively young origin of the variant and a likely recessive mode of inheritance. Thus, we assumed that the causative variant should be absent from breeds unrelated to the Leonbergers. A total of 32 variants in the interval unique to the sequenced case remained after filtering against 201 control genomes of 66 different dog breeds and three wolves (Supplementary Table S2). Only a single variant (chr18:g.16987520 G > C) was predicted to affect the coding sequence of an annotated gene (Supplementary Table S3). Sanger sequencing confirmed the presence of this variant (Fig. 2a) and its nearly perfect association to the LEMP phenotype (Supplementary Fig. S2). This private non-synonymous variant

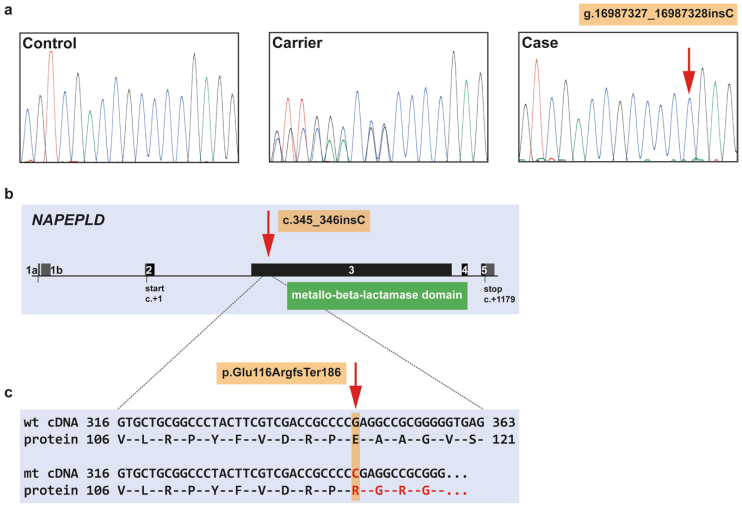

Figure 3. The *NAPEPLD* frameshift variant detected in LEMP-affected Rottweilers. (**a**) Chromatograms of wild type, carrier, and an affected dog indicate the c.345_346insC variant. (**b**,**c**) The schematic representation of the canine *NAPEPLD* gene indicates that the 1 bp insertion is located in exon 3 and leads to a frameshift which is predicted to produce a novel 186 amino acid long C-terminus of NAPEPLD and replaces the metallo-beta-lactamase domain of the wild type protein.

Breed	Status	Total	wt/wt (homozygous normal; N/N)	ins/wt (heterozygous; D/N)	ins/ins (homozygous mutant; D/D)
Rottweiler		233	212	17	4[1]
	Affected[2]	4			4[1]
	Non-affected[2]	72	64	8	
	Unknown[§]	157	148	9	
Great Danes[§]		262	238	24	
St. Bernard[§]		47	47		
Newfoundland[§]		10	10		
Entlebucher Mountain dog[§]		10	10		
Appenzeller Mountain dog[§]		8	8		
Bernese Mountain dog[§]		7	7		
Greater Swiss Mountain dog[§]		6	6		

Table 2. *NAPEPLD* c.345_346insC genotype frequencies in Rottweilers, Great Danes and 6 other related breeds. [§]Dogs without known phenotype status that were submitted for diagnostic purposes. [1]Includes the two cases reported before[11,12]. [2]Allele frequency difference $p < 2.4 \times 10^{-17}$.

in the *N-acyl phosphatidylethanolamine phospholipase D* (*NAPEPLD*) gene (c.538 G > C) is located within the metallo-beta-lactamase domain of the encoded protein (Fig. 2b). It is predicted to alter the sequence of codon 180 resulting in the replacement of alanine by proline (p.Ala180Pro). Multiple species amino acid sequence alignment showed that the wild type residue at the affected position is conserved across NAPEPLD orthologues in vertebrates including the zebrafish (Fig. 2c). Software-based analysis of the NAPEPLD amino acid exchange characterized the variant as probably damaging (PolyPhen 2), deleterious (SIFT), pathogenic (MutPred2) or disease causing (Mutation Taster). An mRNA-seq experiment on a spinal cord sample of a LEMP-affected Leonberger was carried out and revealed no evidence for alternative splicing of *NAPEPLD* in comparison to a spinal cord sample from a control dog (not shown).

To further investigate the frequency of the *NAPEPLD* variant in Leonbergers, and its segregation with disease, we genotyped in total 7,086 dogs (Table 1). This includes all 200 dogs from the GWAS, two previously described cases[13], 11 additional LEMP-affected Leonbergers, as well as 574 dogs with owner reported health updates indicating no neurological problems within the first eight years of life. There was a highly significant difference in *NAPEPLD* allele frequencies between LEMP cases and these controls using a standard chi-square test ($p < 1.5 \times 10^{-89}$). The absence of affected heterozygotes in this large cohort supports a recessive mode of inheritance (standard chi-square test $p < 7.9 \times 10^{-108}$). However, approximately 1% (6/574) of non-affected dogs at eight years of age were homozygous for the mutant allele, indicating reduced penetrance (Table 1). Furthermore,

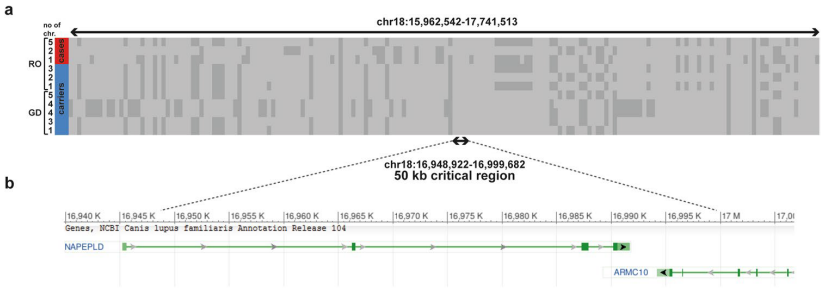

Figure 4. Across breed haplotype analysis in Rottweilers and Great Danes indicates an old mutation event. (**a**) Diverse haplotypes were detected exploring SNP array genotypes of four LEMP-affected Rottweilers (RO) and 23 heterozygous carriers of the *NAPEPLD* frameshift variant in Rottweilers and Great Danes (GD). Each line represents a unique haplotype. (**b**) A 50 kb-sized identical haplotype in all dogs contains segments of the canine *NAPEPLD* and *ARMC10* genes.

we genotyped nearly 6,500, mostly young, Leonbergers with unknown phenotype status, that were submitted for diagnostic purposes, which well represents the global population of the breed. The genotype frequency of homozygous mutant dogs in this cohort was 0.6%. Altogether, the mutant allele frequency in the global population of 7,086 Leonbergers was estimated as 8.5%.

Finally, the analysis of 28 heterozygous dogs carrying the LEMP-associated *NAPEPLD* allele identified shorter versions of the variant-containing haplotype due to recombination events. This enabled narrowing of the shared region surrounding the *NAPEPLD* allele on chromosome 18 to approximately 0.9 Mb (bp position 16,611,402 to 17,559,984; Fig. 1b). This shared haplotype contains 13 annotated protein-encoding genes and one pseudogene (Fig. 1c).

A frameshift variant in the *NAPEPLD* gene is associated with LEMP in Rottweilers and also occurs in Great Danes. We utilized a total of four LEMP-affected Rottweilers to evaluate whether the *NAPEPLD* gene harbored possible disease-causing variants in this breed. This cohort includes the two previously described cases[11,12] and two additional dogs (Supplementary Table S4). All four coding exons of the *NAPEPLD* gene were Sanger sequenced in the LEMP-affected Rottweilers. All were clear of the Leonberger missense variant, but a frameshift variant in exon 3 (c.345_346insC) was discovered in which the four LEMP-affected Rottweilers were homozygous (Fig. 3a). This 1 bp insertion is predicted to replace the C-terminal metallo-beta-lactamase domain of the wild type protein by a recoded peptide of 186 amino acids (p.Glu116ArgfsTer186) without any sequence similarity (Fig. 3b,c). The *NAPEPLD* frameshift allele (chr18: g.16987327_16987328insC) frequency in a population of 229 non-affected Rottweilers was 3.7% with no homozygous mutant dogs observed (Table 2). To confirm the focus on the *NAPELD* gene we performed a GWAS with 3 affected Rottweilers with passing SNP genotyping call rate and identified the same chromosome 18 locus ($p_{corrected} = 7.06 \times 10^{-15}$) as found in Leonbergers, that spanned the *NAPEPLD* gene (Supplementary Fig. S3). Further, WGS of an LEMP-affected Rottweiler confirmed the 1 bp insertion which was identified via Sanger sequencing and showed no further variants in the *NAPEPLD* gene. We also screened related breeds for the presence of this frameshift variant and identified non-affected heterozygous dogs in Great Danes, and an allele frequency of 4.5% in this breed cohort (Table 2). This variant was not present among the 201 sequenced control genomes of 66 different dog breeds and three wolves (Supplementary Table S2).

Subsequent haplotype analysis of 170k SNP array genotypes revealed that the *NAPEPLD* frameshift variant was present on three diverse haplotypes in affected Rottweilers (Fig. 4a). We then genotyped additional heterozygous Rottweilers and Great Danes to study the haplotype diversity in a 1.5 Mb interval surrounding the *NAPEPLD* gene. This revealed a collection of extended haplotypes associated with the frameshift variant (Fig. 4a). Nonetheless, an identical 50 kb sub-haplotype containing segments of the canine *NAPEPLD* and *ARMC10* genes was identified in all homozygous and heterozygous dogs (Fig. 4b).

Variable histopathological phenotype in *NAPEPLD* homozygous dogs within and across breeds. After variant identification we histopathologically re-evaluated eight neurologically affected Leonbergers genotyped as homozygous mutant for the c.538 G > C variant, including the two previously described cases[13], and three neurologically affected Rottweilers genotyped as homozygous mutant for the c.345_346insC variant, including two reported LEMP-affected dogs from Germany[11] and the US[12]. Stained tissue sections available for re-evaluation varied between dogs, however clear variation in histological lesions was observed. Nine out of these eleven dogs had histopathological lesions compatible with LEMP (Leonberger cases L1-6 (Supplementary Table S1), Rottweiler cases R1-3 (Supplementary Table S4)). In all nine dogs the spinal cord was affected. Additionally, in four LEMP-affected dogs, of which the brain was also available, a specific pattern with lesions involving the spinal tract of the trigeminus, cerebellar peduncles, cerebellar medulla, pyramids, crus cerebri, and optic tract was observed. No lesions were present in the spinal nerve roots. Generally, the nine dogs compatible with LEMP suffered marked loss of myelin (Fig. 5). The myelin was replaced by numerous fibrillary and

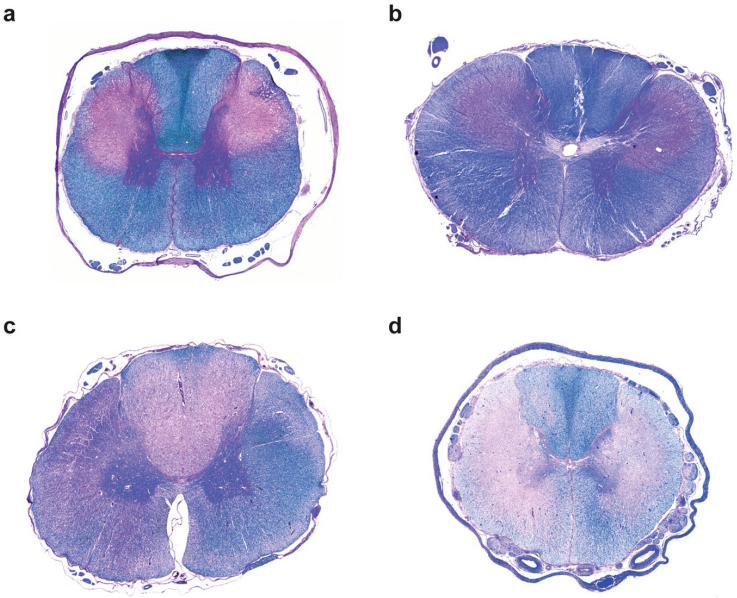

Figure 5. Phenotypic variability of transverse spinal cord sections of LEMP-affected dogs. Combined luxol fast blue/hematoxylin & eosin stain of paraffin sections. (**a**) Thoracic spinal cord of a Leonberger case L1 with typical LEMP lesions as previously described[13]. Note the bilateral-symmetrical loss of myelin in the corticospinal tracts as indicated by the loss of the blue color. (**b**) Cervical spinal cord of a previously described[12] Rottweiler case R1 with similar lesions as in (**a**). This dog had severe white matter loss in some areas combined with axonal loss, infiltration by macrophages and capillary hypertrophy. (**c**) Thoracic spinal cord of Leonberger case L3. The quality of lesions is similar as in (**a**), but the distribution is different with lesions being most severe in the dorsal tracts. (**d**) Thoracic spinal cord of Rottweiler case R3 showing similar lesions, which are less defined and more widespread encroaching into lateroventral tracts.

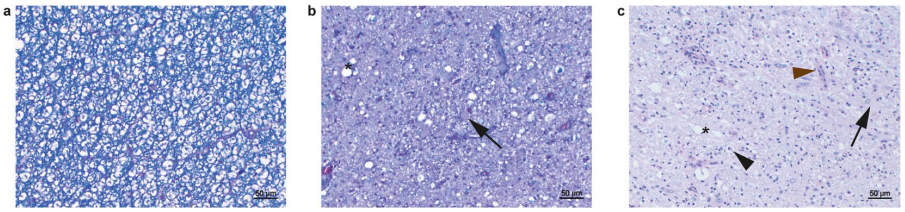

Figure 6. Histopathology of the spinal cord of LEMP-affected dogs. Combined luxol fast blue/hematoxylin & eosin stain of paraffin sections. (**a**,**b**) Spinal cord sections of previously described[11] Rottweiler case R2 showing deep blue staining of the normal myelin in an unaffected area (**a**) and an affected spinal cord area (**b**) exhibiting loss of blue myelin staining, vacuoles (asterisk), and large gemistocytic astrocytes (arrow). (**c**) Affected spinal cord area of the previously described Rottweiler case R1[12] exhibiting severe loss of blue myelin staining, vacuoles (asterisk), infiltration by macrophages (black arrowhead), capillary hypertrophy (brown arrowhead) and large gemistocytic astrocytes (arrow).

gemistocytic astrocytes (Fig. 6). Within and around the gliotic area scattered dilated myelin sheaths containing gitter cells were present, and a few scattered swollen axons were observed. Minimal Wallerian-like degeneration was observed in areas distant from the areas of myelin loss. Lesions were predominantly bilaterally symmetrical (Fig. 5), although in some areas a marked asymmetry could be observed.

There was variation in severity and topography between cases. Three cases (L1, L2, and R2) exhibited lesions predominantly affecting the lateral corticospinal tract and encroaching on the dorsal spinocerebellar, rubrospinal and the lateral spinothalamic tracts as previously described. In three dogs (R1, L5, and L6) lesions were similar in distribution, but very severe, being characterized by severe loss of white matter with infiltration of macrophages and capillary prominence. In three other cases (R3, L3 and L4) the distribution was different with lesions being most severe in the dorsal funiculi or being more widespread. In most dogs, axonal changes were relatively mild

and confined to the areas of severe myelin loss indicating a primary myelin disorder. In contrast, one Rottweiler case (R3) with severe white matter lesions exhibited more conspicuous axonal degeneration, and Bielschowsky stain revealed reduced axonal density in these areas.

Interestingly, two Leonbergers exhibited histopathological lesions not compatible with LEMP. One Leonberger case (L8) suffered axonal degeneration in the peripheral nerves associated with denervation atrophy of the skeletal muscle (Supplementary Fig. S4) compatible with Leonberger polyneuropathy[15]. Additionally, this dog exhibited scattered axonal degeneration in the spinal cord without conspicuous myelin loss (Supplementary Fig. S4). The second deviating Leonberger case (L7) had no lesions in the examined spinal cord section, but a well-defined, plaque-like area of demyelination and gliosis in the corona radiata (Supplementary Fig. S5).

Discussion

We investigated a possible genetic basis for leukoencephalomyelopathy (LEMP) in Leonbergers and Rottweilers. The pathological lesions in both breeds were reported clearly to be demyelinating but the distribution, restricted to the white matter of brain and spinal cord, remained unclear[11,13]. Our studies revealed two independent non-synonymous variants affecting the canine *NAPEPLD* gene, a gene with no known role in myelination or myelinogenesis of oligodendrocytes.

The endocannabinoid system consists of endocannabinoids, cannabinoid receptors and enzymes, such as NAPEPLD, involved in the synthesis and degradation of endogenous ligands[16]. It is a widespread neuromodulatory system which plays important roles in central nervous system development, synaptic plasticity, and the response to endogenous and environmental insults[16]. Some endocannabinoids are supposed to have a neuroprotective function[17]. The NAPEPLD protein is a membrane-bound phospholipase D type enzyme that catalyzes the release of N-acylethanolamine from N-acyl-phosphatidylethanolamine[18], and in so doing generates N-arachidonoylethanolamine, a ligand of cannabinoid and vanilloid receptors[19]. NAPEPLD is the enzyme catalyzing the production of arachidonoylethanolamine in animal tissues, and is structurally different from other known phospholipases D[19]. Nonetheless, it belongs to the zinc metallohydrolase family of the beta-lactamase fold, which is involved in a variety of biological events including antibiotic resistance, DNA repair and RNA maturation[20]. Several single point mutants, affecting highly conserved histidine and aspartic acid residues in the metallo-beta-lactamase domain involved in the binding to Zn^{2+}, are known to be catalytically inactive or have reduced enzyme activity[20,21]. The NAPEPLD protein forms a homodimer composed of two interconnected subunits, partly separated by an internal channel, and uniquely adapted to associate with phospholipids[22]. Recently, it was found that binding of bile acid enhances dimer assembly, and stimulates the NAPEPLD enzyme to favor the selective production of the endocannabinoid anandamide and other fatty acid ethanolamides[23].

Mice with a targeted disruption of the *NAPEPLD* gene were viable and healthy, but displayed reductions in very long-chain saturated and mono-unsaturated N-acyl ethanolamines in the central nervous system[24]. Brain tissue from mice lacking *GDE1* and *NAPEPLD* showed a near-complete loss in N-acyl-phosphatidylethanolamine conversion to N-acyl ethanolamines, but bulk brain levels of N-acyl ethanolamines were unaltered[25]. It was concluded that, both GDE1 and NAPEPLD make partial contributions to the biosynthesis of anandamide and other N-acyl ethanolamines *in vivo*[25]. No brain pathology was investigated in these knock-out mice. Pharmacological administration of agonists and antagonists of cannabinoid receptors in rats showed that the activation of both receptors is needed to augment the expression of myelin basic protein in the subcortical white matter[26]. Although most studies have focused on the role of endocannabinoids in neuronal differentiation, essential data suggest that these signalling lipids are also important in myelination of long-range axons to increase their conductance velocity[27]. Therefore we speculate that disruption in the production of endocannabinoids by a mutation in the *NAPELPD* gene could be a mechanism of demyelination. Recently, a marked upregulation of type-2 cannabinoid receptors in the spinal cord of dogs showing *SOD1*-associated degenerative myelopathy was reported, supporting the assumed neuroprotective function of the endocannabinoid system in neurodegenerative disorders[28]. For the first time, our study implicates a significant role of the NAPEPLD enzyme, and thereby the endocannabinoid system, in maintaining the white matter. In LEMP-affected dogs it was previously speculated that the primary defect might be located in neurons or axons[13]. Therefore the NAPEPLD enzyme could be involved in the axon-myelin crosstalk and might improve the understanding of the currently poorly characterized molecular and cellular mechanisms that impact the differentiation of oligodendrocytes and myelination[29]. Furthermore, there is as yet no association known for involvement of the *NAPEPLD* gene in myelin disorders in people[19]. The group of heterogeneous leukoencephalopathies, characterized by white matter abnormalities affecting the CNS, is associated with various genes; still, the genetic cause in about every second patient is unknown[30,31]. We suggest the *NAPEPLD* gene is worthy of investigation as a possible molecular basis of genetic leukoencephalopathies.

There are 173 observed human *NAPEPLD* coding region variants depicted in Exac Browser[32]: 43 are synonymous, 123 are missense, and 4 are loss of function; with 3 CNVs. The synonymous, missense and CNV variants all have positive Z scores, indicating increased intolerance to variation and fewer variants in the gene than expected. The probability of loss of function intolerance is pLI = 0.06, which places *NAPEPLD* gene variants as most likely being recessive, where heterozygous loss of function variants are often tolerated[32]. One missense human SNP (rs367936558) alters the same position of the human NAPEPLD protein as the variant found in Leonbergers; however the alanine residue is exchanged by valine (p.Ala180Val), not by proline as found here. This human variant is present in only 1 out of 60,704 patients, and in heterozygous state (allele frequency of $8.237e^{-06}$)[32].

The two disease-associated canine *NAPEPLD* variants affect the highly conserved metallo-beta-lactamase domain: changing a single amino acid in LEMP-affected Leonbergers, and removing the entire domain by a frameshift in Rottweilers. Interestingly, with the exception of two dogs, the general disease phenotype in these two breeds is very similar, although we have noticed a certain variability comparing dogs within the affected breeds sharing the identical homozygous mutant genotype. This could be explained either by disease progression over time, or by the individual genetic background and epistatic effects to unknown variation in possible modifier genes[33].

Recently it was shown that variations in SP110-mediated gene transcription may underlie, at least in part, the variability in risk for developing canine degenerative myelopathy among Pembroke Welsh Corgis that are homozygous for the disease-related *SOD1* mutation[34]. Similar efforts might be carried out to identify modifier genes explaining the observed variability in developing LEMP in Leonbergers. In addition, by genotyping several thousand Leonbergers, we noticed about 1% of dogs harboring the homozygous mutant genotype with no reported clinical manifestation (Table 1). This indicates a reduced penetrance of the *NAPEPLD* missense variant and supports recent findings in human genetics indicating that incomplete penetrance for presumed Mendelian diseases is likely more common than previously believed[35].

Furthermore, the neuropathological examination of eight *NAPEPLD* homozygous mutant Leonbergers indicated that one dog had a primary axonal disorder, not a myelin disorder, which was first noticed clinically as late as 7.5 years of age (Leonberger case L8, Supplementary Table S1). As we have learned recently in polyneuropathy-affected Leonbergers, where only approximately every third polyneuropathy-diagnosed Leonberger can be explained by the reported *ARHGEF10* or *GJA9* variants[36,37], this case could possibly be explained by an independent mutation in an unknown gene causing polyneuropathy, although a possible effect of the *NAPEPLD* genotype could not be ruled out. Altogether, this nicely highlights the limits to precisely clinically diagnosing neurological diseases in dogs. Furthermore, we noticed in Leonbergers that some dogs were initially diagnosed as polyneuropathy-affected, although in fact they were suffering from LEMP. Disease awareness has to be taken into account as well, when dogs that have not been diagnosed by a board-certified neurologist are used for genetic studies. Finally, variation in histopathological phenotypes and genotype-phenotype correlations in the population may also be influenced by the fact that the samples available were examined retrospectively from dogs of varying age and geographic locations.

In the Rottweilers it is unclear whether the truncated NAPEPLD protein produced by the frameshift variant, with more than 70% of the normal protein missing, is actually expressed. Furthermore, it is very unlikely that the predicted mutant protein with a recoded peptide of 186 amino acids before the newly encoded stop codon, without any sequence similarity to the normal protein (Supplementary Fig. S6), would fulfill any physiological function. This mutant protein would also have the functionally important metallo-beta-lactamase domain missing (Fig. 3). It is therefore more likely that the mutant mRNA is targeted by non-sense-mediated decay, thus the deleterious canine *NAPEPLD* variant represents the most likely causative variant in the four LEMP-affected Rottweilers.

Our results provide strong evidence for allelic heterogeneity in canine LEMP, where two independent *NAPEPLD* variants were found in Leonbergers and Rottweilers. In addition, we found several Great Danes being heterozygous carriers of the variant identified in Rottweilers, but so far we have not seen a homozygous mutant dog in this third breed. Although all three breeds are part of the Molossian section of the World Canine Organization classification, Leonbergers belong to the Mountain type dogs, and both Rottweilers and Great Danes represent the group of Mastiff type dogs. According to a recent study on the development of modern dog breeds all three breeds belong to separate clades, but show significant haplotype sharing[38]. This example of a similar canine disease occurring in different breeds caused by independent variants affecting the same gene, is comparable to what has been shown before, for example, in canine *NDRG1*-related polyneuropathy (OMIA 002120-9615)[39,40]. The missense variant detected in Leonbergers is probably caused by a recent mutation event as the associated haplotype encompasses a 0.9-Mb-sized region. On the other hand, the frameshift variant detected in LEMP-affected Rottweilers and in Great Dane carriers most likely has a quite old origin, as it exists on a relatively small-sized common haplotype of 50 kb.

In conclusion, here we report the identification of two *NAPEPLD*-associated variants in LEMP-affected dogs. Our results indicate a recessive mode of inheritance in each, albeit with a slightly reduced penetrance, and enable the development of genetic tests for veterinary diagnostic and breeding purposes. Our study describes a canine neurological disease with distinctive pathological features and implicates the NAPEPLD protein as an important enzyme in myelin homeostasis. Finally, our results reveal a novel candidate gene for myelin disorders such as the genetic leukoencephalopathies in humans.

Methods

Animals. Written consent was obtained from all dogs' owners. Dog samples were obtained primarily via elective owner submission for diagnostic purposes or were submitted for genotyping of the previously reported polyneuropathy-associated variants in Leonbergers. Most of these dogs do not have complete medical information and were used only for a population study. Blood collection from dogs does not require anesthesia and the study was approved according to the national guidelines for animal welfare by the Institutional Animal Care and Use Committees (IACUC) of the University of Minnesota, and by the Cantonal Committee for Animal Experiments (Canton of Bern; permits 23/10, 48/13 and 75/16) for the University of Bern.

All invasive procedures were performed *post-mortem* either on animals that had died of natural causes, or after euthanasia, thus no ethical evaluation was required. All methods were performed in accordance with the relevant guidelines and regulations of the University of Minnesota and the University of Bern.

Genomic DNA was isolated from blood using either the Gentra PureGene blood kit (Qiagen) or the Maxwell RSC whole blood DNA kit (Promega). The phenotypic characterization of LEMP in Leonbergers and Rottweilers has been described elsewhere and the previously established criteria to select cases and controls were applied[36]. Samples from a total of 7,086 Leonbergers, including 213 dogs with detailed phenotype records (Supplementary Table S1), and 233 Rottweilers, including four dogs with detailed phenotype records (Supplementary Table S4) were used during this study. All Leonbergers were genotyped for the polyneuropathy-associated *ARHGEF10* and *GJA9* variants as previously described[36,37] to test for underlying neurological disease. Furthermore, DNA samples of 262 Great Danes, 47 St. Bernards, 10 Newfoundlands, 10 Entlebucher Mountain dogs, 8 Appenzeller Mountain dogs, 7 Bernese Mountain dogs, and 6 Greater Swiss Mountain dogs were taken from the Vetsuisse Biobank.

Histopathology. Histopathological samples of eight Leonbergers and four Rottweilers were examined. Two to five µm thick sections from 10% neutral-buffered formalin fixed and paraffin-embedded brain and spinal cord were retrieved from histopathological archives of different diagnostic labs. These included combined luxol fast blue/hematoxylin and eosin stained sections or combined luxol fast blue/periodic acid Schiff stained sections. In five cases, Bielschowsky silver stained sections and in one case sections labeled with PGP 9.5 immunohistochemistry were available for evaluation of axonal density. Tissue sections were examined by light microscopy.

SNP array genotyping. Genomic DNA samples of 200 Leonbergers, 10 Rottweilers, and 17 Great Danes were genotyped with the Illumina CanineHD BeadChip array by GeneSeek/Neogen for 173,662 SNP markers. In Leonbergers we performed pruning of genotyping data as described previously[36] and 112,833 SNPs remained for genome-wide association study (GWAS). Fourteen LEMP-affected dogs and 186 controls (i. e. dogs eight years and older that showed no signs of neurological disease) were analyzed with the mixed model from the GenABEL library[41] and the hglm package[42] in the R environment that corrects for the population stratification. Multidimensional scaling analysis was carried with the GenABEL[41]. We used 173k SNP data of additional 147 Rottweilers as controls, which were publically available from previous projects[43,44], to perform GWAS. Due to the limited number of cases in Rottweilers, only a fast score test (GenABEL) for association corrected for possible stratification by principal components was performed. Haplotypes around the significantly associated locus were constructed using fastPHASE[45]. All genome positions refer to the CanFam3.1 reference sequence assembly.

Whole genome sequencing. We performed a whole-genome sequencing of a LEMP-affected Leonberger. Briefly, we prepared a fragment library with 300 bp insert size and collected ~200 million 2 × 100 bp paired-end reads on a HiSeq. 2000 instrument (Illumina, San Diego, USA), which corresponds to roughly 17x coverage. The reads were mapped against the dog reference genome assembly (CanFam3.1) as described before[36]. The annotation version CanFam3.1.75 (http://www.ensembl.org) was used to predict the functional effects of detected variants as described previously[36]. In addition, whole-genome sequencing of a LEMP-affected Rottweiler (case R2; 13× coverage) was performed accordingly.

The IGV-viewer software[46] was used for visual inspection of sequence variants to exclude any structural variants in the critical region. For variant filtering we used 204 control genomes, which were either publicly available[47] or produced during other projects of our group. A list of these control genomes is given in Supplementary Table S2.

Gene analysis. We used the dog CanFam 3.1 reference genome assembly for all genomic analyses. The chromosome 18 reference sequence has a gap after *NAPEPLD* exon 4 and does not include exon 5 (Supplementary Fig. S7). A partial sequence of intron 3, the entire exon 4, intron 4 and exon 5 are retrievable on the unplaced contig Un_JH373889 of the CanFam 3.1 assembly (Supplementary Fig. S7). Numbering within the canine *NAPEPLD* transcript corresponds to the mRNA sequences within study accession PRJEB22251. The predicted effects of the mutations were evaluated by PolyPhen2[48], SIFT[49], Mutation Taster[50], and MutPred2[51].

RNA-seq. We isolated total RNA from spinal cord samples from a single LEMP-affected Leonberger and a Labrador control using the RNeasy Fibrous Tissue Mini kit (Qiagen). Prior to RNA extraction, the tissue was mechanically disrupted using the TissueRuptor device (Qiagen). The RNA samples were transformed into Illumina TruSeq libraries and 2 × 150 bp sequencing reads were obtained on a HiSeq3000 instrument (Illumina). RNA-seq data analysis was done as described before[52].

Sanger sequencing. We used Sanger sequencing to confirm the candidate *NAPEPLD* variant c.538 G > C in the affected Leonberger and to amplify the coding exons of the *NAPEPLD* gene in the affected Rottweilers. We amplified PCR products (primers are shown in Supplementary Table S5) using AmpliTaqGold360Mastermix (Life Technologies) and purified PCR products were directly sequenced on an ABI3730 capillary sequencer (Life Technologies). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes).

Data Availability. Genome sequencing data were deposited in the European Nucleotide Archive (ENA, http://www.ebi.ac.uk/ena): The LEMP-affected Leonberger (sample accession number SAMEA103935360 within study accession PRJEB16012) and the LEMP-affected Rottweiler (sample accession number SAMEA3121337 within study accession PRJEB7735). RNAseq data were deposited in the ENA under sample accession number SAMEA103936001 within study accession PRJEB20118. The mRNA sequences for the canine *N-acyl phosphatidylethanolamine phospholipase D (NAPEPLD)* gene were deposited in the ENA under sample accession number LT906616 and LT906617 within study accession PRJEB22251.

References

1. Vanderver, A. *et al.* Case definition and classification of leukodystrophies and leukoencephalopathies. *Mol. Genet. Metab.* **114**, 494–500 (2015).
2. Di Rocco, M., Biancheri, R., Rossi, A., Filocamo, M. & Tortori-Donati, P. Genetic disorders affecting white matter in the pediatric age. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **129B**, 85–93 (2004).
3. Schiffmann, R. & Van Der Knaap, M. S. Invited article: an MRI-based approach to the diagnosis of white matter disorders. *Neurology* **72**, 750–759 (2009).
4. Kaye, E. M. Update on genetic disorders affecting white matter. *Pediatr. Neurol.* **24**, 11–24 (2001).
5. Wang, X. *et al.* The use of targeted genomic capture and massively parallel sequencing in diagnosis of Chinese leukoencephalopathies. *Sci. Rep.* **6**, 35936 (2016).
6. Summers, B. A., Cummings, J. F. & de Lahunta, A. Hereditary, familial, and idiopathic degenerative diseases in *Veterinary Neuropathology* 281–300 (Mosby, 1995).
7. Sisó, S. *et al.* Neurodegenerative diseases in domestic animals: a comparative review. *Vet. J.* **171**, 20–38 (2006).

8. Averill, D. R. & Bronson, R. T. Inherited necrotizing myelopathy of Afghan hounds. *J. Neuropathol. Exp. Neurol.* **36**, 734–747 (1977).
9. Gamble, D. A. & Chrisman, C. L. A leukoencephalomyelopathy of rottweiler dogs. *Vet. Pathol.* **21**, 274–280 (1984).
10. Baum, F. 3rd, de Lahunta, A. & Trotter, E. J. Cervical fibrotic stenosis in a young Rottweiler. *J. Am. Vet. Med. Assoc.* **201**, 1222–1224 (1992).
11. Hirschvogel, K. *et al.* Magnetic resonance imaging and genetic investigation of a case of Rottweiler leukoencephalomyelopathy. *BMC Vet. Res.* **9**, 57 (2013).
12. Eagleson, J. S., Kent, M., Platt, S. R., Rech, R. R. & Howerth, E. W. MRI findings in a rottweiler with leukoencephalomyelopathy. *J. Am. Anim. Hosp. Assoc.* **49**, 255–261 (2013).
13. Oevermann, A., Bley, T., Konar, M., Lang, J. & Vandevelde, M. A novel leukoencephalomyelopathy of leonberger dogs. *J. Vet. Intern. Med.* **22**, 467–471 (2008).
14. Mandigers, P. J. J., van Nes, J. J., Knol, B. W., Ubbink, G. J. & Gruys, E. Hereditary necrotising myelopathy in Kooiker dogs. *Res. Vet. Sci.* **54**, 118–123 (1993).
15. Shelton, G. D. *et al.* Inherited polyneuropathy in Leonberger dogs: a mixed or intermediate form of Charcot-Marie-Tooth disease? *Muscle Nerve.* **27**, 471–477 (2003).
16. Lu, H. C. & MacKie, K. An introduction to the endogenous cannabinoid system. *Biol. Psychiatry* **79**, 516–525 (2016).
17. Cassano, T. *et al.* Cannabinoid receptor 2 signaling in neurodegenerative disorders: From pathogenesis to a promising therapeutic target. *Front. Neurosci.* **11** (2017).
18. Okamoto, Y., Morishita, J., Tsuboi, K., Tonai, T. & Ueda, N. Molecular characterization of a phospholipase D generating anandamide and its congeners. *J. Biol. Chem.* **279**, 5298–5305 (2004).
19. Hussain, Z., Uyama, T., Tsuboi, K. & Ueda, N. Mammalian enzymes responsible for the biosynthesis of N -acylethanolamines. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids.* **1862**, 1546–1561 (2017).
20. Pettinati, I., Brem, J., Lee, S. Y., McHugh, P. J. & Schofield, C. J. The chemical biology of human metallo-β-lactamase fold proteins. *Trends Biochem. Sci.* **41**, 338–355 (2016).
21. Okamoto, Y., Wang, J., Morishita, J. & Ueda, N. Biosynthetic pathways of the endocannabinoid anandamide. *Chem. Biodivers.* **4**, 1842–1857 (2007).
22. Magotti, P. *et al.* Structure of human N-acylphosphatidylethanolamine-hydrolyzing phospholipase D: regulation of fatty acid ethanolamide biosynthesis by bile acids. *Structure* **23**, 598–604 (2015).
23. Margheritis, E. *et al.* Bile acid recognition by NAPE-PLD. *ACS Chem. Biol.* **11**, 2908–2914 (2016).
24. Leung, D., Saghatelian, A., Simon, G. M. & Cravatt, B. F. Inactivation of N-acyl phosphatidylethanolamine phospholipase D reveals multiple mechanisms for the biosynthesis of endocannabinoids. *Biochemistry* **45**, 4720–4726 (2006).
25. Simon, G. M. & Cravatt, B. F. Characterization of mice lacking candidate N-acyl ethanolamine biosynthetic enzymes provides evidence for multiple pathways that contribute to endocannabinoid production *in vivo*. *Mol. Biosyst.* **6**, 1411–1418 (2010).
26. Arévalo-Martín, A. *et al.* Cannabinoids modulate Olig2 and polysialylated neural cell adhesion molecule expression in the subventricular zone of post-natal rats through cannabinoid receptor 1 and cannabinoid receptor 2. *Eur. J. Neurosci.* **26**, 1548–1559 (2007).
27. Maccarrone, M., Guzmán, M., Mackie, K., Doherty, P. & Harkany, T. Programming of neural cells by (endo)cannabinoids: from physiological rules to emerging therapies. *Nat. Rev. Neurosci.* **15**, 786–801 (2014).
28. Fernández-Trapero, M. *et al.* Upregulation of CB2 receptors in reactive astrocytes in canine degenerative myelopathy, a disease model of amyotrophic lateral sclerosis. *Dis. Model. Mech.* **10**, 551–558 (2017).
29. Tanaka, T. & Yoshida, S. Mechanisms of remyelination: recent insight from experimental models. *Biomol. Concepts.* **5**, 289–298 (2014).
30. Zhang, J. *et al.* A founder mutation in VPS11 causes an autosomal recessive leukoencephalopathy linked to autophagic defects. *PLoS Genet.* **12**, e1005848 (2016).
31. Arai-Ichinoi, N. *et al.* Genetic heterogeneity in 26 infants with a hypomyelinating leukodystrophy. *Hum. Genet.* **135**, 89–98 (2016).
32. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* **536**, 285–291 (2016).
33. Riordan, J. D. & Nadeau, J. H. From peas to disease: modifier genes, network resilience, and the genetics of health. *Am. J. Hum. Genet.* **101**, 177–191 (2017).
34. Ivansson, E. L. *et al.* Variants within the SP110 nuclear body protein modify risk of canine degenerative myelopathy. *Proc. Natl. Acad. Sci. USA* **113**, E3091–E3100 (2016).
35. Chen, R. *et al.* Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat. Biotechnol.* **34**, 531–538 (2016).
36. Becker, D. *et al.* A *GJA9* frameshift variant is associated with polyneuropathy in Leonberger dogs. *BMC Genomics* **18**, 662 (2017).
37. Ekenstedt, K. J. *et al.* An *ARHGEF10* deletion is highly associated with a juvenile-onset inherited polyneuropathy in Leonberger and Saint Bernard dogs. *PLoS Genet.* **10**, e1004635 (2014).
38. Parker, H. G. *et al.* Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Rep.* **19**, 697–708 (2017).
39. Drögemüller, C. *et al.* A deletion in the N-Myc downstream regulated gene 1 (*NDRG1*) gene in Greyhounds with polyneuropathy. *PLoS ONE* **5**, e11258 (2010).
40. Bruun, C. *et al.* A Gly98Val mutation in the N-Myc downstream regulated gene 1 (*NDRG1*) in Alaskan malamutes with polyneuropathy. *PLoS ONE* **8**, e54547 (2013).
41. Aulchenko, Y. S., Ripke, S., Isaacs, A. & van Duijn, C. M. GenABEL: An R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).
42. Ronnegard, L., Shen, X. & Alam, M. hglm: A package for fitting hierarchical generalized linear models. *R J.* **2**, 20–28 (2010).
43. Vaysse, A. *et al.* Identification of Genomic Regions Associated with Phenotypic Variation between Dog Breeds using Selection Mapping. *PLoS Genet.* **7**, e1002316 (2011).
44. Karlsson, E. *et al.* Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome Biol.* **14**, R132 (2013).
45. Scheet, P. & Stephens, M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* **78**, 629–644 (2006).
46. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative genomics viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
47. Bai, B. *et al.* DoGSD: the dog and wolf genome SNP database. *Nucleic Acids Res.* **43**, D777–D783 (2015).
48. Adzhubei, I. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
49. Kumar, P., Henikoff, S. & Ng, P. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
50. Schwarz, J., Cooper, D., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362 (2014).
51. Pejaver, V. *et al.* MutPred2: inferring the molecular and phenotypic impact of amino acid variants. bioRxiv 134981 (2017).
52. Waluk, D. P. *et al.* A splice defect in the EDA gene in dogs with an X-linked hypohidrotic ectodermal dysplasia (XLHED) phenotype. *G3 (Bethesda)* **6**, 2949–2954 (2016).

Acknowledgements

The authors are grateful to the dog owners who donated samples and shared pedigree data of their dogs. We thank Nathalie Besuchet Schmutz, Muriel Fragnière, and Sabrina Schenk for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the high-throughput sequencing experiments, and the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure. We thank the Dog Biomedical Variant Database Consortium (Gus Aguirre, Catherine André, Oliver Forman, Steven Friedenberg, Urs Giger, Christophe Hitte, Marjo Hytönen, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Claire Wade) for sharing whole genome sequencing data from control dogs and wolves. We also acknowledge all canine researchers who deposited dog whole genome sequencing data into public databases. Partial funding for E.F. was provided by the Office of the Director, National Institutes of Health (NIH) under award number K01OD019912.

Author Contributions

K.M.M. and A.L. performed phenotype assessment, GWAS mapping, haplotype analysis, and genotyping. D.B., M.D., S.R.B., E.F. and K.J.E. supported genotyping. A.O., S.R.P., K.M. and G.D.S. evaluated biopsies of spinal cord. P.J.J.M., P.A.J.L., Q.E.M.S., K.P., A.F., T.F., E.E.P., P.A.K., J.P.C., K.L., D.L.B., H.M., C.F.M., S.L.P. and D.H. contributed samples of LEMP-affected dogs. V.J. performed bioinformatics. J.R.M. and C.D. designed the study. A.L., A.O., C.D. and J.R.M. drafted the paper. All authors participated in writing the manuscript and have read and approved the final version.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-23938-7.

Competing Interests: Both the University of Minnesota and the University of Bern offer a genotyping test for LEMP, as well as several other genetic conditions in dogs, in their respective laboratories. None of the authors gain financially from the proceeds of these tests, which all go to support ongoing canine genetic research. Non-financial competing interests of authors from both Universities include informal advisory relationships concerning use and interpretation of genetic testing, as well as financial support for the research, from non-profit organizations that include the Leonberger Health Foundation (USA), the International Leonberger Union, and the Swiss and German Leonberger Clubs.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A missense mutation in the vacuolar protein sorting 11 (*VPS11*) gene is associated with neuroaxonal dystrophy in Rottweiler dogs

Journal:	G3 Genes|Genomes|Genetics
Manuscript status:	published
Contributions:	Resources, Validation, Writing – review & editing

G3 Genes | Genomes | Genetics

INVESTIGATION

A Missense Mutation in the Vacuolar Protein Sorting 11 (*VPS11*) Gene Is Associated with Neuroaxonal Dystrophy in Rottweiler Dogs

Katherine L. Lucot,* Peter J. Dickinson,† Carrie J. Finno,* Tamer A. Mansour,* Anna Letko,‡ Katherine M. Minor,§ James R. Mickelson,§ Cord Drögemüller,‡ C. Titus Brown,* and Danika L. Bannasch*,1

*Departments of Population Health and Reproduction, and †Surgical and Radiological Sciences, School of Veterinary Medicine, University of California, Davis, Davis, CA 95616, ‡Department of Clinical Research and Veterinary Public Health, University of Bern, 3001 Bern, Switzerland, and §Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, Minneapolis, MN 55455

ORCID ID: 0000-0001-9773-522X (C.D.)

ABSTRACT Canine neuroaxonal dystrophy (NAD) is a recessive, degenerative neurological disease of young adult Rottweiler dogs (*Canis lupus familiaris*) characterized pathologically by axonal spheroids primarily targeting sensory axon terminals. A genome-wide association study of seven Rottweilers affected with NAD and 42 controls revealed a significantly associated region on canine chromosome 5 (CFA 5). Homozygosity within the associated region narrowed the critical interval to a 4.46 Mb haplotype (CFA5:11.28 Mb – 15.75 Mb; CanFam3.1) that associated with the phenotype. Whole-genome sequencing of two histopathologically confirmed canine NAD cases and 98 dogs unaffected with NAD revealed a homozygous missense mutation within the Vacuolar Protein Sorting 11 (*VPS11*) gene (g.14777774T > C; p.H835R) that was associated with the phenotype. These findings present the opportunity for an antemortem test for confirming NAD in Rottweilers where the allele frequency was estimated at 2.3%. *VPS11* mutations have been associated with a degenerative leukoencephalopathy in humans, and *VSP11* should additionally be included as a candidate gene for unexplained cases of human NAD.

KEYWORDS
autophagy
canine
lysosome
neurodegenerative
inherited
genetic

Neuroaxonal dystrophy (NAD) is a relatively non-specific histopathological diagnosis for a group of neurodegenerative disorders characterized by dystrophic changes of the neuron followed by development of axonal swellings or spheroids (Revesz *et al.* 2015). First described by Cajal, (Cajal 1928) axonal swellings may occur in the central or peripheral nervous system and the underlying pathogenesis of the variable structured material found in these swellings is often poorly defined. The dystrophic phenotype may vary depending on age of onset, clinical manifestations, and specific disorder.

NAD can be divided into three major etiological groups: *physiological*, *secondary*, and *primary*. *Physiological* NAD can be seen commonly in humans and domesticated species as a component of aging (Suzuki *et al.* 1979; Saito 1980; Summers *et al.* 1995; Borràs *et al.* 1999; Gavier-Widen *et al.* 2001; Revesz *et al.* 2015), and *secondary* NAD may be seen focally, or more widely throughout the nervous system in response to a wide variety of conditions, including trauma, infection, toxin exposure and metabolic disease such as vitamin E deficiency or organophosphate exposure (Yagishita 1978; Summers *et al.* 1995; Revesz *et al.* 2015). Axonal spheroids have also been described in human patients with amyotrophic lateral sclerosis, Alzheimer's, Parkinson's disease and hereditary spastic paraparesis. *Primary* NAD is generally associated with a group of genetically heterogeneous, inherited neurodegenerative diseases where the presence of neuroaxonal dystrophy is a major pathological component of the

doi: https://doi.org/10.1534/g3.118.200376
Manuscript received May 1, 2018; accepted for publication June 19, 2018; published Early Online June 26, 2018.

Supplemental material available at Figshare: https://doi.org/10.25387/g3.6214010.
[1]Corresponding Author: Department of Population Health and Reproduction, School of Veterinary Medicine, One Shields Ave, University of California, Davis, CA 95616, E-mail: dlbannasch@ucdavis.edu

disease. In human patients, neuroaxonal dystrophic pathology has been associated to varying degrees with several genetically defined disease syndromes, most prominently in Infantile Neuroaxonal Dystrophy (INAD) and Pantothanate Kinase-associated Neurodegeneration associated with alterations in the *PLA2G6*, and *PANK2* genes respectively. Many of the human NAD syndromes are also associated with brain iron accumulation, including alterations in the *PLA2G6*, *PANK2*, *FTL*, *C19orf12*, *FA2H*, and *WDR45* genes (Revesz *et al.* 2015; Arber *et al.* 2016). Neuroaxonal dystrophy in humans without iron accumulation is seen in Wilson's disease and Nasu-Hakola disease involving the *ATP7B* and *DAP12/TREM2* genes respectively, and in "neuroaxonal leukoencephalopathy with axonal spheroids" which has no defined genetic cause to date (Revesz *et al.* 2015). *Primary* NAD has been reported in most domesticated species including dogs (Clark *et al.* 1982; Chrisman *et al.* 1984; Blakemore and Palmer 1985; Evans *et al.* 1988; Sacre *et al.* 1993; Franklin *et al.* 1995; Bennett and Clarke 1997; Siso *et al.* 2001; Nibe *et al.* 2007; Fyfe *et al.* 2010; Fyfe *et al.* 2011; Hahn *et al.* 2015; Pintus *et al.* 2016; Degl'Innocenti *et al.* 2017; Tsuboi *et al.* 2017), cats (Woodard *et al.* 1974; Carmichael *et al.* 1993; Rodriguez *et al.* 1996; Résibois and Poncelet 2004), cattle (Hanshaw *et al.* 2015), sheep (Cordy *et al.* 1967; Nuttall 1988; Harper and Morton 1991; Kessell *et al.* 2012; Finnie *et al.* 2014), pigeons (Barrows *et al.* 2017), mice (Bouley *et al.* 2006), and horses (Beech 1984; Blythe *et al.* 1991; Aleman *et al.* 2011; Finno *et al.* 2013; Finno *et al.* 2016), where an association with vitamin E deficiency, along with a genetic susceptibility, has been reported.

In dogs, breed related NAD has been reported as fetal onset in Giant Schnauzer-Beagle mix breed dogs (Fyfe *et al.* 2010; Fyfe *et al.* 2011), juvenile onset in Dachshund mix breed dogs (Pintus *et al.* 2016), Border collies (Clark *et al.* 1982), Chihuahuas (Blakemore and Palmer 1985; Degl'Innocenti *et al.* 2017), Jack Russell Terriers (Sacre *et al.* 1993), Papillons (Franklin *et al.* 1995; Nibe *et al.* 2007; Tsuboi *et al.* 2017), Spanish Water Dogs (Hahn *et al.* 2015), and young adult or adult onset in Rottweilers (Cork *et al.* 1983; Chrisman *et al.* 1984; Evans *et al.* 1988; Bennett and Clarke 1997; Siso *et al.* 2001) and English Cocker Spaniels (McLellan *et al.* 2003). NAD in Cocker Spaniels is accompanied by retinal degeneration and is associated with vitamin E deficiency (McLellan *et al.* 2003). Specific genetic mutations associated with the *PLA2G6*, *TECPR2* and *MFN2* genes have been identified in the Papillon, Spanish Water Dog and Schnauzer-Beagle cross dogs respectively (Fyfe *et al.* 2011; Hahn *et al.* 2015; Tsuboi *et al.* 2017).

Rottweiler NAD was first reported in the early 1980s and is characterized by a young adult age of onset with mild progression of clinical signs, typically including postural deficits, ataxia, hypermetria, intention tremor and nystagmus (Cork *et al.* 1983; Chrisman *et al.* 1984). Clinical signs reflect the predominantly sensory topographical distribution of pathology within the central nervous system (CNS) consisting of mild cerebellar atrophy, large numbers of axonal spheroids, and demyelination of axons in the vestibular nucleus, lateral and medial geniculate nuclei, sensory nucleus of the trigeminal nerve, gracilis and cuneate nuclei, and in the spinal cord dorsal horn (Cork *et al.* 1983; Chrisman *et al.* 1984). Based on a small pedigree, it was hypothesized to be an autosomal recessive disorder (Cork *et al.* 1983).

Defining underlying genetic mechanisms for breed related neuroaxonal dystrophies in dogs has the potential to provide biological insight and potential translational models for the heterogeneous disease phenotypes seen in human patients (Shearin and Ostrander 2010; Hytönen and Lohi 2016). A genome-wide association was therefore performed with samples from the original four reported Rottweiler cases (Cork *et al.* 1983; Chrisman *et al.* 1984) and three additional cases, together with whole-genome sequencing of selected cases to identify candidate genes for NAD in Rottweilers.

MATERIALS AND METHODS

Canine Samples

Buccal swabs or blood samples were collected from privately owned dogs through the William R. Pritchard Veterinary Medical Teaching Hospital at UC Davis. Collection of canine blood samples was approved by the University of California, Davis Animal Care and Use Committee (protocol #16892). Additional Rottweiler DNA samples were provided by the University of Minnesota and the University of Bern, Switzerland. These studies were approved according to the national guidelines for animal welfare by the Institutional Animal Care and Use Committees (IACUC) of the University of Minnesota, and by the Cantonal Committee for Animal Experiments (Canton of Bern; permits 23/10, 48/13 and 75/16) for the University of Bern. Owners specified the breed of each dog. Genomic DNA was extracted using the Qiagen kit (QIAGEN, Valencia, CA). Neurological phenotypes were determined by a veterinarian and confirmed postmortem via necropsy when available.

Genome-wide SNP Genotyping

Genome-wide SNP genotyping was performed on seven cases and 42 controls, using the Illumina CanineHD 220k BeadChip (Illumina, San Diego, CA, USA). All samples had a genotyping rate of $\geq$ 90%. 62,193 SNPs were excluded due to a minor allele frequency $\leq$ 5% and 7,421 SNPs were excluded due to a high genotype failure rate ($\geq$ 10%), leaving 151,799 SNPs after quality control. A Chi-square analysis and a genomic inflation factor (λ) was calculated with PLINK (Purcell *et al.* 2007). Homozygosity throughout the associated interval was analyzed by visual inspection assisted by color-coding homozygous genotypes in Excel. Homozygosity in the affected dogs, that passed the Bonferroni threshold ($P \leq 3.29\times10^{-7}$), was used to narrow down the regions of association and was visualized using Haploview (Barrett *et al.* 2005; Barrett 2009; Clarke *et al.* 2011). Figures were made in R using the ggplot2 package (Wickham 2009).

Whole-Genome Sequencing

Whole-genome sequencing was performed as described on 100 canine genomes, (Brown *et al.* 2017) with two histopathologically confirmed Rottweiler cases and 98 dogs unaffected with NAD, across 25 different breeds, including two Rottweilers. Sequencing was performed on the Illumina HiSeq 2000 using 100bp paired-end reads with approximately 8.7x coverage per sample. The reads were aligned to the canine reference genome (CanFam3.1) (Lindblad-Toh *et al.* 2005). Local realignment and variant calls were performed using the Genome Analysis Tool Kit (GATK version 2.5-2gf57256b) pipeline (McKenna *et al.* 2010). Biological consequences of variants were predicted using Ensembl's Variant Effect Predictor (VEP), PolyPhen-2 (v2.2.2r398), and SIFT (Adzhubei *et al.* 2010; Sievers *et al.* 2011; Sim *et al.* 2012; Adzhubei *et al.* 2013; McLaren *et al.* 2016;).

Genotyping by Sanger Sequencing

Primers were designed using Primer3 (Rozen and Skaletsky 2000) to validate the putative functional mutation uncovered in *VPS11* (F: CTGCAGGTCCCTGTCCTAAG; R: TGTACCTGGCTCTTGGCTCT). PCR products were sequenced using the Big Dye termination kit on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Sequences were evaluated using Chromas (Technelysium, South Brisbane, QLD, Australia). Sequences were aligned to CanFam3.1 using BLAT (UCSC Genome Browser). Allele frequency was calculated excluding the seven affected cases.

RNA Extraction and cDNA Sequencing

RNA was isolated from liver using Qiagen QIAamp Blood Mini Kit tissue protocols (QIAGEN, Valencia, CA). RNA was reverse transcribed into cDNA using Qiagen QuantiTect Reverse Transcription Kit. Ubiquitously expressed *VPS11* (F: TGGTCCAAAAACTGCAGAAA; R: CTCAAAGCAGTGTTGGTGGA) and the housekeeping gene *RPS5* (Brinkhof *et al.* 2006) cDNA were PCR amplified from liver tissue from two affected Rottweilers, one Gordon Setter, and one mixed breed dog. *RPS5* was amplified in liver to ensure equivalent amounts of cDNA were produced. The PCR products were sequenced on an ABI 3500 Genetic Analyzer and analyzed using Chromas (Technelysium, South Brisbane, QLD, Australia). The sequences were aligned to Can Fam3.1 using BLAT (UCSC Genome Browser) to confirm the missense mutation in the cDNA of *VPS11*.

Data Availability

The SNP genotyping data can be found in files Supplemental 1 (File S1), and Supplemental 2 (File S2). Whole-genome sequencing files reported in this paper can be found in the NCBI Sequence Read Archive (SRA Bioproject no. PRJNA377155). Sequences from four Pugs were made available in 2012 by TGEN (https://www.tgen.org/patients/canine/). Supplemental material available at Figshare: https://doi.org/10.25387/g3.6214010.

RESULTS

Case Definition

DNA samples were available from the four original NAD affected Rottweilers from Cork *et al.* (Cork *et al.* 1983) and three additional cases. One case was evaluated at the Veterinary Medical Teaching Hospital VMTH (University of California, Davis) with neurologic deficits and histopathological findings at necropsy consistent with previously reported cases. Blood samples from two additional dogs were submitted by their respective owners. Both dogs were presented to veterinarians with a history and clinical signs consistent with NAD and were noted to be "clumsy" as puppies. One dog was presented at approximately 1 year of age with generalized ataxia and hypermetria and absent menace responses. The second dog was presented for progressive ataxia and hypermetria that had been present for several years. Both dogs were alive and ambulatory at 2 and 5 years of age respectively. The second dog had been tested previously for mutations associated with two other neurodegenerative diseases reported in Rottweilers (leukoencephalomyelopathy (Minor *et al.* 2018), laryngeal paralysis-polyneuropathy (Mhlanga-Mutangadura *et al.* 2016)) and was negative for both mutations.

Genome-Wide Association Study

To identify loci associated with the NAD phenotype in the Rottweiler dog, a genome-wide association study was performed, followed by homozygosity analysis using seven cases affected with NAD and 42 healthy controls. Four of the seven cases were directly related resulting in a genomic inflation (λ) value of 1.52. A chi-square analysis of the 151,799 SNPs, identified preliminary associations on canine chromosomes (CFA) 4, 5, 12, 14, 16, 19, and 37 (Figure 1A). The lowest *P* value was on CFA 5 ($P = 1\times10^{-14}$) and 26 SNPs in this region were more associated than on the next highest chromosomal location. Since this disease is uncommon and pedigree analysis was consistent with a recessive mode of inheritance, a homozygosity analysis was performed in the cases. To identify regions of homozygosity in the cases, *P* values from markers with an allele frequency of 1 in the cases were plotted. There were 45 markers that met the Bonferroni correction ($P \leq 1\times10^{-4}$), and all but one (CFA 34; $P = 3.71\times10^{-5}$) were on CFA 5, making it the only statistically significant region of association that met the allele frequency criteria (Figure 1A-C). Homozygosity throughout the associated interval was used to narrow down the region of interest to 4.46 Mb (Chr5:11,282,754-15,754,443; CanFam3.1).

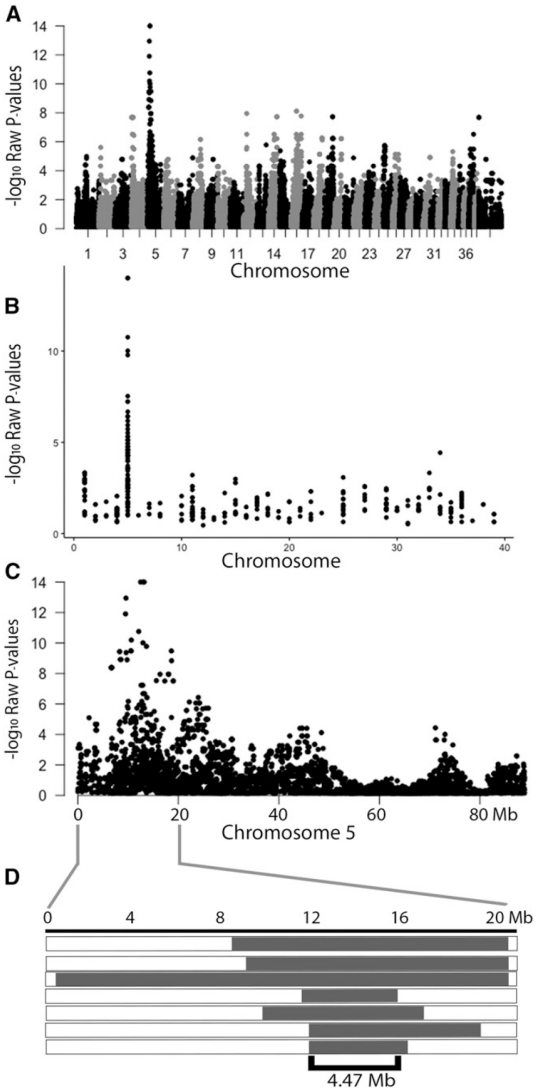

Figure 1 GWAS for Rottweiler NAD. A) Manhattan plot for the NAD GWAS showing the –log10 of the raw P values (Y axis) for each genotyped SNP by chromosome (X axis). Genomic inflation (λ) was 1.52. B) SNPs with an allele frequency of 100% in cases were plotted; with the –log10 of the raw P values (Y axis) for each SNP by chromosome (X axis). C) Plot of the –log10 of the raw P values (Y axis) for each SNP on canine chromosome 5 (CFA 5). D) Haplotypes observed in the seven cases, showing homozygosity throughout the associated region. Runs of homozygosity are marked by the gray horizontal bars. The critical interval is marked by the shared homozygous haplotype in between the black bracket (CFA5: 11.29 Mb – 15.75 Mb).

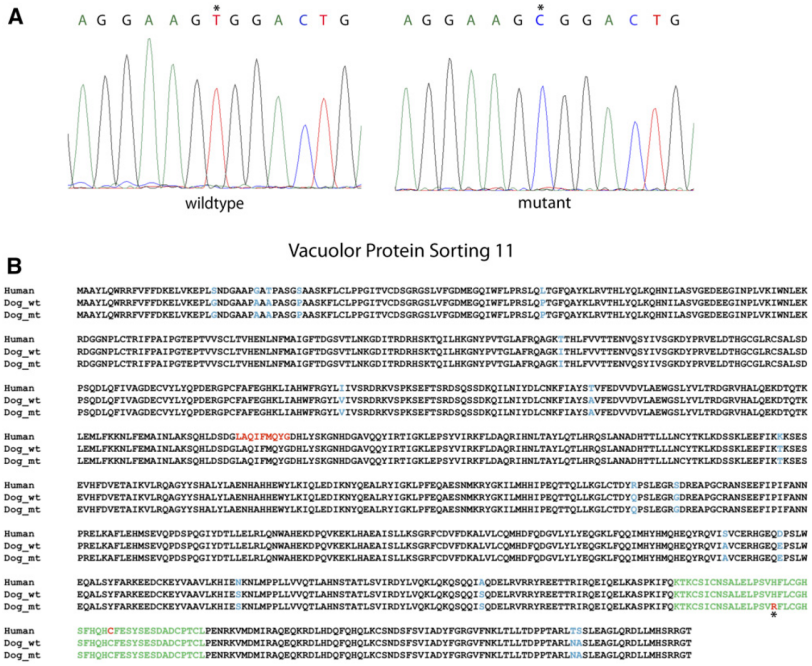

Figure 2 *VPS11* sequence electropherograms and amino acid alignment with human VPS11. A) Electropherogram of the missense mutation (*) (g.14777774T > C) within the cDNA of *VPS11*. B) Amino acid sequence alignment of human and dog (wild type and mutant) Vacuolar Protein Sorting 11 showing 98.2% amino acid conservation across species. The Zinc RING finger domain is in green with the location of the missense variant denoted by an asterisk below the aligned sequence (specific amino acid is highlighted in red). Non-conserved amino acids are in blue, and known disease causing mutations in human patients are in red.

Whole-Genome Sequencing

Variants within the critical interval, identified in the GWAS, were analyzed for association using 98 dogs unaffected with NAD and two Rottweilers histopathologically confirmed to have NAD. Within the critical interval, there were 73 genes, and 31,749 SNPs and 17,421 indels were identified. 164 SNP variants and 15 indels segregated with the phenotype were within the region identified on CFA 5. Only a single SNP was found to be protein coding (CFA 5:14,777,774) among the segregating variants. The remaining variants were 3′ UTR (n = 4), downstream (n = 23), intergenic (n = 74), intronic (n = 56), a non-coding transcript (n = 5), or upstream variants (n = 16).

VPS11 Non-Synonymous Variant

A non-synonymous variant was identified on CFA 5 (g.14777774T > C; CanFam3.1) in the Vacuolar Protein Sorting 11 (*VPS11*) gene. The variant leads to an amino acid change (p.H835R) in the Zinc RING finger domain of the protein (Figure 2B), which is ultimately predicted to be deleterious (VEP: moderate; PolyPhen-2: 0.999; SIFT: 0). The cDNA of *VPS11* was sequenced from liver from two NAD affected Rottweilers, one Gordon Setter, and one mixed breed dog, (Figure 2.A) to confirm the presence of the mutation in the mRNA (*VPS11*c.2504A > G). *VPS11* is highly conserved across species (Table S1), with humans and dogs having 98.2% conservation at the amino acid level (Figure 2.B).

VPS11 Variant Genotyping

440 dogs, consisting of 288 Rottweilers, and 152 dogs from 19 other breeds, were genotyped for the *VPS11* mutation (Table 1). Of the 288 Rottweilers, 13 were identified as heterozygous for the mutation, seven were homozygous for the mutation (cases as described above), and the remaining 268 Rottweilers, along with the 152 other dogs, were all homozygous for the reference allele. Of the Rottweilers genotyped, 211 were from the United States of America (204 wild type, three heterozygous, and seven homozygous mutants); 75 were from Europe (65 wild type and 10 heterozygous). The frequency of the mutant allele in this population is estimated to be 2.3%.

DISCUSSION

Seven Rottweilers that presented with clinical signs consistent with the NAD phenotype were homozygous for a non-synonymous mutation within the RING-finger domain of the *Vacuolar Protein Sorting 11 (VPS11)* gene (Figure 2). In order to overcome significant population stratification based on relatedness of the affected cases, a genome-wide association followed by homozygosity mapping was used. This type of approach has been used successfully in the past to identify breed specific Mendelian recessive diseases in dogs (Drögemüller *et al.* 2009; Kropatsch *et al.* 2010; Forman *et al.* 2016).

Sorting and degradation of internalized cell surface proteins and lipids in eukaryotic cells is controlled through the "endocytic network" (Balderhaar and Ungermann 2013; Spang 2016), such that surface proteins may progress through early and late endosomes before they are degraded in lysosomes, or may be sorted and recycled. Disposal and recycling of cytoplasmic components is similarly achieved through the autophagosome-lysosome pathway during autophagy (Levine and Klionsky 2004; Nixon 2013) (Figure 3).

Two VPS class C complexes, CORVET and HOPS, each composed of multiple different VPS proteins are essential for control of the membrane fusion machinery and trafficking of material through these endosome-lysosome organelles. CORVET and HOPS act as tethers, in coordination with other key proteins such as RAB5 and RAB7, and bring

Table 1 *VPS11* (g.14777774T > C) MUTATION GENOTYPING RESULTS

BREED	TOTAL	*VPS11/ VPS11*	*VPS11/ vps11*	*vps11/ vps11*
Boston Terrier	**2**	2	0	0
Boxer	**33**	33	0	0
Brittany	**3**	3	0	0
Bulldog	**5**	5	0	0
Dachshund	**1**	1	0	0
French Bulldog	**3**	3	0	0
German Shorthaired Pointer	**4**	4	0	0
Golden Retriever	**39**	39	0	0
Great Dane	**2**	2	0	0
Irish Setter	**1**	1	0	0
Labrador Retriever	**17**	17	0	0
Mixed Breeds	**6**	6	0	0
Newfoundland	**4**	4	0	0
Nova Scotia Duck Tolling Retriever	**13**	13	0	0
Pug	**5**	5	0	0
Rottweiler	**288**	268	13	7
Saluki	**2**	2	0	0
Weimaraner	**5**	5	0	0
West Highland White Terrier	**2**	2	0	0
Whippet	**5**	5	0	0

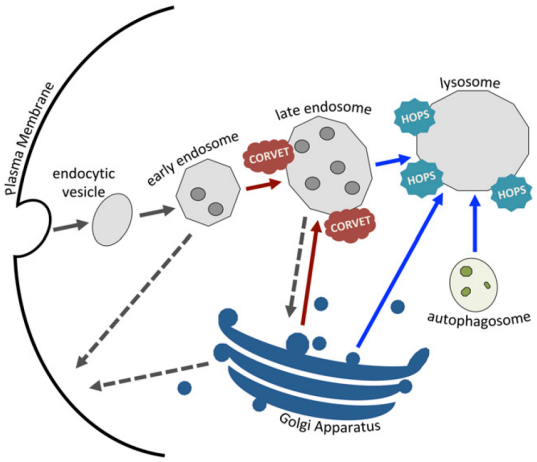

Figure 3 Schematic representation of the endosome-autophagosome-lysosome pathway. VPS11 is a key constituent of the VPS class C complexes CORVET (red) and HOPS (blue). Disruption of the CORVET/HOPS tethering complexes, and subsequently the membrane fusion processes required for appropriate trafficking, would be consistent with both the lysosomal storage and NAD phenotypes seen in human and dog disease, secondary to accumulation of membrane and cytosolic constituents. Blue arrows represent fusion events mediated by HOPS, red arrows represent fusion events mediated by CORVET, dashed gray arrows represent pathways of exocytosis and plain gray lines represent pathways of endocytosis.

appropriately targeted vesicles into close proximity with the target membrane (Richardson *et al.* 2004; Balderhaar and Ungermann 2013; Perini *et al.* 2014; van der Kant *et al.* 2015; Spang 2016). Both CORVET (class C core vacuole/endosome tethering complex) and HOPS (homotypic fusion and protein transport) contain a four-subunit core consisting of VPS11 (PEP5), VPS16, VPS18 (PEP3), and VPS33, which are conserved across yeast, insects, plants, and mammals (Nickerson *et al.* 2009). VPS11 has been shown to have a key role in determining selective binding to either early or late endosomes, and as an integrator of the complex assembly (Plemel *et al.* 2011; van der Kant *et al.* 2015). The RING domain of VPS11 that harbors the non-synonymous mutation in Rottweiler NAD has been shown to be important specifically in fusion at the vacuole (lysosome) in yeast (Plemel *et al.* 2011).

Mutation of the *VPS11* gene in humans is associated with an infantile onset neurological syndrome characterized by hypomyelination and variable neurological deficits including motor and cognitive impairment, dystonia, ataxia, visual deficits, and seizures (Figure 2.B) (Edvardson *et al.* 2015; Hörtnagel *et al.* 2016; Zhang *et al.* 2016). Histopathological characterization has not been done. However, the syndrome is classified as a leukoencephalopathy based on MRI (magnetic resonance imaging) findings, and skin and bone marrow biopsies were suggestive of a lysosomal storage type disease (Hörtnagel *et al.* 2016). Consistent with the known function of VSP11 (Plemel *et al.* 2011), *in vitro* studies of the mutant human protein resulted in disruption of late endosome/vacuole fusion and the autophagic pathway (Edvardson *et al.* 2015; Zhang *et al.* 2016). Although the Rottweiler *VPS11* mutation is in a similar location to one of the documented human mutations within the VPS11 RING finger domain (Figure 2.B) (Edvardson *et al.* 2015; Zhang *et al.* 2016), the clinical phenotypes appear to have distinct differences, most notably the apparent white matter *vs.* gray matter distribution of lesions in humans *vs.* dogs. The human *VPS11* syndrome also appears to be characterized by lysosomal accumulations compared to axonal spheroids (Hörtnagel *et al.* 2016), although detailed histopathological characterization of human CNS lesions is not available. However, this spectrum of intracellular accumulations of varying types is within the rational consequences of disruption of the endosome-autophagosome-lysosome system predicted following VPS11 (CORVET/HOPS) dysfunction. Additionally, species and site specific differences in pathological phenotype for mutations within the same gene are well documented across a broad range of genetic diseases. For example, a variety of clinical and pathological phenotypes have been reported for alterations within the same genes that cause some of the neuroaxonal dystrophy syndromes in humans (Revesz *et al.* 2015; Arber *et al.* 2016); *PLA2G6* gene mutations can give a spectrum of disease phenotypes as well as classical INAD including dystonia-parkinsonism syndromes and spastic paraplegia (Gregory *et al.* 2008; Ozes *et al.* 2017). Similarly, alterations in the *MFN2* gene give rise to fetal onset NAD in dogs (Fyfe *et al.* 2010; Fyfe *et al.* 2011); however, human alterations result in peripheral nervous system syndromes (Charcot-Marie-Tooth disease type 2A2 and hereditary motor and sensory neuropathy type 6A (Del Bo *et al.* 2008)), while mouse knockouts and transgenic overexpression and cattle with altered *MFN2* have CNS and PNS neurodegeneration but do not have NAD (Drögemüller *et al.* 2011). Protein site-specific effects, species differences in pathological responses (such as lack of iron accumulation in canine NAD), and differences in gain or loss of function mutations may all contribute to the phenotypic heterogeneity.

Despite this heterogeneity within the NAD disease phenotype, common pathological pathways are implicated in many NAD syndromes, including the primary and secondary diseases. Vitamin E deficiency has been associated to varying degrees with axonal dystrophy in both experimental and clinical settings, in several species including

dogs, rodents, horses and primates (Nelson *et al.* 1981; Pillai *et al.* 1994; McLellan *et al.* 2003; Finno *et al.* 2013; Finno *et al.* 2016). The importance of the complex relationship between pathways controlling reactive oxygen species (ROS) and autophagy has been well documented (Underwood *et al.* 2010; Fang *et al.* 2017), and the autophagy pathway is particularly important in the context of the highly metabolic neuron (Nixon 2013). Previously defined genes associated with NAD in dogs (*PLA2G6*, *TECPR2* and *MFN2*) as well as many human NAD related genes have been proposed as potential modulators of the autophagy pathway (Fyfe *et al.* 2011; Hahn *et al.* 2015; Meyer *et al.* 2015; Arber *et al.* 2016; Tsuboi *et al.* 2017), and the currently described *VPS11* gene alteration in Rottweiler NAD would be predicted to affect the autophagic, as well as other lysosomal pathways. High conservation of *VPS11* between species, the essential role *VPS11* plays in the endosomal-autophagy-lysosomal pathways, and the impact of mutations in *VPS11* leading to neurodegenerative diseases, provides strong support for the missense mutation identified in Rottweiler NAD to be causative for the disease and a potential candidate for unexplained forms of human NAD. Detailed biological analysis of the Rottweiler *VPS11* specific mutation will be needed to fully understand the apparent species/mutation differences in disease expression and its potential value as a translational model.

ACKNOWLEDGMENTS

It is with extreme gratitude that we acknowledge Dr. Linda Cork. If it were not for her maintaining DNA samples from the four original Rottweiler NAD cases, from 1983, this study would not have been possible.

LITERATURE CITED

Adzhubei, I. A., S. Schmidt, L. Peshkin, V. E. Ramensky, A. Gerasimova *et al.*, 2010 A method and server for predicting damaging missense mutations. Nat. Methods 7: 248–249. https://doi.org/10.1038/nmeth0410-248

Adzhubei, I., D. M. Jordan, and S. R. Sunyaev, 2013 Predicting functional effect of human missense mutations using PolyPhen-2. Curr. Protoc. Hum. Genet. Chapter 7: Unit7.20. https://doi.org/10.1002/0471142905.hg0720s76

Aleman, M., C. J. Finno, R. J. Higgins, B. Puschner, B. Gericota *et al.*, 2011 Evaluation of epidemiological, clinical, and pathological features of neuroaxonal dystrophy in Quarter Horses. J. Am. Vet. Med. Assoc. 239: 823–833. https://doi.org/10.2460/javma.239.6.823

Arber, C. E., A. Li, H. Houlden, and S. Wray, 2016 Review: Insights into molecular mechanisms of disease in neurodegeneration with brain iron accumulation: unifying theories. Neuropathol. Appl. Neurobiol. 42: 220–241. https://doi.org/10.1111/nan.12242

Balderhaar, H. J., and C. Ungermann, 2013 CORVET and HOPS tethering complexes - coordinators of endosome and lysosome fusion. J. Cell Sci. 126: 1307–1316. https://doi.org/10.1242/jcs.107805

Barrett, J. C., B. Fry, J. Maller, and M. J. Daly, 2005 Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265. https://doi.org/10.1093/bioinformatics/bth457

Barrett, J. C., 2009 Haploview: Visualization and analysis of SNP genotype data. *Cold Spring Harb Protoc* 2009 (10): pdb ip71. https://doi.org/10.1101/pdb.ip71

Barrows, M., R. Killick, C. Day, R. Saunders, K. Baiker *et al.*, 2017 Neuroaxonal Dystrophy in a Flock of Pied Imperial Pigeons (Ducula bicolor). J. Comp. Pathol. 156: 451–457. https://doi.org/10.1016/j.jcpa.2017.02.006

Beech, J., 1984 Neuroaxonal dystrophy of the accessory cuneate nucleus in horses. Vet. Pathol. 21: 384–393. https://doi.org/10.1177/030098588402100404

Bennett, P. F., and R. E. Clarke, 1997 Laryngeal paralysis in a rottweiler with neuroaxonal dystrophy. Aust. Vet. J. 75: 784–786. https://doi.org/10.1111/j.1751-0813.1997.tb15650.x

Blakemore, W. F., and A. C. Palmer, 1985 Nervous disease in the chihuahua characterised by axonal swellings. Vet. Rec. 117: 498–499. https://doi.org/10.1136/vr.117.19.498

Blythe, L. L., B. D. Hultgren, A. M. Craig, L. H. Appell, E. D. Lassen *et al.*, 1991 Clinical, viral, and genetic evaluation of equine degenerative myeloencephalopathy in a family of Appaloosas. J. Am. Vet. Med. Assoc. 198: 1005–1013.

Borràs, D., I. Ferrer, and M. Pumarola, 1999 Age-related changes in the brain of the dog. Vet. Pathol. 36: 202–211. https://doi.org/10.1354/vp.36-3-202

Bouley, D. M., J. J. McIntire, B. T. Harris, R. J. Tolwani, G. M. Otto *et al.*, 2006 Spontaneous murine neuroaxonal dystrophy: a model of infantile neuroaxonal dystrophy. J. Comp. Pathol. 134: 161–170. https://doi.org/10.1016/j.jcpa.2005.10.002

Brinkhof, B., B. Spee, J. Rothuizen, and L. C. Penning, 2006 Development and evaluation of canine reference genes for accurate quantification of gene expression. Anal. Biochem. 356: 36–43. https://doi.org/10.1016/j.ab.2006.06.001

Brown, E. A., P. J. Dickinson, T. Mansour, B. K. Sturges, M. Aguilar *et al.*, 2017 FGF4 retrogene on CFA12 is responsible for chondrodystrophy and intervertebral disc disease in dogs. Proc. Natl. Acad. Sci. USA 114: 11476–11481. https://doi.org/10.1073/pnas.1709082114

Cajal, S. R. Y., 1928 *Degeneration and regeneration of the nervous system*, Oxford University Press, London.

Carmichael, K. P., E. W. Howerth, J. E. Oliver, Jr., and K. Klappenbach, 1993 Neuroaxonal dystrophy in a group of related cats. J. Vet. Diagn. Invest. 5: 585–590. https://doi.org/10.1177/104063879300500414

Chrisman, C. L., L. C. Cork, and D. A. Gamble, 1984 Neuroaxonal dystrophy of Rottweiler dogs. J. Am. Vet. Med. Assoc. 184: 464–467.

Clark, R. G., W. J. Hartley, G. S. Burgess, J. S. Cameron, and G. Mitchell, 1982 Suspected inherited cerebellar neuroaxonal dystrophy in collie sheep dogs. N. Z. Vet. J. 30: 102–103. https://doi.org/10.1080/00480169.1982.34897

Clarke, G. M., C. A. Anderson, F. H. Pettersson, L. R. Cardon, A. P. Morris *et al.*, 2011 Basic statistical analysis in genetic case-control studies. Nat. Protoc. 6: 121–133. https://doi.org/10.1038/nprot.2010.182

Cordy, D. R., W. P. Richards, and G. E. Bradford, 1967 Systemic neuroaxonal dystrophy in Suffolk sheep. Acta Neuropathol. 8: 133–140. https://doi.org/10.1007/BF00687690

Cork, L. C., J. C. Troncoso, D. L. Price, E. F. Stanley, and J. W. Griffin, 1983 Canine neuroaxonal dystrophy. J. Neuropathol. Exp. Neurol. 42: 286–296. https://doi.org/10.1097/00005072-198305000-00006

Degl'Innocenti, S., N. Asiag, O. Zeira, C. Falzone, and C. Cantile, 2017 Neuroaxonal Dystrophy and Cavitating Leukoencephalopathy of Chihuahua Dogs. Vet. Pathol. 54: 832–837. https://doi.org/10.1177/0300985817712557

Del Bo, R., M. Moggio, M. Rango, S. Bonato, M. G. D'Angelo *et al.*, 2008 Mutated mitofusin 2 presents with intrafamilial variability and brain mitochondrial dysfunction. Neurology 71: 1959–1966. https://doi.org/10.1212/01.wnl.0000327095.32005.a4

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. https://doi.org/10.1371/journal.pgen.1000579

Drögemüller, C., U. Reichart, T. Seuberlich, A. Oevermann, M. Baumgartner *et al.*, 2011 An unusual splice defect in the mitofusin 2 gene (MFN2) is associated with degenerative axonopathy in Tyrolean Grey cattle. PLoS One 6: e18931. https://doi.org/10.1371/journal.pone.0018931

Edvardson, S., F. Gerhard, C. Jalas, J. Lachmann, D. Golan *et al.*, 2015 Hypomyelination and developmental delay associated with VPS11 mutation in Ashkenazi-Jewish patients. J. Med. Genet. 52: 749–753. https://doi.org/10.1136/jmedgenet-2015-103239

Evans, M. G., T. P. Mullaney, and C. T. Lowrie, 1988 Neuroaxonal dystrophy in a rottweiler pup. J. Am. Vet. Med. Assoc. 192: 1560–1562.

Fang, C., L. Gu, D. Smerin, S. Mao, and X. Xiong, 2017 The Interrelation between Reactive Oxygen Species and Autophagy in Neurological

Disorders. Oxid. Med. Cell. Longev. 2017: 8495160. https://doi.org/10.1155/2017/8495160

Finnie, J. W., I. V. Jerrett, J. Manavis, and J. Cave, 2014 Neuroaxonal dystrophy in Merino-Border Leicester x Polled Dorset lambs. Aust. Vet. J. 92: 389–391. https://doi.org/10.1111/avj.12222

Finno, C. J., T. Famula, M. Aleman, R. J. Higgins, J. E. Madigan *et al.*, 2013 Pedigree analysis and exclusion of alpha-tocopherol transfer protein (TTPA) as a candidate gene for neuroaxonal dystrophy in the American Quarter Horse. J. Vet. Intern. Med. 27: 177–185. https://doi.org/10.1111/jvim.12015

Finno, C. J., M. H. Bordbari, S. J. Valberg, D. Lee, J. Herron *et al.*, 2016 Transcriptome profiling of equine vitamin E deficient neuroaxonal dystrophy identifies upregulation of liver X receptor target genes. Free Radic. Biol. Med. 101: 261–271. https://doi.org/10.1016/j.freeradbiomed.2016.10.009

Forman, O. P., R. J. Hitti, L. Pettitt, C. A. Jenkins, D. P. O'Brien *et al.*, 2016 An Inversion Disrupting FAM134B Is Associated with Sensory Neuropathy in the Border Collie Dog Breed. G3 (Bethesda) 6: 2687–2692. https://doi.org/10.1534/g3.116.027896

Franklin, R. J., N. D. Jeffery, and I. K. Ramsey, 1995 Neuroaxonal dystrophy in a litter of papillon pups. J. Small Anim. Pract. 36: 441–444. https://doi.org/10.1111/j.1748-5827.1995.tb02774.x

Fyfe, J. C., R. A. Al-Tamimi, R. J. Castellani, D. Rosenstein, D. Goldowitz *et al.*, 2010 Inherited neuroaxonal dystrophy in dogs causing lethal, fetal-onset motor system dysfunction and cerebellar hypoplasia. J. Comp. Neurol. 518: 3771–3784. https://doi.org/10.1002/cne.22423

Fyfe, J. C., R. A. Al-Tamimi, J. Liu, A. A. Schaffer, R. Agarwala *et al.*, 2011 A novel mitofusin 2 mutation causes canine fetal-onset neuroaxonal dystrophy. Neurogenetics 12: 223–232. https://doi.org/10.1007/s10048-011-0285-6

Gavier-Widen, D., G. A. Wells, M. M. Simmons, J. W. Wilesmith, and J. Ryan, 2001 Histological observations on the brains of symptomless 7-year-old cattle. J. Comp. Pathol. 124: 52–59. https://doi.org/10.1053/jcpa.2000.0428

Gregory, A., S. K. Westaway, I. E. Holm, P. T. Kotzbauer, P. Hogarth *et al.*, 2008 Neurodegeneration associated with genetic defects in phospholipase A(2). Neurology 71: 1402–1409. https://doi.org/10.1212/01.wnl.0000327094.67726.28

Hahn, K., C. Rohdin, V. Jagannathan, P. Wohlsein, W. Baumgartner *et al.*, 2015 TECPR2 Associated Neuroaxonal Dystrophy in Spanish Water Dogs. PLoS One 10: e0141824. https://doi.org/10.1371/journal.pone.0141824

Hanshaw, D. M., J. W. Finnie, J. Manavis, and A. E. Kessell, 2015 Axonal Spheroid Accumulation In the Brainstem and Spinal Cord of A Young Angus Cow with Ataxia. Aust. Vet. J. 93: 283–286. https://doi.org/10.1111/avj.12346

Harper, P. A., and A. G. Morton, 1991 Neuroaxonal dystrophy in Merino sheep. Aust. Vet. J. 68: 152–153. https://doi.org/10.1111/j.1751-0813.1991.tb03162.x

Hörtnagel, K., I. Krägeloh-Mann, A. Bornemann, M. Döcker, S. Biskup *et al.*, 2016 The second report of a new hypomyelinating disease due to a defect in the VPS11 gene discloses a massive lysosomal involvement. J. Inherit. Metab. Dis. 39: 849–857. https://doi.org/10.1007/s10545-016-9961-x

Hytönen, M. K., and H. Lohi, 2016 Canine models of human rare disorders. Rare Dis. 4: e1241362. https://doi.org/10.1080/21675511.2016.1241362

Kessell, A. E., J. W. Finnie, P. C. Blumbergs, J. Manavis, and I. V. Jerrett, 2012 Neuroaxonal dystrophy in Australian Merino lambs. J. Comp. Pathol. 147: 62–72. https://doi.org/10.1016/j.jcpa.2011.09.006

Kropatsch, R., E. Petrasch-Parwez, D. Seelow, A. Schlichting, W. M. Gerding *et al.*, 2010 Generalized progressive retinal atrophy in the Irish Glen of Imaal Terrier is associated with a deletion in the ADAM9 gene. Mol. Cell. Probes 24: 357–363. https://doi.org/10.1016/j.mcp.2010.07.007

Levine, B., and D. J. Klionsky, 2004 Development by self-digestion: molecular mechanisms and biological functions of autophagy. Dev. Cell 6: 463–477. https://doi.org/10.1016/S1534-5807(04)00099-1

Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe *et al.*, 2005 Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438: 803–819. https://doi.org/10.1038/nature04338

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. https://doi.org/10.1101/gr.107524.110

McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie *et al.*, 2016 The Ensembl Variant Effect Predictor. Genome Biol. 17: 122. https://doi.org/10.1186/s13059-016-0974-4

McLellan, G. J., R. Cappello, I. G. Mayhew, R. Elks, P. Lybaert *et al.*, 2003 Clinical and pathological observations in English cocker spaniels with primary metabolic vitamin E deficiency and retinal pigment epithelial dystrophy. Vet. Rec. 153: 287–292. https://doi.org/10.1136/vr.153.10.287

Meyer, E., M. A. Kurian, and S. J. Hayflick, 2015 Neurodegeneration with Brain Iron Accumulation: Genetic Diversity and Pathophysiological Mechanisms. Annu. Rev. Genomics Hum. Genet. 16: 257–279. https://doi.org/10.1146/annurev-genom-090314-025011

Mhlanga-Mutangadura, T., G. S. Johnson, A. Ashwini, G. D. Shelton, S. A. Wennogle *et al.*, 2016 A Homozygous *RAB3GAP1:c.743delC* mutation in Rottweilers with Neuronal Vacuolation and Spinocerebellar Degeneration. J. Vet. Intern. Med. 30: 813–818. https://doi.org/10.1111/jvim.13921

Minor, K. M., A. Letko, D. Becker, M. Drögemüller, P. J. J. Mandigers *et al.*, 2018 Canine NAPEPLD-associated models of human myelin disorders. Sci. Rep. 8: 5818. https://doi.org/10.1038/s41598-018-23938-7

Nelson, J. S., C. D. Fitch, V. W. Fischer, G. O. Broun, Jr., and A. C. Chou, 1981 Progressive neuropathologic lesions in vitamin E-deficient rhesus monkeys. J. Neuropathol. Exp. Neurol. 40: 166–186. https://doi.org/10.1097/00005072-198103000-00008

Nibe, K., C. Kita, M. Morozumi, Y. Awamura, S. Tamura *et al.*, 2007 Clinicopathological features of canine neuroaxonal dystrophy and cerebellar cortical abiotrophy in Papillon and Papillon-related dogs. *The Journal of Veterinary Medical Science / The Japanese Society of Veterinary Science* 69 (10): 1047–1052.

Nickerson, D. P., C. L. Brett, and A. J. Merz, 2009 Vps-C complexes: gatekeepers of endolysosomal traffic. Curr. Opin. Cell Biol. 21: 543–551. https://doi.org/10.1016/j.ceb.2009.05.007

Nixon, R. A., 2013 The role of autophagy in neurodegenerative disease. Nat. Med. 19: 983–997. https://doi.org/10.1038/nm.3232

Nuttall, W. O., 1988 Ovine neuroaxonal dystrophy in New Zealand. N. Z. Vet. J. 36: 5–7. https://doi.org/10.1080/00480169.1988.35462

Ozes, B., N. Karagoz, R. Schule, A. Rebelo, M. J. Sobrido *et al.*, 2017 PLA2G6 mutations associated with a continuous clinical spectrum from neuroaxonal dystrophy to hereditary spastic paraplegia. Clin. Genet. 92: 534–539. https://doi.org/10.1111/cge.13008

Perini, E. D., R. Schaefer, M. Stoter, Y. Kalaidzidis, and M. Zerial, 2014 Mammalian CORVET is required for fusion and conversion of distinct early endosome subpopulations. Traffic 15: 1366–1389. https://doi.org/10.1111/tra.12232

Pillai, S. R., M. G. Traber, H. J. Kayden, N. R. Cox, M. Toivio-Kinnucan *et al.*, 1994 Concomitant brainstem axonal dystrophy and necrotizing myopathy in vitamin E-deficient rats. J. Neurol. Sci. 123: 64–73. https://doi.org/10.1016/0022-510X(94)90205-4

Pintus, D., M. G. Cancedda, S. Macciocu, C. Contu, and C. Ligios, 2016 Pathological findings in a Dachshund-cross dog with neuroaxonal dystrophy. Acta Vet. Scand. 58: 37. https://doi.org/10.1186/s13028-016-0218-3

Plemel, R. L., B. T. Lobingier, C. L. Brett, C. G. Angers, D. P. Nickerson *et al.*, 2011 Subunit organization and Rab interactions of Vps-C protein complexes that control endolysosomal membrane traffic. Mol. Biol. Cell 22: 1353–1363. https://doi.org/10.1091/mbc.e10-03-0260

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. https://doi.org/10.1086/519795

Résibois, A., and L. Poncelet, 2004 Purkinje cell neuroaxonal dystrophy similar to nervous mutant mice phenotype in two sibling kittens. Acta Neuropathol. 107: 553–558. https://doi.org/10.1007/s00401-004-0846-y

Revesz, T., H. B. Clark, J. L. Holton, H. H. Houlden, P. G. Ince et al., 2015 Extrapyramidal diseases of movement: Neuroaxonal dystrophy, disorders of metal biometabolism and related disorders, pp. 778–786 in *Greenfield's Neuropathology*, edited by Love, S., H. Budka, J. W. Ironside, and A. Perry. CRC Press, Boca Raton, FL.

Richardson, S. C., S. C. Winistorfer, V. Poupon, J. P. Luzio, and R. C. Piper, 2004 Mammalian late vacuole protein sorting orthologues participate in early endosomal fusion and interact with the cytoskeleton. Mol. Biol. Cell 15: 1197–1210. https://doi.org/10.1091/mbc.e03-06-0358

Rodriguez, F., A. Espinosa de los Monteros, M. Morales, P. Herraez, J. R. Jaber *et al.*, 1996 Neuroaxonal dystrophy in two siamese kitten littermates. Vet. Rec. 138: 548–549. https://doi.org/10.1136/vr.138.22.548

Rozen, S., and H. Skaletsky, 2000 Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. 132: 365–386.

Sacre, B. J., J. F. Cummings, and A. De Lahunta, 1993 Neuroaxonal dystrophy in a Jack Russell terrier pup resembling human infantile neuroaxonal dystrophy. Cornell Vet. 83: 133–142.

Saito, K., 1980 Spheroids and altered axons in the spinal gray matter of the normal cat. An electron-microscopic study. Acta Neuropathol. 52: 213–222. https://doi.org/10.1007/BF00705809

Shearin, A. L., and E. A. Ostrander, 2010 Leading the way: canine models of genomics and disease. Dis. Model. Mech. 3: 27–34. https://doi.org/10.1242/dmm.004358

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus *et al.*, 2011 Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7: 539. https://doi.org/10.1038/msb.2011.75

Sim, N. L., P. Kumar, J. Hu, S. Henikoff, G. Schneider *et al.*, 2012 SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* 40 (Web Server issue): W452–457. https://doi.org/10.1093/nar/gks539

Siso, S., I. Ferrer, and M. Pumarola, 2001 Juvenile neuroaxonal dystrophy in a Rottweiler: accumulation of synaptic proteins in dystrophic axons. Acta Neuropathol. 102: 501–504.

Spang, A., 2016 Membrane Tethering Complexes in the Endosomal System. Front. Cell Dev. Biol. 4: 35. https://doi.org/10.3389/fcell.2016.00035

Summers, B. A., J. F. Cummings, and A. de Lahunta, 1995 Principles of neuropathology: Neurons, pp. 4–10 in *Veterinary Neuropathology*, edited by Summers, B. A., J. F. Cummings, and A. de Lahunta. Mosby, St. Louis.

Suzuki, Y., K. Ohta, and S. Suu, 1979 Correlative studies of axonal spheroids and Lafora-like bodies in aged dogs. Acta Neuropathol. 48: 77–81. https://doi.org/10.1007/BF00691796

Tsuboi, M., M. Watanabe, K. Nibe, N. Yoshimi, A. Kato *et al.*, 2017 Identification of the PLA2G6 c.1579G>A Missense Mutation in Papillon Dog Neuroaxonal Dystrophy Using Whole Exome Sequencing Analysis. PLoS One 12: e0169002. https://doi.org/10.1371/journal.pone.0169002

Underwood, B. R., S. Imarisio, A. Fleming, C. Rose, G. Krishna *et al.*, 2010 Antioxidants can inhibit basal autophagy and enhance neurodegeneration in models of polyglutamine disease. Hum. Mol. Genet. 19: 3413–3429. https://doi.org/10.1093/hmg/ddq253

van der Kant, R., C. T. Jonker, R. H. Wijdeven, J. Bakker, L. Janssen *et al.*, 2015 Characterization of the Mammalian CORVET and HOPS Complexes and Their Modular Restructuring for Endosome Specificity. J. Biol. Chem. 290: 30280–30290. https://doi.org/10.1074/jbc.M115.688440

Wickham, H., 2009 *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Woodard, J. C., G. H. Collins, and J. R. Hessler, 1974 Feline hereditary neuroaxonal dystrophy. Am. J. Pathol. 74: 551–566.

Yagishita, S., 1978 Morphological investigations on axonal swellings and spheroids in various human diseases. Virchows Arch. A Pathol. Anat. Histol. 378: 181–197. https://doi.org/10.1007/BF00427359

Zhang, J., V. Lachance, A. Schaffner, X. Li, A. Fedick *et al.*, 2016 A Founder Mutation in VPS11 Causes an Autosomal Recessive Leukoencephalopathy Linked to Autophagic Defects. PLoS Genet. 12: e1005848. https://doi.org/10.1371/journal.pgen.1005848

Communicating editor: B. Andrews

A missense variant in *SCN8A* in Alpine dachsbracke dogs affected by spinocerebellar ataxia

Journal:	Genes
Manuscript status:	published
Contributions:	Investigation, Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing

genes

MDPI

Article

A Missense Variant in *SCN8A* in Alpine Dachsbracke Dogs Affected by Spinocerebellar Ataxia

Anna Letko [1], Elisabeth Dietschi [1], Marco Nieburg [2], Vidhya Jagannathan [1], Corinne Gurtner [3], Anna Oevermann [4] and Cord Drögemüller [1,*]

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; anna.letko@vetsuisse.unibe.ch (A.L.); elisabeth.dietschi@vetsuisse.unibe.ch (E.D.); vidhya.jagannathan@vetsuisse.unibe.ch (V.J.)
2 Tierarztpraxis Nieburg, 18546 Sassnitz, Germany; marco@vetmed-berlin.de
3 Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; corinne.gurtner@vetsuisse.unibe.ch
4 Division of Neurological Sciences, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; anna.oevermann@vetsuisse.unibe.ch
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch; Tel.: +41-316-312-529

Received: 12 April 2019; Accepted: 6 May 2019; Published: 10 May 2019

Abstract: Spinocerebellar ataxias is an umbrella term for clinically- and neuropathologically-heterogeneous early-onset hereditary neurodegenerative diseases affecting several dog breeds. The purpose of this study is to identify the causative genetic variant associated with ataxia, tremor, and loss of balance in Alpine Dachsbracke dogs. We investigated two related litters in which four cases were reported. Neuropathology of two dogs revealed spongy degeneration associated with axonal degeneration. Combined genetic linkage and autozygosity analyses in four cases and eight related controls showed one critical disease-associated interval on chromosomes 27. Private whole-genome sequence variants of one ataxia case against 600 unrelated controls revealed one protein-changing variant within the critical interval in the *SCN8A* gene (c.4898G>T; p.Gly1633Val). Perfect segregation with the phenotype was confirmed by genotyping >200 Alpine Dachsbracke dogs. *SCN8A* encodes a voltage-gated sodium channel and the missense variant was predicted deleterious by three different in silico prediction tools. Pathogenic variants in *SCN8A* were previously reported in humans with ataxia, pancerebellar atrophy, and cognitive disability. Furthermore, cerebellar ataxia syndrome in the 'jolting' mutant mice is caused by a missense variant in *Scn8a*. Therefore, we considered the *SCN8A*:c.4898G>T variant to be the most likely cause for recessively inherited spinocerebellar ataxia in Alpine Dachsbracke dogs.

Keywords: *Canis familiaris*; sodium channel; central nervous system; ataxia; Alpine Dachsbracke

1. Introduction

Hereditary ataxias in humans are described as a clinically- and genetically-heterogeneous group of neurodegenerative diseases usually associated with cerebellar degeneration. They are phenotypically characterized by features such as gait abnormalities, imbalance, and associated movement abnormalities including incoordination of eye and hand movements, visual loss, seizures, behavioral symptoms, and peripheral neuropathy [1]. Different modes of inheritance have been described in cerebellar ataxias, with autosomal dominant (usually adult-onset) and autosomal recessive (typically juvenile-onset) being the most prevalent groups. Causative variants in more than 50 genes, which have been previously reported [2,3], reflect the large amount of clinical heterogeneity. Several pathophysiological mechanisms have been implicated including impaired ion channels function, failure of protein homeostasis, defects in the DNA repair system, or mitochondrial dysfunction [4].

Similar to humans, cerebellar ataxias in dogs are neurodegenerative disorders with variable age of onset, neuropathology, and disease severity. Cerebellar dysfunction is considered an important cause of movement disorders in purebred dogs [5]. However, to date, the number of described genes harboring the underlying genetic variants is significantly lower in dogs than in humans (*VLDLR*, *SPTBN2*, *SNX14*, *SEL1L*, *RAB24*, *KCNJ10*, *ITPR1*, *GRM1*, *CAPN1*, *ATP1B2*, *ATG4D*). These have been implicated in autosomal recessive ataxias of several dog breeds (OMIA 001947, 002092, 002034, 001692, 001913, 002089, 002097, 000078, 001820, 002110, 001954). Currently, canine neurodegenerative disease classification is based mainly on clinicopathological features, which due to lack of consensus may potentially lead to confusion and misdiagnoses. Therefore, it is important to improve the understanding of the underlying genetic and molecular mechanisms of inherited ataxias [5]. In this study, we aimed to identify the causative genetic variant associated with cerebellar ataxia, tremor, and loss of balance in Alpine Dachsbracke dogs.

2. Materials and Methods

2.1. Ethics Statement

All animal experiments were performed according to the local regulations. All animals in this study were examined with the consent of their owners. Sample collection was approved by the Cantonal Committee for Animal Experiments (Canton of Bern; permit 75/16).

2.2. DNA Samples and Single Nucleotide Variant Genotyping

A total of 216 blood samples from Alpine Dachsbracke dogs were collected. These included animals from two litters in which puppies affected by ataxia were observed (one dam, two sires, four cases, and eight healthy littermates). Genomic DNA was isolated from EDTA blood samples using the Maxwell RSC Whole Blood DNA kit (Promega, Dübendorf, Switzerland). DNA from the 12 selected dogs (4 affected and 8 closely related dogs, Supplementary Table S1) were genotyped using Illumina CanineHD BeadChip array (Illumina, San Diego, CA, USA) by GeneSeek (Neogen, Lincoln, NE, USA) for 220,853 single nucleotide variant (SNV) markers. Quality control filtering steps, as well as parentage control, were carried out using PLINK v1.9 [6]. Markers with call rate <90% were excluded, all individuals had call rates >90%. The pruned dataset consisted of 12 individuals and 213,288 markers.

2.3. Pathology

Brain from two affected puppies (Supplementary Table S1) and an eye from one affected puppy were collected and fixed in 4% buffered formaldehyde solution, embedded in paraffin, and sectioned at 4 µm. Sections were stained with hematoxylin and eosin (HE) and examined by light microscopy. Furthermore, selected sections of brains were stained with Bielschowsky silver stain.

2.4. Linkage and Autozygosity Analyses

Linkage and autozygosity analyses were performed in order to find critical disease-associated intervals. Parametric linkage analysis was carried out under a fully penetrant, recessive model of inheritance using the Merlin software [7] to test for co-segregation of any chromosomal regions and the ataxia phenotype. The dataset for linkage analysis included one complete family of seven dogs representing the first litter (sire, dam, two affected and three unaffected offspring). Five additional littermates from the second litter (two affected and three unaffected) were added to the analysis carried out with PLINK v1.9 [6] to determine intervals of extended homozygous regions with alleles shared by all four affected dogs as well as individual homozygous intervals in the control animals (Figure 1). The circos plot was created using OmicCircos package [8].

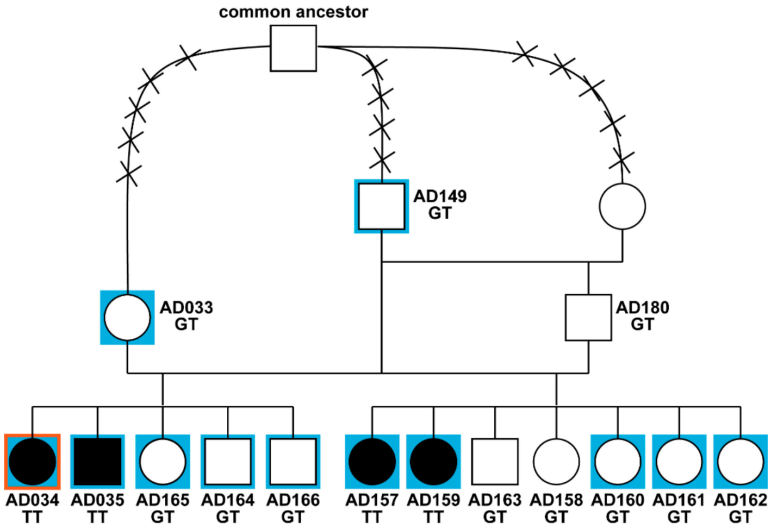

Figure 1. Pedigree of Alpine Dachsbracke dogs used for genetic mapping of spinocerebellar ataxia. Filled symbols represent spinocerebellar ataxia-affected dogs. The blue background indicates animals, which were genotyped on the single nucleotide variant (SNV) array, and the red contour indicates an affected dog selected for whole-genome resequencing. A common ancestor in both maternal and paternal lineages was identified. The crosses intersecting the connection lines represent the number of generations to the common ancestor. Note that the male 'AD149' sired the first litter and is also the grandsire of the second litter produced by his son 'AD180'.

2.5. Whole-Genome Resequencing

Whole-genome sequence (WGS) data at 23x coverage was obtained from a single affected Alpine Dachsbracke dog (Supplementary Table S1) after preparation of a PCR-free fragment library with a 450 bp insert size and collection of 206,013,034 paired-end reads (2 × 150 bp) using NovaSeq6000 Sequencing System (Illumina). The sequence data analysis and variant (SNVs and small indels) calling including the prediction of functional effects were performed as described before [9]. The dog reference genome assembly CanFam3.1 and NCBI annotation release 105 was used. The whole-genome sequence data from the Alpine Dachsbracke case was compared to the Boxer reference genome (CanFam3.1), 592 publically available control dogs of 124 various breeds, and 8 wolves (Supplementary Table S2). The Integrative genomics viewer (IGV) software [10] was used for visual inspection and screening for structural variants in the region of interest of the affected dog's WGS.

2.6. Targeted Genotyping

Polymerase chain reaction (PCR) and Sanger sequencing were used to validate and genotype the variants identified from whole-genome resequencing. PCR products from genomic DNA were amplified using AmpliTaq Gold 360 Master Mix (Applied Biosystems, Foster City, CA, USA) and the purified PCR amplicons were directly sequenced on an ABI3730 capillary sequencer (Applied Biosystems). The *SCN8A* missense variant (XM_022411522.1:c.4898G>T) was genotyped using the following primers: GGCCAATGTTGAACAGAGCA (forward) and ACTTAAGGGCTCCAGTGTCA (reverse). The sequence data were analyzed using Sequencher 5.1 software (GeneCodes).

2.7. Protein Predictions

PON-P2 [11], PROVEAN [12], and MutPred2 [13] were used to predict biological consequences of the discovered variant on protein. All references to the canine *SCN8A* gene correspond to the accessions NC_006609.3 (NCBI accession), XM_022411522.1 (mRNA), and XP_022267230.1 (protein). The canine

protein has the same length as the human SCN8A protein (NP_055006.1): 1980 amino acids from which 1966 (99.3%) are identical between dog and human. The Genome Aggregation Database (gnomAD) [14] was searched for corresponding variants in the human *SCN8A*.

2.8. Availability of Data and Material

The whole-genome data of an affected Alpine Dachsbracke dog has been made freely available (sample accession number SAMEA4848713) at the European Nucleotide Archive (ENA). All accession numbers of the control genomes are available in the Supplementary Table S2.

3. Results

3.1. Phenotype and Pedigree Analysis

We investigated two closely related litters of Alpine Dachsbracke dogs in which four puppies affected by ataxia were reported. All three parents of the affected dogs were phenotypically normal. A single female was sired with two related males (father and son) and pedigree analysis revealed a common ancestor within a maximum of five generations in both maternal and paternal lineages (Figure 1). Therefore, we assumed a monogenic autosomal recessive mode of inheritance.

Clinical signs of cerebellar dysfunction in the four puppies (one male, three females) were observed immediately when their normal littermates started to move in a coordinated fashion, so after approximately three weeks of age. The affected dogs exhibited ataxia, tremors, loss of balance, falling and other movement problems (Supplementary Video S1). Furthermore, the dog breeder reported that the vision of the affected dogs might be impaired. The severity of the clinical signs resulted in euthanasia of all cases by the age of 10–12 weeks.

Pedigree of Alpine Dachsbracke dogs used for genetic mapping of spinocerebellar ataxia.

3.2. Pathological Findings

Neuropathological analysis revealed mild and scattered white and grey matter vacuolization indicating myelin splitting in the entire brain including the optic chiasm (Figure 2). The gross architecture of the cerebellum appeared normal, and no significant degeneration or loss of Purkinje cells and granule layer neurons was observed (Figure 2a). However, mild astrogliosis of the molecular layer and scattered vacuoles were present (Figure 2b). Additionally, mild to severe diffuse astrogliosis was observed in both white and grey matter, rarely associated with the presence of large axonal spheroids or axonal degeneration (Figure 2d,e). These changes were most severe in the vestibulocochlear nucleus, cerebellar nuclei, thalamus, and brainstem. In addition, the histopathological examination of the eye of one affected dog was carried out and showed no abnormalities.

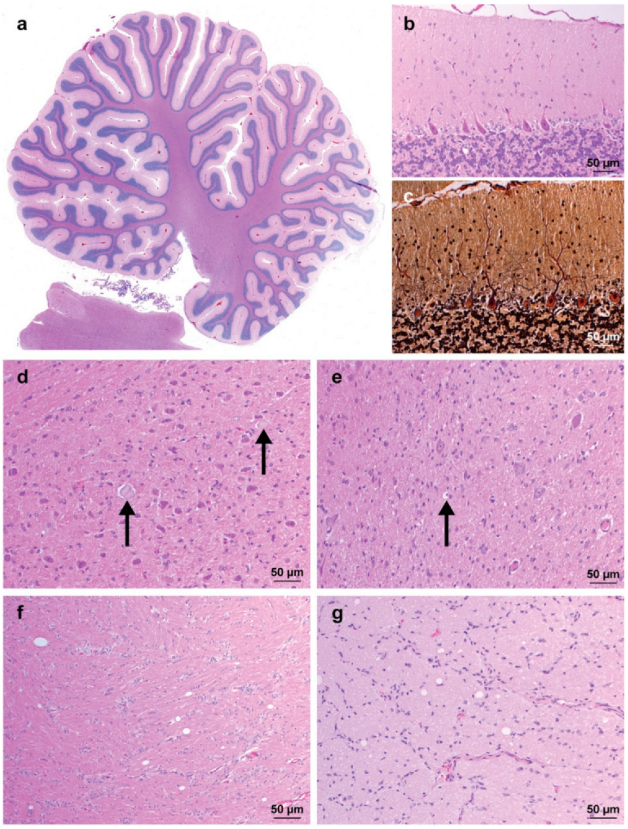

Figure 2. Neuropathology of a spinocerebellar ataxia-affected Alpine Dachsbracke dog. (**a**) Overview of the cerebellum showing a normal architecture. HE staining; (**b**) Cerebellar cortex with normal Purkinje cell layer. The molecular layer exhibits a slightly increased cellularity due to mild astrogliosis. Astrocytes have rather open-faced nuclei. Note the single, clearly delineated empty vacuole in the granule cell layer. Hematoxylin and eosin (HE) staining; (**c**) Cerebellar cortex with normal Purkinje cell layer. No empty baskets are observed. Bielschowsky stain; (**d**) Vestibulocochlear nucleus with marked astrogliosis and two axonal spheroids (arrows). HE staining; (**e**) Cerebellar nucleus showing marked astrogliosis and a single dilated myelin sheath with a central myelinophage (arrow). HE staining; (**f**) Corona radiata exhibiting diffuse astrogliosis and multiple clearly delineated empty vacuoles. HE staining; (**g**) Optic chiasm with multiple clearly delineated empty vacuoles. HE staining.

3.3. Positional Cloning of the Spinocerebellar Ataxia-Associated Locus

Parametric linkage analysis for a recessive trait was performed in the complete litter and resulted in 23 linked genome regions with a positive logarithm of the odds (LOD) score (Supplementary Table S3). Based on the pedigree analysis we assumed that the affected dogs were inbred to a common founder and identical by descent (IBD) for the causative genetic variant and flanking chromosomal regions. Therefore, autozygosity mapping approach was used to determine regions of homozygosity shared across all four cases. Four genome regions with a total of ~22.4 Mb were identified (Supplementary Table S4). By overlapping the linked and homozygous intervals, four chromosomal segments (one on chromosomes 1, two on chromosome 27, and one on chromosome 31) remained as plausible regions of interest (Figure 3). Analyzing the homozygous regions in eight genotyped control dogs including six normal littermates and two parents revealed 144 blocks of homozygosity across the whole genome (Supplementary Table S5). We further looked for intervals showing both linkage and homozygosity

in affected dogs and no homozygosity in any control dog. Only a single interval on chromosome 27, containing 269 SNVs and spanning ~5.5 Mb (chr27:3,154,712-8,679,881), remained as the most likely critical region for spinocerebellar ataxia (Figure 3).

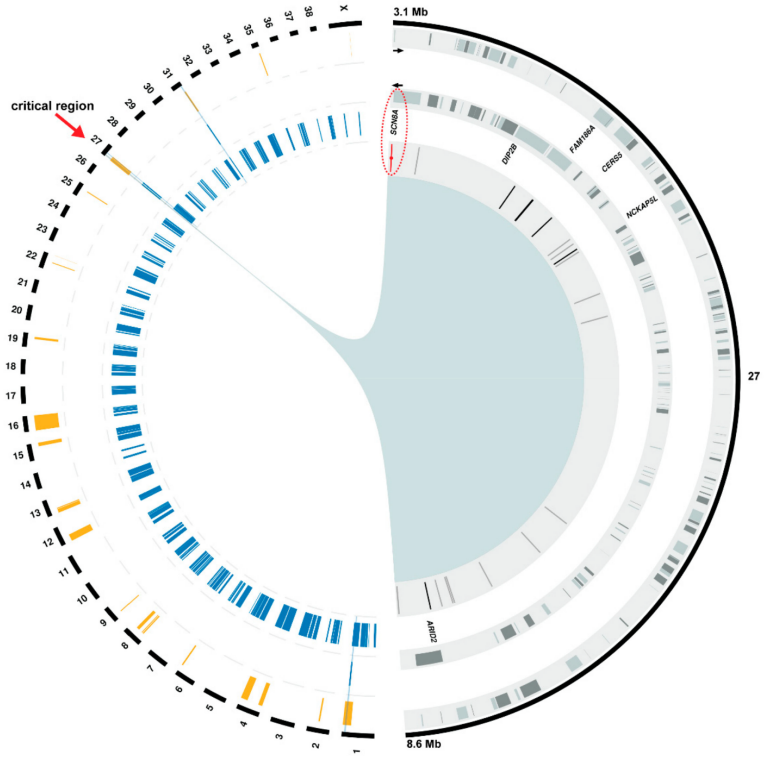

Figure 3. Positional cloning of the spinocerebellar ataxia-associated variant. The canine chromosomes are depicted in the left half of the circle as black bars. Below, the 3 circular tracks from outside to inside indicate: (i) 23 genome segments from linkage analysis in orange, (ii) four homozygous blocks shared in affected dogs in blue, (iii) homozygous blocks found in unaffected dogs in blue. The initial four segments overlapping in cases are indicated by a blue background. Only a single interval on chromosome 27 showed both linkage and homozygosity in affected dogs and no homozygosity in unaffected dogs. Therefore, it was considered the critical region for spinocerebellar ataxia (red arrow). The right half of the circle displays a close-up view of this region (chr27:3,154,712-8,679,881) with 3 circular tracks below showing from outside to inside: (i, ii) gene content showing the 165 genes (grey boxes) annotated on both, the positive (outer circle) and negative (inner circle) strand, (iii) location of private intergenic (grey), intronic (black), and exonic (red) SNVs depicted as vertical lines. For clarity, only the names of six genes, in which an intronic or exonic SNV was found, are shown. Note the red ellipse highlighting the single protein-changing SNV in the *SCN8A* gene.

3.4. Identification of the Causative Variant

In total, 165 genes were annotated in the disease-associated ~5.5 Mb critical interval on chromosome 27 (Figure 3). No structural variants were found by visual investigation of this genomic region in the whole-genome sequence of an affected dog. Filtering of the WGS data for private variants present in a homozygous state in the affected dog and absent in 600 control dog and wolf genomes yielded 21 SNVs within the critical interval (Supplementary Table S6). Six variants were located in introns of five genes (*DIP2B*, *FAM186A*, *CERS5*, *NCKAP5L*, and *ARID2*) while 14 others were intergenic. Only one exonic,

protein-changing variant in a compelling functional candidate gene *SCN8A* (chr27:g.3,179,029C>A) was found (Figure 3).

The detected variant (*SCN8A*:c.4898G>T) alters the encoded amino acid of SCN8A residue 1633 (p.Gly1633Val) with a predicted moderate impact on the resulting protein (Figure 4a). The missense variant is located in the last exon of the sodium voltage-gated channel α subunit 8 gene (Figure 4b) and affects the conserved ion transport domain 4 of SCN8A protein (Figure 4c). The α subunit consists of four homologous ion transport domains, each of which is organized in six transmembrane segments and contains a pore loop between the last two segments [15] (Figure 4c). The amino acid is highly conserved across species (Table S7) and the glycine to valine substitution was predicted pathogenic and deleterious by several in silico prediction tools (PON-P2 probability for pathogenicity: 0.891, MutPred2 score: 0.930, PROVEAN score: −7.256). Furthermore, there is no non-synonymous variant in the human *SCN8A* coding region at the corresponding position depicted in the gnomAD [14].

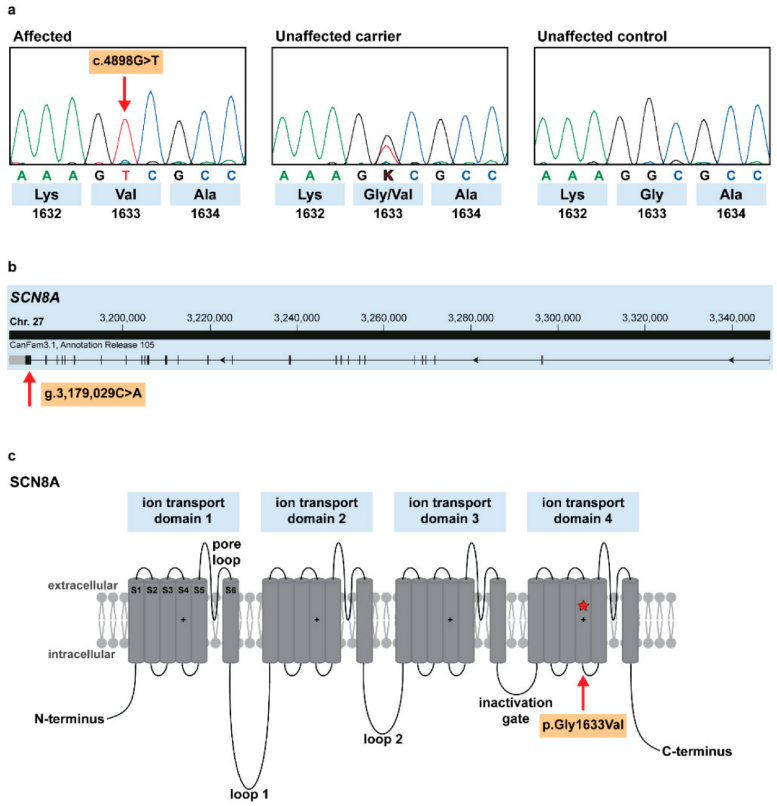

Figure 4. Spinocerebellar ataxia-associated SCN8A missense variant in Alpine Dachsbracke dogs. (**a**) Chromatograms of an affected, carrier, and wild type individual dogs, indicate the c.4898G>T variant, which changes codon 1633 (shown below); (**b**) *SCN8A* gene structure showing the variant location in exon 27 (red arrow); (**c**) Schematic representation of SCN8A protein and its four ion transport domains, which are separated by cytoplasmic loops and a short inactivation gate, each domain contains six transmembrane segments S1–S6 (adapted from [15]). Our p.Gly1633Val variant affects the positively charged S4 of the last domain (red star and arrow).

The perfect segregation of the detected *SCN8A* variant with the observed disease phenotype was further confirmed by genotyping all 216 available Alpine Dachsbracke dogs. Only the four affected dogs were homozygous for the variant allele, three obligate carriers, eight tested littermates in addition

to eight controls were heterozygous carriers of the variant, whereas it was absent in 193 controls (Table 1).

Table 1. Segregation of the *SCN8A*:c.4898G>T genotypes with spinocerebellar ataxia in Alpine Dachsbracke dogs.

Disease Status	G/G	G/T	T/T
Affected (n = 4)	0	0	4
Non-affected (n = 212)	193	19 [1]	0

[1] includes 3 obligate carriers and 8 littermates of the affected dogs.

4. Discussion

Based on clinicopathological data a novel congenital form of spinocerebellar ataxia was diagnosed in four Alpine Dachsbracke dogs. The neuropathological examination of two affected dogs indicated clear, even though not extremely severe, structural changes in the brain. Nevertheless, the observed pathological changes showed similarities to other spinocerebellar ataxias, such as *KCNJ10*-related spongy degeneration with cerebellar ataxia (SDCA1) in Belgian Shepherd Dogs (OMIA 002089) [9], and the vacuolization with marked astrogliosis indicated intramyelinic edema, which was compatible with an ion channel defect. We have used genome-wide SNV data for positional cloning of the disease-associated locus. Because of the population structure of the studied purebred Alpine Dachsbracke dogs, we were able to narrow down the locus by linkage analysis and autozygosity mapping to four chromosomal segments. Additionally, the normal related control dogs shared three of these homozygous intervals with identical alleles, which allowed us to focus on the ~5.5 Mb region on chromosome 27. Further molecular genetic analysis using current sequencing methods revealed a single most likely causative variant in a functional candidate gene.

SCN8A encodes the α subunit of a voltage-gated sodium channel $Na_v1.6$, which is important for sodium ion transport to neurons in the central as well as peripheral nervous systems. However, it is most abundant in the maturing nodes of Ranvier in myelinated axons of the central nervous system, in which it contributes to nerve conduction velocity [16,17]. $Na_v1.6$ was also found at the axon initial segment of excitatory and inhibitory neurons, where it regulates the formation of action potentials, which is essential for membrane depolarization [18]. Voltage-gated sodium channels consist of one α subunit forming the pore and up to two β subunits. The α subunit is organized in four repeated domains linked by cytoplasmic loops, and each domain is made up of six transmembrane segments (S1–S6). First four segments of each homologous domain form the primary voltage sensor for activation, and last two segments together with the pore loop form the ion pore. After depolarization, the positively charged S4 rotates and initiates a conformational change, which is important for opening the sodium channel pore. [16].

Altering of the sodium channels leads to abnormally, either increased or decreased, neuronal signaling as previously described in human patients suffering from epileptic encephalopathy [19]. Mutations of the $Na_v1.6$ channel are described in ~1% of almost 1500 children affected by early-onset epileptic encephalopathies [20]. Protein-changing *SCN8A* variants have been previously associated with clinically similar movement disorders including cerebellar ataxia, pancerebellar atrophy, and cognitive disability in humans. Interestingly, all known *SCN8A*-associated human neurodegenerative diseases are caused by dominant acting variants (OMIM 600702).

The effects of loss of $Na_v1.6$ channel have been extensively studied in spontaneous mouse mutants and can cause movement abnormalities, including ataxia, tremor, muscle weakness, and atrophy, or dystonia. Mice with null mutations exhibit hind limb paralysis, motor impairments, and early death, while heterozygous carriers show milder phenotypes [21]. A missense mutation in *Scn8a* is associated with cerebellar ataxia in the so-called 'jolting' mutant (Supplementary Table S7) and is a result of a shift in voltage dependence of the channel opening [22].

The herein described substitution in the Alpine Dachsbracke dogs alters a conserved position predicted to be in the transmembrane segment S4 in the fourth homologous domain; *in silico* analysis predicts the variant to be probably damaging to the protein structure/function. Heterozygous carriers do not exhibit a visible clinical phenotype, as they can most likely compensate due to the presence, albeit at a reduced level, of the normal proteins. As we identified a non-synonymous variant in a highly plausible functional candidate gene, our mapping analyses and whole-genome resequencing data combined with the current knowledge on SCN8A function in humans and mice strongly support the causality of *SCN8A*:c.4898G>T variant in the Alpine Dachsbracke breed. However, follow-up functional experiments would be needed to fully show the deleterious effects of the putative causal variant on the gene function.

In conclusion, our results provide strong evidence for a breed-specific missense variant in *SCN8A* gene as the most likely causative genetic variant for monogenic recessive spinocerebellar ataxia in Alpine Dachsbracke dogs. This is the first report of an *SCN8A*-associated form of ataxia in dogs. Our results enable the development of a genetic test for veterinary diagnostic and breeding purposes.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/10/5/362/s1, Table S1: Detailed information and genotypes of 216 Alpine Dachsbracke dogs, Table S2: Sample designations and breed information of all whole-genome sequences, Table S3: Genome regions with positive LOD scores for linkage to ataxia in the Alpine Dachsbracke family, Table S4: Shared homozygous regions across the four analyzed Alpine Dachsbracke ataxia cases, Table S5: Genome-wide homozygous regions in the eight analyzed control Alpine Dachsbracke dogs, Table S6: List of private variants of the sequenced Alpine Dachsbracke case in the region of interest on chromosome 27, Table S7: Alignment of SCN8A in vertebrate species. Video S1: Video illustrating the clinical phenotype of an affected Alpine Dachsbracke dog.

Author Contributions: Conceptualization, C.D.; Data curation, V.J.; Formal analysis, A.L.; Funding acquisition, C.D.; Investigation, A.L., C.G. and A.O.; Methodology, A.L.; Project administration, C.D.; Resources, E.D. and M.N.; Software, V.J.; Supervision, C.D.; Visualization, A.L.; Writing—original draft, A.L.; Writing—review & editing, A.L., E.D., C.G., A.O. and C.D.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to the breeder of the two affected litters and all other owners and breeders and the Verein Dachsbracke e.V., namely Roger Hörr, who provided samples and shared pedigree data and other valuable information of their dogs. We acknowledge collaborators of the Dog Biomedical Variant Database Consortium (Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Oliver Forman, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Anita Oberbauer, Jeffrey Schoenebeck, Sheila Schmutz, Claire Wade) for sharing whole-genome sequence data of control dogs. We thank Nathalie Besuchet Schmutz for expert technical assistance, the Next Generation Sequencing Platform of the University of Bern for performing the whole-genome resequencing experiments, and the Interfaculty Bioinformatics Unit of the University of Bern for providing computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Manto, M.; Marmolino, D. Cerebellar ataxias. *Curr. Opin. Neurol.* **2009**, *22*, 419–429. [CrossRef]
2. Jayadev, S.; Bird, T.D. Hereditary ataxias: Overview. *Genet. Med.* **2013**, *15*, 673–683. [CrossRef]
3. Sandford, E.; Burmeister, M. Genes and genetic testing in hereditary ataxias. *Genes* **2014**, *5*, 586–603. [CrossRef] [PubMed]
4. Klockgether, T.; Paulson, H. Milestones in Ataxia. *Mov. Disord.* **2011**, *26*, 1134–1141. [CrossRef] [PubMed]
5. Urkasemsin, G.; Olby, N.J. Canine Hereditary Ataxia. *Vet. Clin. N. Am. Small Anim. Pract.* **2014**, *44*, 1075–1089. [CrossRef]
6. Chang, C.C.; Chow, C.C.; Tellier, L.C.A.M.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 1–16. [CrossRef]
7. Abecasis, G.R.; Cherny, S.S.; Cookson, W.O.; Cardon, L.R. Merlin—Rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **2002**, *30*, 97–101. [CrossRef] [PubMed]
8. Hu, Y.; Yan, C.; Hsu, C.H.; Chen, Q.R.; Niu, K.; Komatsoulis, G.A.; Meerzaman, D. OmicCircos: A simple-to-use R package for the circular visualization of multidimensional Omics data. *Cancer Inform.* **2014**, *13*, 13–20. [CrossRef] [PubMed]

9. Mauri, N.; Kleiter, M.; Leschnik, M.; Högler, S.; Dietschi, E.; Wiedmer, M.; Dietrich, J.; Henke, D.; Steffen, F.; Schuller, S.; et al. A Missense Variant in KCNJ10 in Belgian Shepherd Dogs Affected by Spongy Degeneration with Cerebellar Ataxia (SDCA1). *G3 Genes Genomes Genetics* **2017**, *7*, 663–669. [CrossRef] [PubMed]
10. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [CrossRef]
11. Niroula, A.; Urolagin, S.; Vihinen, M. PON-P2: Prediction method for fast and reliable identification of harmful variants. *PLoS ONE* **2015**, *10*. [CrossRef] [PubMed]
12. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef] [PubMed]
13. Pejaver, V.; Urresti, J.; Lugo-Martinez, J.; Pagel, K.A.; Lin, G.N.; Nam, H.-J.; Mort, M.; Cooper, D.N.; Sebat, J.; Iakoucheva, L.M.; et al. MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. *bioRxiv* **2017**, 134981. [CrossRef]
14. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *bioRxiv* **2019**, 531210. [CrossRef]
15. Catterall, W.A. International Union of Pharmacology. XLVII. Nomenclature and Structure-Function Relationships of Voltage-Gated Sodium Channels. *Pharmacol. Rev.* **2005**, *57*, 397–409. [CrossRef]
16. Catterall, W.A. Sodium Channels. *Encycl. Life Sci.* **2006**, 65–71. [CrossRef]
17. Kearney, J.A.; Buchner, D.A.; De Haan, G.; Adamska, M.; Levin, S.I.; Furay, A.R.; Albin, R.L.; Jones, J.M.; Montal, M.; Stevens, M.J.; et al. Molecular and pathological effects of a modifier gene on deficiency of the sodium channel Scn8a (Na(v)1.6). *Hum. Mol. Genet.* **2002**, *11*, 2765–2775. [CrossRef]
18. Hu, W.; Tian, C.; Li, T.; Yang, M.; Hou, H.; Shu, Y. Distinct contributions of Nav1.6 and Nav1.2 in action potential initiation and backpropagation. *Nat. Neurosci.* **2009**, *12*, 996–1002. [CrossRef]
19. Liu, Y.; Schubert, J.; Sonnenberg, L.; Helbig, K.L.; Hoei-hansen, C.E.; Koko, M.; Rannap, M.; Lauxmann, S.; Huq, M.; Schneider, M.C.; et al. Neuronal mechanisms of mutations in SCN8A causing epilepsy or intellectual disability. *Brain* **2019**, 376–390. [CrossRef]
20. Wagnon, J.L.; Meisler, M.H. Recurrent and non-recurrent mutations of SCN8A in epileptic encephalopathy. *Front. Neurol.* **2015**, *6*, 1–7. [CrossRef]
21. Chen, K.; Godfrey, D.A.; Ilyas, O.; Xu, J.; Preston, T.W. Cerebellum-related characteristics of Scn8a-Mutant mice. *Cerebellum* **2009**, *8*, 192–201. [CrossRef] [PubMed]
22. Kohrman, D.C.; Smith, M.R.; Goldin, A.L.; Harris, J.; Meisler, M.H. A missense mutation in the sodium channel Scn8a is responsible for cerebellar ataxia in the mouse mutant jolting. *J. Neurosci.* **1996**, *16*, 5993–5999. [CrossRef] [PubMed]

A missense variant in *ALDH5A1* associated with canine succinic semialdehyde dehydrogenase deficiency (SSADHD) in the Saluki dog

Journal:	Genes
Manuscript status:	published
Contributions:	Investigation, Formal analysis, Visualization, Writing – review & editing

genes

MDPI

Article

A Missense Variant in *ALDH5A1* Associated with Canine Succinic Semialdehyde Dehydrogenase Deficiency (SSADHD) in the Saluki Dog

Karen M. Vernau [1,*], Eduard Struys [2], Anna Letko [3], Kevin D. Woolard [4], Miriam Aguilar [5], Emily A. Brown [5], Derek D. Cissell [1], Peter J. Dickinson [1], G. Diane Shelton [6], Michael R. Broome [7], K. Michael Gibson [8], Phillip L. Pearl [9], Florian König [10], Thomas J. Van Winkle [11], Dennis O'Brien [12], B. Roos [2], Kaspar Matiasek [13], Vidhya Jagannathan [3], Cord Drögemüller [3], Tamer A. Mansour [5,14], C. Titus Brown [5] and Danika L. Bannasch [5,*]

1 Department of Surgical and Radiological Sciences, University of California Davis, Davis, CA 95616, USA; ddcissell@ucdavis.edu (D.D.C.); pjdickinson@ucdavis.edu (P.J.D.)
2 Department of Clinical Chemistry, VU University Medical Center, 1081 HV Amsterdam, The Netherlands; E.Struys@vumc.nl (E.S.); b.roos@amsterdamumc.nl (B.R.)
3 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland; anna.letko@vetsuisse.unibe.ch (A.L.); vidhya.jagannathan@vetsuisse.unibe.ch (V.J.); cord.droegemueller@vetsuisse.unibe.ch (C.D.)
4 Department of Pathology, Microbiology and Immunology, University of California Davis, Davis, CA 95616, USA; kdwoolard@ucdavis.edu
5 Department of Population Health and Reproduction, University of California Davis, Davis, CA 95616, USA; miraguilar@ucdavis.edu (M.A.); eabrown@ucdavis.edu (E.A.B.); drtamermansour@gmail.com (T.A.M.); ctbrown@ucdavis.edu (C.T.B.)
6 Department of Pathology, University of California San Diego, La Jolla, CA 92093, USA; gshelton@health.ucsd.edu
7 Advanced Veterinary Medical Imaging, Tustin, CA 92780, USA; mbroome@avmi.net
8 College of Pharmacy and Pharmaceutical Sciences, Washington State University, Spokane, WA 99202, USA; mike.gibson@wsu.edu
9 Harvard Medical School, Boston, MA 02115, USA; Phillip.Pearl@childrens.harvard.edu
10 Fachtierarzt fur Kleintiere, Am Berggewann 13, 65199 Wiesbaden, Germany; fk@neurovet.de
11 Department of Pathobiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA; tomvw@vet.upenn.edu
12 College of Veterinary Medicine, University of Missouri, Columbia, MO 65211, USA; ObrienD@missouri.edu
13 Clinical and Comparative Neuropathology, Ludwig-Maximilians-Universitaet München, 80539 Munchen, Germany; kaspar.matiasek@neuropathologie.de
14 Department of Clinical Pathology, School of Medicine, Mansoura University, Mansoura 35516, Egypt
* Correspondence: kmvernau@ucdavis.edu (K.M.V.); dlbannasch@ucdavis.edu (D.L.B.)

Received: 7 August 2020; Accepted: 27 August 2020; Published: 2 September 2020

check for updates

Abstract: Dogs provide highly valuable models of human disease due to the similarity in phenotype presentation and the ease of genetic analysis. Seven Saluki puppies were investigated for neurological abnormalities including seizures and altered behavior. Magnetic resonance imaging showed a diffuse, marked reduction in cerebral cortical thickness, and symmetrical T2 hyperintensity in specific brain regions. Cerebral cortical atrophy with vacuolation (status spongiosus) was noted on necropsy. Genome-wide association study of 7 affected and 28 normal Salukis revealed a genome-wide significantly associated region on CFA 35. Whole-genome sequencing of three confirmed cases from three different litters revealed a homozygous missense variant within the aldehyde dehydrogenase 5 family member A1 (*ALDH5A1*) gene (XM_014110599.2: c.866G>A; XP_013966074.2: p.(Gly288Asp). *ALDH5A1* encodes a succinic semialdehyde dehydrogenase (SSADH) enzyme critical in the gamma-aminobutyric acid neurotransmitter (GABA) metabolic pathway. Metabolic screening

of affected dogs showed markedly elevated gamma-hydroxybutyric acid in serum, cerebrospinal fluid (CSF) and brain, and elevated succinate semialdehyde in urine, CSF and brain. SSADH activity in the brain of affected dogs was low. Affected Saluki dogs had striking similarities to SSADH deficiency in humans although hydroxybutyric aciduria was absent in affected dogs. *ALDH5A1*-related SSADH deficiency in Salukis provides a unique translational large animal model for the development of novel therapeutic strategies.

Keywords: inborn error of metabolism; encephalopathy; SSADHD; *ALDH5A1*; GABA; 4-hydroxybutyric acid; succinic semialdehyde; encephalopathy; whole-genome sequencing; precision medicine; GWAS; inherited

1. Introduction

Inborn errors of metabolism (IEMs) are a group of diseases caused by an enzymatic deficiency in a metabolic pathway, most commonly caused by a genetic mutation. While individually these diseases are rare, as a group they are relatively common, with more than 500 IEM diseases reported in people [1]; in animals, they are becoming increasingly recognized [2–6]. In diseases caused by an IEM, clinical signs are due to the reduced or lack of production of a biochemical product, or accumulation of an abnormal amount of substrate or substrates produced by alternative metabolic pathways, secondary to the enzymatic deficiency. The diagnosis of an IEM may be a challenging, as clinical signs can be vague and non-specific, and targeted diagnostic testing is required [7]. IEMs are often recognized in young people and animals, and many have neurological manifestations [8,9].

Seizures are a common neurological sign in dogs [10]. Disorders causing seizures arise either extracranially (reactive seizures), or intracranially [11] Epilepsy is a brain disease characterized by a lasting predisposition to generate seizures, which is classified in dogs as structural epilepsy or idiopathic epilepsy (OMIA 000344-9615) [11]. Causes of structural epilepsy include inflammation (e.g., granulomatous meningoencephalitis), neoplasia, nutritional alterations (e.g., thiamine deficiency), infection, anomalous entities (e.g., hydrocephalus), inborn errors of metabolism and trauma [11]. Dogs with idiopathic epilepsy (IE) are typically 6 months to 6 years of age and usually have normal physical and neurological examinations between seizures [12]. Dogs younger than 6 months or older than 6 years of age usually have reactive seizures or structural epilepsy, rather than idiopathic epilepsy [12].

A seizure disorder reported in Salukis is called central nervous system status spongiosus in Saluki dogs (SSSD). There are only brief reports of this disease in the literature [13,14]. One affected 8-month-old male puppy from a litter of 9 was reported with a 5 month history of seizures and behavioral changes. The sire and dam were full siblings. All nine puppies and the sire were euthanized; pathological changes were noted in the affected puppy and in two clinically normal littermates, and the rest of the puppies and the sire were pathologically normal. Pathological changes in the clinically affected puppy included widespread bilaterally symmetrical status spongiosis of the cerebrum, brainstem and cerebellum at the grey–white matter junction, which extended into both the grey and white matter. There were also lesions in the thalamus, optic nerve and internal capsule but no lesions were noted in the spinal cord [13].

Recognized causes of early-onset symmetrical brain lesions include metabolic, nutritional and toxin-induced diseases [15]. In Saluki dogs with SSSD, the clinical signs and lesions on MRI and pathology appear to be breed specific, identical in distribution and type, and diagnosed in multiple dogs over a long period of time (1987 [13] to 2020). Clinical signs developed while puppies were still with the breeders, making toxicity a less likely cause; pathology differed from previously reported nutritional [16] or toxic [17–19] central nervous system problems, and thus a genetic cause was considered most likely. Although a metabolic disorder was not identified by routine diagnostic testing

in affected Salukis, an underlying genetic abnormality causing a metabolic problem was most likely based on the age of onset of clinical signs.

The purpose of this study was to define the phenotype of Salukis with SSSD and to determine the underlying genetic cause in this breed. Comprehensive evaluations including MRI and necropsy, as well as metabolic and enzyme activity testing, were performed on urine, serum, cerebrospinal fluid and brain tissue from four affected puppies from two litters from the USA and a litter with three affected puppies from Germany. All seven affected dogs were used for a genome-wide association study (GWAS) followed by whole-genome sequence analysis of three affected puppies, which identified a private homozygous missense variant in the canine *ALDH5A1* gene.

2. Materials and Methods

2.1. Affected Dogs

From 2005 to 2015, seven Saluki dogs affected with SSSD from the USA (4) and Germany (3) had DNA collected. Four dogs were examined—three dogs at the William R. Pritchard Veterinary Medical Teaching Hospital at the University of California Davis (UCD) and one dog was evaluated at Fachtierarzt fur Kleintiere, in Wiesbaden, Germany. All 4 dogs were presented by their breeders for examination. A fifth Saluki dog affected with SSSD from the USA had a necropsy completed at UCD Two additional German dogs were not evaluated clinically beyond the breeder's description of the clinical signs.

2.2. Control Dogs

2.2.1. MRI Evaluation

Four unaffected Saluki dogs related to the affected USA Saluki dogs were examined and had magnetic resonance imaging of their brain and completed at Advanced Veterinary Medicine Imaging in Los Angeles, California.

2.2.2. Targeted Metabolic Testing

Archived urine = 4, serum = 4, cerebrospinal fluid (3) and brain tissue (4) from 15 different non-Saluki dogs unaffected by SSSD were utilized as control samples.

2.3. Affected Saluki Dogs

Blood work (complete blood count, and serum biochemical profile) was performed by the referring veterinarians in two dogs (dogs 3 and 4). Cerebrospinal fluid sample (CSF) was collected in two dogs (dogs 1 and 5). One CSF sample was routinely analyzed in Germany and the other sample was collected at UCD and frozen at −80 degrees for further analysis. Two dogs had quantitative urine organic acid testing completed at the University of California San Diego Biochemical Genetics Laboratory (dogs 1 and 2). Urine was shipped to the lab by the breeder and was analyzed by the lab 18 days later. Four dogs had complete necropsies completed at UCD, and one had a necropsy at Ludwig Maximilians Universität München in Germany. The owners consented to the necropsy and processing of postmortem samples. Following necropsy, the brain was immediately immersed in 10% neutral buffered formalin followed by standard paraffin embedding. Selected regions were sectioned at 5 µm slice thickness and stained with hematoxylin-eosin and luxol fast blue-cresyl violet.

MRI and Histopathology

Six Saluki dogs underwent MRI of the brain—two affected Saluki dogs underwent magnetic resonance imaging (MRI) of the brain at UC Davis, and four unaffected Saluki dogs had imaging of the brain at Advanced Veterinary Medicine Imaging in Los Angeles, California. Both locations used a 1.5 T MRI system (GE Signa, GE Healthcare, Waukesha, WI, USA), with paired 5" general purpose

radiofrequency coils. Sagittal T1-weighted (T1W) and T2-weighted (T2W) images, transverse T1W, T2W, fluid attenuating inversion recovery (FLAIR), and T2*-weighted (T2*W) images, and dorsal T2W images were acquired of the brain. Sagittal, transverse, and dorsal T1W images were repeated after intravenous administration of 0.1 mmol/kg gadopentetate dimeglumine (Magnevist, Bayer, Whippany, NJ, USA).

2.4. Sample Collection and DNA Extraction

Blood samples, pedigree, and phenotype information were collected from 7 affected dogs and 18 close relatives of affected dogs including 4 parents (Figure 1). Additional samples were collected from 48 healthy Saluki dogs. Healthy dogs of other breeds (n = 228) were used that were part of a DNA repository at UC Davis. DNA was extracted from EDTA whole blood samples using Gentra Puregene DNA purification extraction kit (Qiagen, Valencia, CA, USA). Collection of canine samples was approved by the University of California, Davis Animal Care and Use Committee (protocol #18561) and the Cantonal Committee for Animal Experiments (Canton of Bern; permit 75/16).

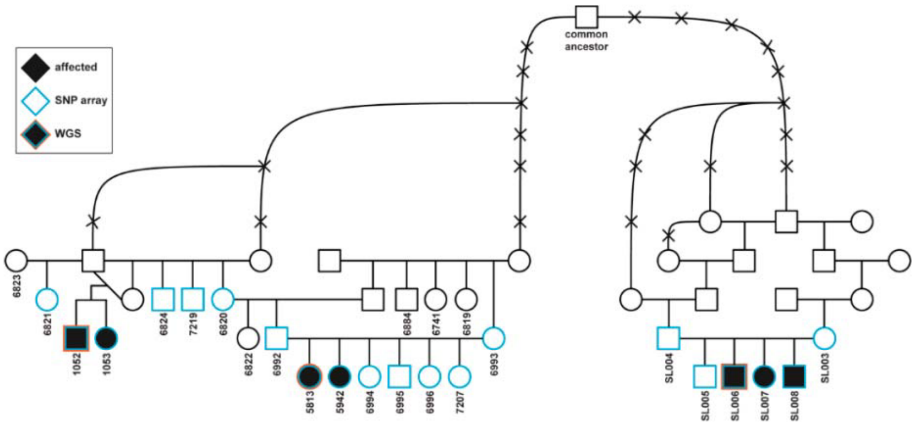

Figure 1. Pedigree of seven succinic semialdehyde dehydrogenase deficiency (SSADHD)-affected Saluki dogs. Females are depicted by circles and males by squares. Black fill indicates affected puppies. Numbers indicate 25 dogs from which samples were available. The blue contour indicates animals that were genotyped on SNP array, and the red contour the three affected dogs selected for the WGS. Note that the two litters on the left were seen in the USA and the third litter on the right in Germany. A common male ancestor illustrates the genealogical relatedness.

2.5. Genome-Wide Association Scan

SNP genotyping was performed using the Illumina Canine HD 174,000 SNP array (Illumina, San Diego, CA, USA) for 7 affected cases and 28 neurologically normal adult Saluki controls. Genome-wide association analysis was performed using Plink [20]. SNPs were pruned from analysis if the minor allele frequency was <5% and the call rate <90%. Chi-square association analysis, Bonferroni adjustments, and genomic inflation calculations were performed in Plink. Figure 5 was made in R using ggplot2 [21,22].

2.6. Whole-Genome Sequence

Whole-genome sequencing (WGS) was performed on the two affected Salukis from the USA and compared to 98 controls dogs from various breeds as reported, and coverage was 6.4× for 1052 and 5.3× for 5813 SRA: SRR5311685 and SRR5311664 (study: PRJNA377155) [23]. Segregation of variants was performed using 2 cases compared to 98 controls within the homozygous interval identified by visual inspection of genotype calls from the array data. Variant Effect Predictor (VEP), as employed in Ensembl using Refseq annotation and additional EST/CCDs, was used to predict the effect of

segregating variants [24]. Polyphen2 [25] and SIFT [26] were used to evaluate the severity of the missense variants.

In one German Saluki (dog 5), whole-genome sequencing using genomic DNA isolated from the blood sample of the affected dog was performed as described previously [27]. Data corresponding to approximately 15× coverage of the genome was collected on an Illumina HiSeq2000 instrument (2 × 100 bp). Read mapping and variant calling were carried out as previously described [28], with respect to the CanFam3.1 genome reference assembly and the NCBI annotation release 105. Variant filtering was performed against 581 dog and 8 wolf genomes which were publicly available [28]. WGS data of the affected dog was made available under study accession PRJEB16012 at the European Nucleotide Archive (www.ebi.ac.uk/ena; sample accession SAMEA4504825).

Annotations within the canine *ALDH5A1* gene refer to the NCBI mRNA accession no. XM_014110599.2 and the protein accession no. XP_013966074.2. Annotations within the *GPLD1* gene refer to the NCBI mRNA accession no. XM_005640079.3 and the protein accession number XP_005640136.1. Annotations for the putative *PTCHD3* gene refer to the mRNA accession no. ENSCAFT00000039442.3 and the protein accession no. ENSCAFP00000035301.

2.7. Genotyping

The variant in *ALDH5A1* disrupted a Sau96I restriction enzyme site, allowing rapid genotyping by PCR-RFLP analysis. PCR primers were designed using primer 3 [29] to amplify an 872 bp product, which upon digestion with Sau96I produced 700 and 150 bp fragments for the variant allele and 550 and 150 bp fragments for the wild-type allele. All PCR was carried out using Qiagen HotStart DNA polymerase kit (Qiagen, Valencia, CA, USA) at an annealing temperature of 58 degrees using the following PCR primers: F_TCCCGAGTTAGGGGTTCTTT, R_TCACGTTTTCCTGATTTCACC. The same primers were used to verify the mutation by Sanger sequencing on an Applied Biosystems 3500 Genetic Analyzer using the Big Dye Terminator Sequencing Kit (Life Technologies, Burlington, ON, Canada).

2.8. RT-PCR

RT-PCR was performed for liver cDNA from a case and control. *RPS5* was included as a housekeeping gene control [27]. Primers were designed using Primer3 (SAL_2F: TTGTATTTGACAGC GCCAAC, SAL_2R: CAAGGCCAGATTGCTTCAC) except for RPS5, in which the primers were as recommended [30]. Each reaction included 13.9 µL of water, 2 µL of 10× buffer with $MgCl_2$, 1 µL of dNTP, 1 µL of each forward and reverse primers (20 µM), 1 µL of HotStarTaq DNA Polymerase (Qiagen, Valenica, CA, USA), and 1 µL of cDNA made from 1000 ng of RNA. Amplified products were visualized on a 2% agarose gel.

2.9. Targeted Metabolic Testing

Gamma-hydroxybutyrate (GHB) and succinate semialdehyde (SSA):

GHB and SSA in fluids and brain tissue were quantified by isotope dilution mass spectrometric methodology, as previously described [31,32].

The 4,5-dihydroxyhexanoic acid (DHHA):

Analysis of DHHA in fluids and brain was comparable to that for GHB as previously described [31], with some modifications. $^2H_3$-DHHA was used as the internal standard and the samples were extracted a single time with ethylacetate. For quantitation, positive chemical ionization was employed.

Succinic semialdehyde dehydrogenase (SSADH) activity:

SSADH activity was quantified fluorometrically in brain tissue samples using the NADH/NAD couple and SSA as substrate, as previously described [33].

3. Results

3.1. Affected Dogs

Four dogs from two litters from the USA were closely related, and the third litter from Germany was distantly related to the other two litters. There were four females and three males affected (Figure 1).

3.2. Clinical Phenotype

The breeders of affected Saluki puppies noted that puppies were first abnormal between six and ten weeks of age. Historical clinical signs included seizures, abnormal behavior such as episodes of vocalization (Video S1, Supplementary Materials), and difficulty being aroused from sleep. Four puppies (dogs 1 and 3–5) were evaluated by a board-certified veterinary neurologist (Table 1). No abnormalities were noted on physical examination. On neurological examination, puppies had mild generalized ataxia with thoracic limb hypermetria (two puppies) (Video S2, Supplementary Materials), absent menace reflex in both eyes and delayed proprioceptive positioning in all four limbs, consistent with a multifocal disease process. Two dogs (dogs 3 and 4) had a normal CBC and serum biochemical profile completed by their referring veterinarian, and two dogs had normal quantitative urine organic acid analysis (Biochemical Genetics Laboratory, University of California San Diego, San Diego, CA, USA). One dog (dog 5) had a normal cisternal cerebrospinal fluid analysis. Five dogs were treated for seizures with oral phenobarbital or levetiracetam. Although clinical signs did not progress, all affected dogs were euthanized as puppies at the request of the breeders when they were still in their care. Puppies were euthanized between three and nine months of age for quality of life concerns, primarily due to the recurrent episodes of vocalization. Five dogs (dogs 1–5) had necropsies completed.

3.3. MRI and Histopathology

Two affected dogs (dogs 1 and 2) and four other related but unaffected dogs had an MRI of the brain. All four unaffected dogs had unremarkable MR images. Both affected dogs exhibited prominent sulci (Figure 2A,C–E) compared to normal dogs (Figure 2F), consistent with diffuse cortical atrophy. Bilateral, symmetrical, T2 and FLAIR hyperintensity was present in the diencephalon, deep cerebellar nuclei (Figure 2A–C), midbrain (Figure 2D), and multiple basal nuclei (Figure 2E). Multifocal, symmetrical T2 and FLAIR hyperintensity was also present in the deep cortical laminae of the grey matter throughout the cerebral cortex (Figure 2A–F).

No hypointense lesions or signal voids were observed associated within the brain parenchyma on T2*W images.

Histopathologically, there was severe bilaterally symmetric spongiform change, worse within the mesencephalon (Figure 3), brainstem, and deep cerebellar nuclei, but also severe in the thalamic nuclei and deep cortical grey matter. The corpus striatum was less affected but exhibited similar lesions most notably in the entopeduncular nuclei and putamen. Neurons exhibited single to multiple, clear, well-demarcated vacuoles that compressed and displaced the nucleus (Figure 4). There was marked proliferation of enlarged astrocytes associated with the spongiform change, and both neurons and astrocytes appear affected. The grey matter was more severely affected, particularly at the grey–white matter junction. The spinal cord was not affected.

Table 1. Clinical information for seven SSADHD-affected Saluki dogs. (OU = oculus uterque (both eyes)).

Dog No.	Sample ID	Country	Sex	Age of Onset	Clinical Signs	Neurological Examination	Outcome
1	5813	USA	F	10 weeks	Generalized epileptic seizures, episodes of vocalization, abnormal behavior, generalized ataxia with thoracic limb hypermetria	Mild generalized ataxia with thoracic limb hypermetria. Delayed proprioceptive positioning present in all 4 limbs	Treated with anticonvulsants (phenobarbital), euthanized at 32 weeks of age
2	4942	USA	F	10 weeks	Generalized epileptic seizures, episodes of vocalization, abnormal behavior, generalized ataxia with thoracic limb hypermetria	Not done	Treated with anticonvulsants (phenobarbital), euthanized at 39 weeks of age
3	1053	USA	F	6 weeks	Focal epileptic seizures, episodes of vocalization, normal between episodes Unable to arouse when sleeping	Absent menace response OU	Treated with anticonvulsants (phenobarbital), euthanized at 17 weeks of age
4	1052	USA	M	6 weeks	Generalized and focal epileptic seizures, episodes of vocalization, normal between episodes. Unable to arouse when sleeping	Absent menace response OU	Treated with anticonvulsants (phenobarbital), euthanized at 17 weeks of age
5	SL006	Germany	M	9 weeks	Focal epileptic seizures, deep sleep	Thoracic limb hypermetria, mild ataxia, reduced proprioceptive positioning, absent menace	Treated with levetiracetam, euthanized at 4 months of age
6	SL008	Germany	M	9 weeks	Focal epileptic seizures, episodes of vocalization	Not done	Euthanized at unknown age
7	SL007	Germany	F	9 weeks	Focal epileptic seizures	Not done	Euthanized at unknown age

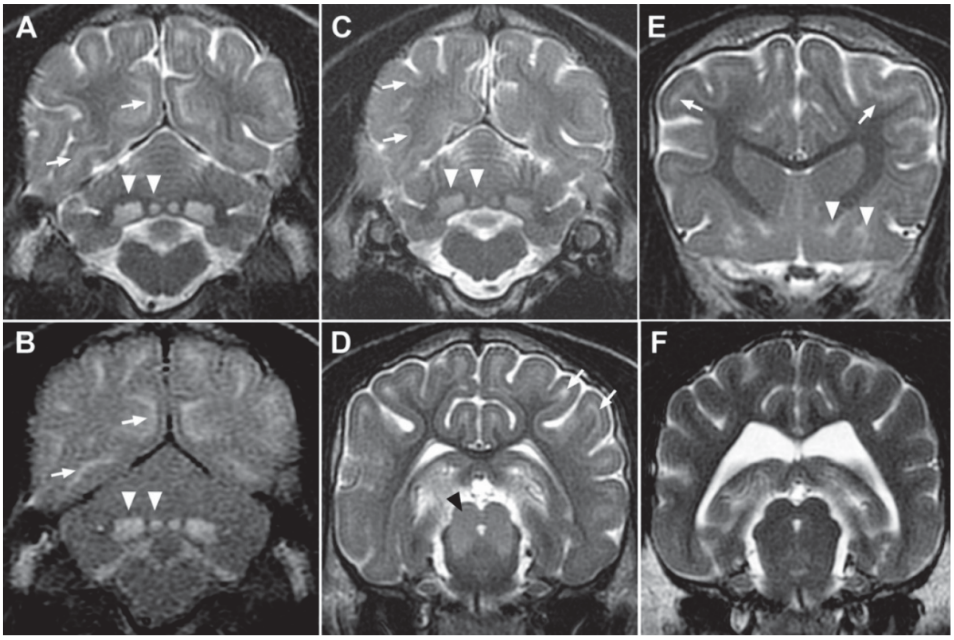

Figure 2. MRI abnormalities in SSADHD-affected Saluki Dogs. Transverse T2-weighted (**A**,**C**–**F**) and FLAIR (**B**) MR images at the level of the cerebellum (**A**–**C**), midbrain (**D**,**F**) and corpus striatum (**E**) demonstrating symmetrical involvement of predominantly grey matter structures. Images from dog 1 (**A**,**B**,**D**,**E**), dog 2 (**C**) and an unaffected littermate (**F**). Consistent bilateral symmetrical T2 hyperintensity of the deep cerebellar nuclei ((**A**–**C**); white arrowheads) is the most prominent finding. Similar bilaterally symmetrical hyperintensity is seen involving the tectum and dorsal tegmentum ((**D**); black arrowhead) and endopeduncular (medial) and lentiform nuclei ((**E**); white arrowheads) but not present in unaffected dog images (**F**). Sulci are prominent (**D**) compared to an unaffected age matched control (**F**), consistent with atrophy of cortical grey matter. Hyperintensity of deep cortical grey mater laminae is evident on T2-weighted and FLAIR images at all levels (white arrows) is not present on unaffected dog MR images (**F**).

3.4. Genetic Analysis

Both sexes are affected and in-depth pedigree analysis revealed the presence of a common male ancestor connecting the American and European families (Figure 1). As all parents of affected offspring show no clinical signs, it could be speculated that the observed disease phenotype follows autosomal monogenic recessive inheritance. Genome-wide association was performed using DNA samples from the seven affected dogs (Figure 1) and 28 phenotypically normal Saluki controls. After quality control, there were 108579 SNPs available for association. A single genome-wide significant association signal based on a $p_{Bonferroni}$ (0.006) on CFA 35 (chr35: g23,654,869; p_{raw} 5.27×10^{-8}) was identified (Figure 5). Furthermore, a 2.683 Mb region of homozygosity was identified in the seven affected dogs on CFA 35: 21,925,974–24,608,949 bp (CanFam3.1).

In order to identify a causative variant, initially paired-end whole-genome sequences of 2 affected puppies from two American litters (1052, 5813: Figure 1) and 98 unaffected controls from various breeds were investigated in the associated interval. There were 35,982 single-nucleotide variants (SNVs) and 16,832 insertion/deletion (indel) variants identified within the critical interval defined by homozygosity: CFA 35: 21,925,974–24,608,949 bp (CanFam3.1). There were 259 SNVs and 41 indels that segregated with the phenotype in the 100 animals. There were three coding variants identified: a synonymous variant (g.22,506,956G>A) in the *GPLD1* gene, and two protein-changing missense

variants, g.22,572,768G>A in the *ALDH5A1* gene, and g.23,908,560T>C in the putative *PTCHD3* gene. The synonymous variant in *GPLD1* was not investigated further.

The two missense variants in *ALDH5A1* (XM_014110599.2: c.866G>A; XP_013966074.2: p.(Gly288Asp)) and *PTCHD3* (ENSCAFT00000039442.3: c.1247T>C; ENSCAFP00000035301: p.(Iso416Thr)) were evaluated to identify whether the substitutions were potentially deleterious. There is a gap in the canine genome assembly that likely contains at least one additional exon of the *ALDH5A1* gene. Aligning the predicted canine ALDH5A1 protein sequence with the human protein sequence places the canine missense variant at amino acid 381 in human (NP_001071.1). Both PolyPhen2 [25] (probably damaging—1.0) and SIFT [26] (deleterious—0.0) predicted this amino acid substitution to be deleterious. It occurs in a well -conserved portion of the protein (Figure 6B). The missense variant in *PTCHD3* is not predicted to affect the protein based on PolyPhen2 [25] (benign-0.436) and SIFT [26] (tolerated-0.05). In addition, based on the known functions of these two proteins and the independent findings presented below, the *ALDH5A1* variant was the only one pursued further.

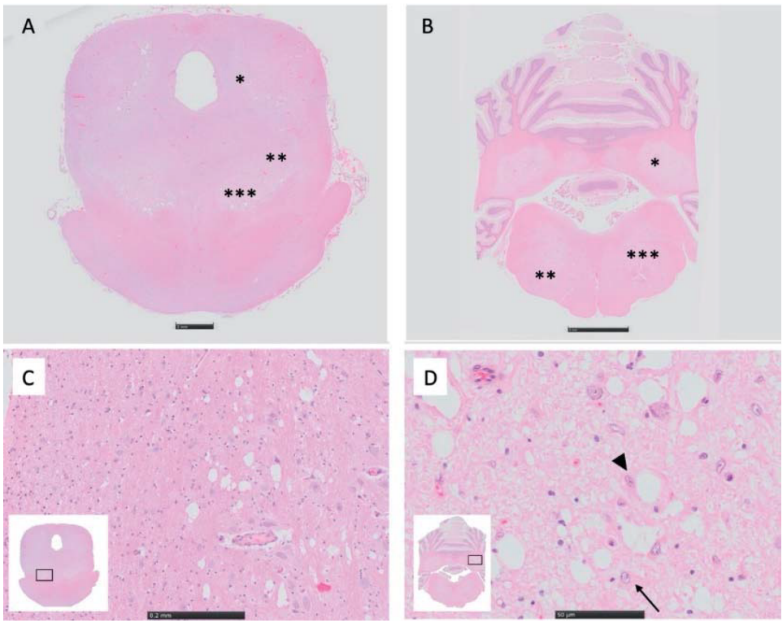

Figure 3. Histopathology of mesencephalon and brainstem from dog 1. (**A**) Bilaterally, the mesencephalic nuclei of cranial nerve V (*), the red nuclei (**), and the substantia nigra (***) exhibit decreased staining intensity (H&E). (**B**) Bilaterally, the deep cerebellar nuclei (*), the dorsal nuclei of the trapezoid body (**), and the reticular formations (***) exhibit decreased staining intensity (H&E). (**C**) Higher magnification of (**A**), inset. The substantia nigra shows prominent vacuolation of affected neurons. (**D**) Higher magnification of (**B**), inset. The interposital nucleus shows reactive astrocytes (arrow), some of which also contain prominent cytoplasmic vacuolation (arrowhead).

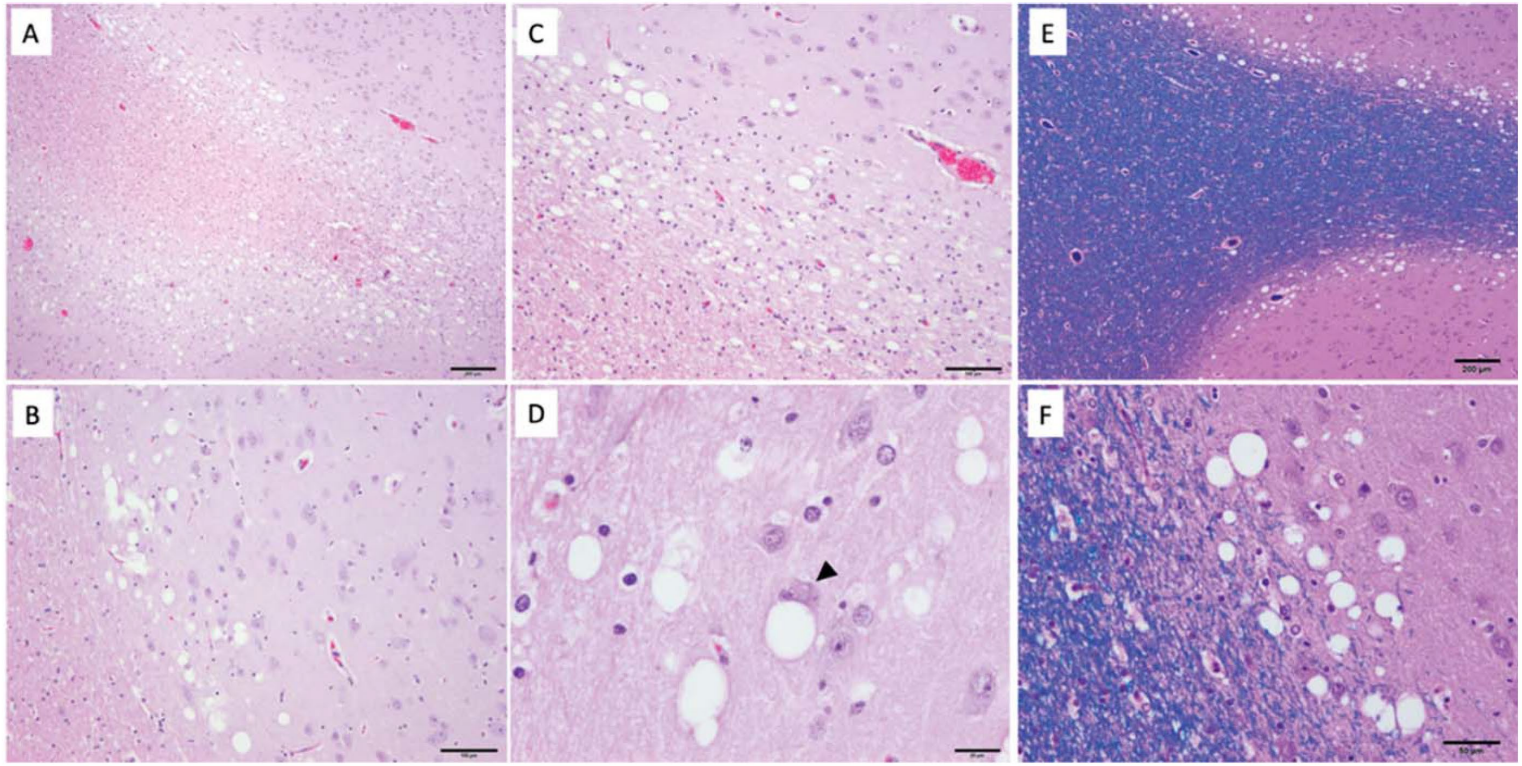

Figure 4. Histopathology of forebrain from dog 1. (**A**) Within the frontal cortex, the grey matter is predominantly affected by spongiotic change, with gliosis. (**B**) The spongiosis within the forebrain is most severe in the deep laminar cortex. (**C**) The caudate nucleus is also affected at the grey–white matter junction. (**D**) Affected neurons are characterized by enlarged, vacuolated cytoplasm with a peripheralized nucleus (arrowhead). Luxol fast blue (LFB) staining highlights the deep cortical nature of the vacuolation (**E**) within the cerebrum. (**F**) Vacuolation is discrete and often displaces cellular nuclei LFB staining.

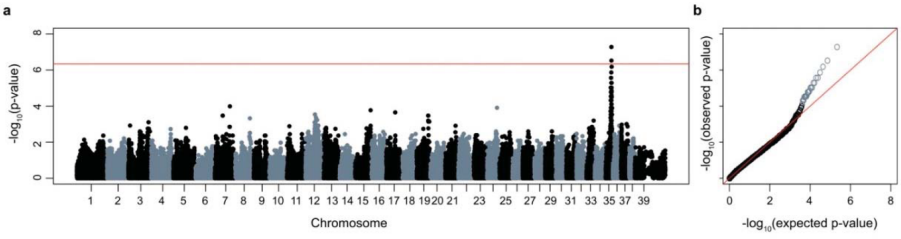

Figure 5. GWAS for SSADHD-affected Saluki dogs. (**a**) Manhattan plot showing –log10 of the raw p-values for each genotyped SNP by chromosome (*x*-axis). Genomic inflation was 1.25. Line denotes genome-wide significance based on Bonferroni-corrected p-values. (**b**) Q–Q plot of samples used in GWAS showing the –log10 of the expected versus the observed p-values. The SNPs on CFA35 are shown in light grey.

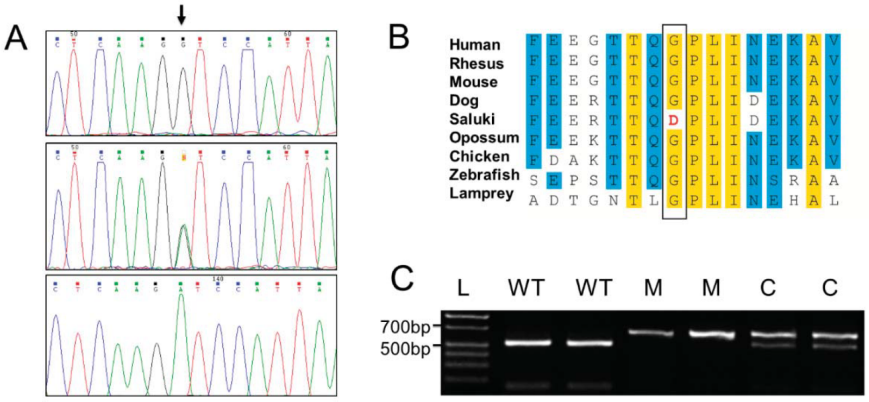

Figure 6. SSADHD-associated *ALDH5A1* missense variant in Saluki dogs. (**A**) The electropherograms from a normal dog (top panel), a heterozygous dog (middle panel) and a dog homozygous for the variant in *ALDH5A1* indicated by an arrow. (**B**) The amino acid alignment around the missense variant in ALDH5A1 (XP_013966074.2: p.(Gly288Asp)). Yellow boxes indicate 100% conservation across the species listed to the left and blue boxes indicate 75% conservation. The variant amino acid residue is boxed and the variant allele detected in affected Salukis is shown in red. (**C**) The PCR-RFLP genotyping assay for the *ALDH5A1* missense variant is shown. After PCR amplification, the products were digested with Sau96I. L is the DNA ladder, WT stands for wild type (542 bp), C for carrier and M for mutant (702 bp).

Independently, the genome of an affected puppy (dog 5) from the German litter was sequenced. No private variants were found in the *GPLD1* and *PTCHD3* genes. Furthermore, only one protein-changing missense variant (*ALDH5A1*: XM_014110599.2: c.866G>A) remained after filtering for homozygous private variants in the region of interest on chromosome 35 against the 589 control genomes from the Dog Biomedical Variant Database Consortium (DBVDC) variant catalog [27].

The *ALDH5A1* variant was confirmed by Sanger sequencing of genomic DNA (Figure 6A). In order to genotype the *ALDH5A1* missense variant, a PCR-RFLP genotyping assay was used (Figure 6C).

Genotyping of the *ALDH5A1* missense variant was performed in the seven affected dogs used for the GWAS, and the four available parents (Figure 1). The variant was homozygous in all cases and heterozygous in the parents. Siblings and other relatives, as well as unrelated Saluki dogs, were genotyped; 13 were heterozygous carriers and 48 were homozygous wild type. The segregation of the *ALDH5A* variant fits perfectly with the assumed monogenic recessive Mendelian inheritance within the studied

family. In-depth pedigree analysis revealed the presence of a common male ancestor connecting the American and European families (Figure 1). To experimentally determine whether the *ALDH5A* variant was a common canine variant, 228 dogs from various other breeds were genotyped and all were found to have the wild-type allele, which was also confirmed by the absence of the variant in 581 dogs from 125 breeds and eight wolves of the DBVDC cohort [27].

The presence of the variant in cDNA of an affected dog was verified by Sanger sequencing of RT-PCR product from liver of an affected Saluki compared to a control unaffected dog. There was no obvious difference in expression level between the case and the control. Quantitative evaluation was not possible since only one affected dog sample was available. Our results have been integrated in the Online Mendelian Inheritance in Animals (OMIA) database (https://omia.org/OMIA002250/9615/).

3.5. Targeted Metabolic Testing

Targeted quantitative organic acids were analyzed on urine, serum, CSF, and brain tissue (Table 2) in affected (n = 1 to 4) and control dogs (n = 2 to 4). Compared to control dogs, there were marked elevations in urine succinate semialdehyde (SSA) and urine 4,5-dihydroxyhexanoic acid (DHHA) but levels of gamma hydroxybutyrate (GHB) in the urine were normal. Serum GHB and serum DHHA from affected dogs were markedly elevated compared to controls. Serum SSA could not be measured in either affected or control dogs. In cisternal cerebrospinal fluid (CSF) and brain, SSA, GHB and DHHA were markedly elevated in the affected dog compared to controls, with the CSF GHB having the highest elevation (by a factor of at least 4800). Activity of succinate semialdehyde dehydrogenase (SSADH) was absent or markedly reduced to 0.18% of normal in the affected dogs compared to control dogs.

Table 2. Specific quantitative organic acids in urine, serum, CSF, and brain tissue in affected and control dogs (nd = not done).

Dog Number	Urine SSA, mmol/mol Creatinine	Urine GHB, mmol/mol Creatinine	Urine DHHA mmol/mol Creatinine	Serum GHB, µmol/L	Serum DHHA, µmol/L	CSF SSA, µmol/L	CSF GHB, µmol/L	CSF DHHA, µmol/L	Brain GHB, nmol/mg Brain	Brain DHHA, nmol/mg Brain	Brain SSA, nmol/mg Brain	Brain SSA Activity, pmol/min/mg Protein
1	9.23	1.06	5.85	6.59	0.45	69	>1500	43.3	2.43	0.22	0.23	10
2	nd	nd	nd	nd	nd	nd	nd	nd	2.93	0.28	0.18	0
5	38.7	nd	11.8	nd	0.41	nd	nd	nd	nd	nd	nd	nd
6	30.9	0.67	10.4	nd	0.61	nd	nd	nd	nd	nd	nd	nd
7	nd	nd	nd	nd	0.56	nd	nd	nd	nd	nd	nd	nd
Number of affected dogs	3	2	3	1	4	1	1	1	2	nd	2	2
Number of control dogs	4	4	4	4	4	2	3	3	4	nd	nd	4
median affected	30.9	0.87	10.4	n/a	0.51	n/a	n/a	n/a	2.68	0.25	0.21	5
range affected	9.23–38.7	0.67–1.06	5.85–11.8	n/a	0.41–0.61	n/a	n/a	n/a	2.43–2.93	0.22–0.28	0.18–0.23	0–10
median control range control	0.86 0.64–0.9	0.82 0.29–2.04	0.29 0.18–0.65	0.38 0.28–0.59	0.08 0.07–0.1	0.24 0.02–0.46	0.31 0.23–0.8	0.11 0.1–0.2	0.03 0–0.05	0 0–0	0.11 0.06–0.14	5587 4214–5942

4. Discussion

A pathogenic variant in the canine *ALDH5A1* gene associated with recessive SSADH deficiency, formerly known as status spongiosus in Saluki dogs (SSSD) [13]. Seven Saluki puppies from two continents had an onset of multifocal cranial neurological signs at 10 weeks of age or younger. Blood work was normal. On MRI of the brain, lesions were similar in all dogs, with bilateral and symmetrical lesions of the cerebrum, brainstem and cerebellum, predominantly affecting grey matter structures. Because this disorder occurred in purebred Saluki puppies from different environments with a consistent clinical phenotype and a normal extracranial work up, structural epilepsy from an inborn error of metabolism was considered the most likely etiology. Using genome-wide association, followed by whole-genome sequencing, a missense mutation in the *ALDH5A1* gene was identified as the presumed cause of status spongiosus in Saluki dogs, previously reported in Saluki puppies [13]. The mutation segregated completely in family members as an apparently fully penetrant monogenic recessive disorder.

The *ALDH5A1* gene encodes the mitochondrial enzyme succinic semialdehyde dehydrogenase (NAD+) (SSADH), which is involved in the catabolism of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) (Figure 7). GABA is the major inhibitory neurotransmitter in the central nervous system(CNS), where it is utilized in up to 30% of cerebral synapses [34]; it is also found in non-nervous-system tissue. GABA is synthesized from L-glutamate via glutamate decarboxylase (GAD). The first step in GABA metabolism is by GABA-transaminase (GABA-T) to form succinic semialdehyde (SSA). SSA is then oxidized by the mitochondrial protein succinic semialdehyde dehydrogenase (NAD+) (SSADH) to succinic acid, which then enters the tricarboxylic acid (TCA) cycle for energy generation (Figure 7). In people, recognized disorders of GABA synthesis are GAD deficiency, and recognized disorders of GABA degradation are GABA transaminase deficiency and succinic semialdehyde dehydrogenase (NAD+) (SSADH) deficiency.

The SSADH protein is expressed in the mammalian brain, as well as liver, pituitary, heart, ovary and kidney [35]. In people, the *ALDH5A1* gene is located on chromosome 6p22, and is 10 exons long extending over 38 kb [36]. Succinic semialdehyde dehydrogenase (NAD+) is an enzyme which is member of the aldehyde dehydrogenase family of proteins. In people with SSADH enzyme deficiency, SSA is not catabolized to succinic acid, and thus excess levels of SSA build up in tissues and fluids (Figure 7). Excess succinic semialdehyde is converted to 4-hydroxybutyric acid (GHB) by succinic semialdehyde reductase. Excess SSA may also interact with an intermediate in the pyruvate dehydrogenase complex to form 4,5-dihydroxyhexanoic acid (DHHA). People with a deficiency of SSADH have elevations of SSA, DHHA, and GHB in body fluids [35]. The activity of SSADH is reduced in people, and thus levels of SSA rise, with associated high levels of GHB and DHHA (Figure 7) [37]. The hallmark of SSADH deficiency in people is persistent and elevated levels of the GHB in urine, plasma and CSF [37]. The diagnosis is confirmed by molecular genetics by sequencing the *ALDH5A1* gene for pathogenic variants. There is no effective therapy [37].

In Saluki dogs with SSADH deficiency, levels of SSA and DHHA are elevated in urine, serum, CSF and brain, and GHB is elevated in serum, CSF and brain. Unlike in people, where GHB is elevated in urine, the level of GHB in urine in Saluki dogs with SSADH deficiency is normal. Since the activity of succinate semialdehyde dehydrogenase (NAD+) (SSADH) was absent or markedly reduced, along with elevated levels of SSA, DHHA and GHB, we believe that the previously described central nervous system status spongiosus in Saluki dogs should be more appropriately termed succinic semialdehyde dehydrogenase deficiency (SSADHD).

In people, SSADH deficiency is a rare autosomal recessive neurological disorder caused by a mutation in the *ALDH5A1* gene, reported in 1981 (OMIM 271980) [38]. There are 44 unique mutations in the *ALDH5A1* gene, which occur in exons 1–10; there are no other mutations in genes other than *ALDH5A1* associated with SSADH deficiency in people [36]. The clinical features in people include developmental delay, hypotonia, intellectual disability, ataxia, seizures, hyperkinetic behavior, aggression and sleep disturbances. Approximately 50% of patients have seizures, 45% have neuropsychiatric problems such

as sleep disturbances, and many patients also have behavioral abnormalities [34], which all worsen with age [39]. The encephalopathy is considered non-progressive and has wide phenotypic heterogeneity from mild to severe. Symptoms are first noted at a mean of 11 months (range 0–44 months of age), with a mean age at diagnosis of 6.6 years (range of 0 to 25 years) [40]. On MR imaging in people with SSADHD, MR images may be normal, or there may be hyperintensities on T2-weighted imaging in the globus pallidus, cerebellar dentate nuclei and brainstem [41]. A small percentage of people are reported with cerebral white matter hyperintensity on MR imaging as well [41]. There is one single case report of the pathology of SSADH deficiency in a young adult, where there was discoloration of the globus pallidi, congestion of the leptomeninges and scar tissue in the cerebral cortex [42].

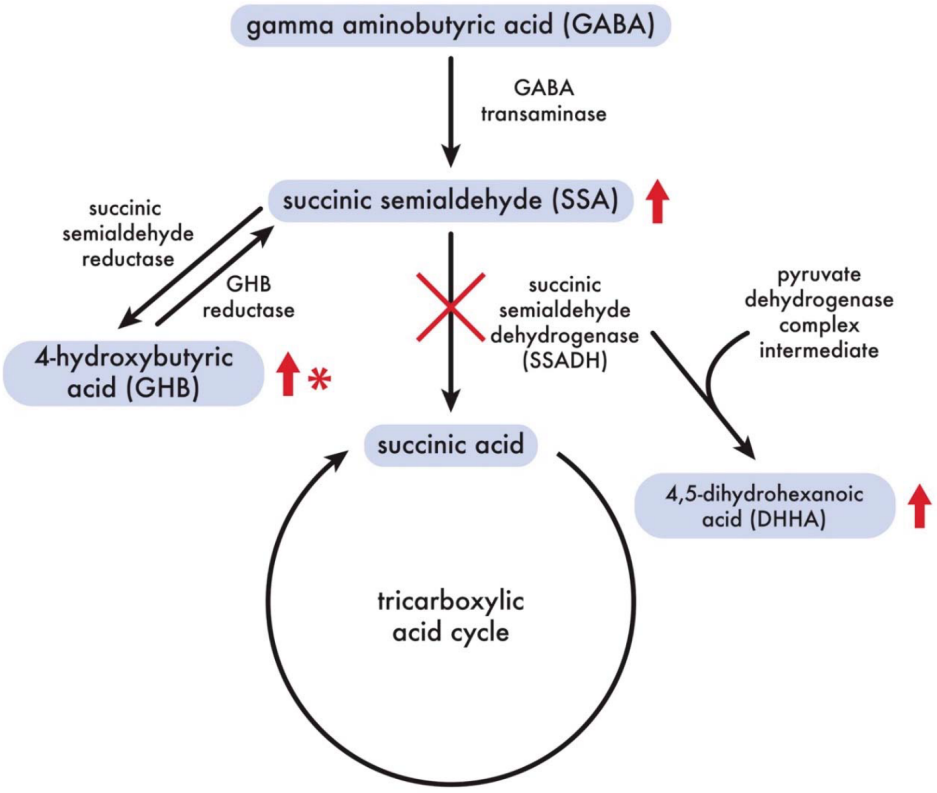

Figure 7. GABA catabolism pathway. In Saluki dogs with SSADH deficiency, levels of SSA and DHHA are elevated in urine, serum, CSF and brain, and GHB is elevated in serum, CSF and brain (red arrows) as in people with SSADH deficiency. Unlike in people, where GHB is elevated in urine (red arrow), the level of GHB in urine (red *) in Saluki dogs with SSADH deficiency is normal.

The pathophysiology of SSADH deficiency in people is complex. The disease is thought to be caused primarily by the elevation of GHB in the brain, particularly during neurodevelopment or due to the imbalance of neurotransmitters [43] but also potentially from oxidative stress in the brain [44]. GHB is a neuromodulator with a wide array of pharmacological effects [45]. It was initially produced as an injectable anesthetic agent but is now longer utilized for this purpose due to adverse effects but is prescribed to treat cataplexy in narcolepsy/cataplexy, opiate dependency and alcoholism, and is a drug of abuse, where its street names include Grievous Bodily Harm, Liquid Ecstasy and Soap [46]. GHB may cause anxiolytic, hypnotic and euphoric effects as well as short-term memory loss and CNS depression, causing sedation [44,47]; intoxication may cause bradycardia, myoclonus and

seizures, hypoventilation, coma and death from respiratory depression [33]. GHB is a monocarboxylate that is primarily cleared from the plasma via metabolism in a dose-dependent fashion through the Kreb's cycle. Renal clearance of GHB is minor and non-linear, due to the carrier-mediated saturable renal reabsorption of GHB through the proximal tubules via sodium-dependent and pH-dependent monocarboxylate transporters. Methods to increase the renal excretion of GHB have been investigated in animals as part of a treatment plan for people with GHB intoxication, such as with the intravenous administration of L-Lactate which increases the renal excretion of GHB [48,49].

As there are no effective specific treatments for SSADH deficiency in people, treatment is currently aimed at managing the clinical signs of seizures and neurobehavioral disturbances [34]. Broad-spectrum anticonvulsants are generally utilized, avoiding valproate which inhibits SSADH which may worsen GHB accumulation and clinical signs [50]. However, there are many therapeutic options under investigation which include pharmacological (e.g., targeting neurotransmitter receptors), enzyme-replacement therapy, gene therapy and treatment with pharmacological chaperones [51,52].

In animals, SSADH has been produced in knockout mice; clinical signs are progressive ataxia, failure to thrive, seizure and death at a young age [53,54], and thus they are utilized as a model for the severe and poor survival phenotype in humans. There are no reports of the MR imaging features of SSADH-knockout mice [51] and, on pathology, there are no reported abnormalities on routine hematoxylin and eosin staining, but detailed neuropathological examinations have not been performed [55]. There is a single case report of a dog with suspected SSADH deficiency, which had a progressive encephalopathy with profound and persistent lactic acidosis, elevated urine GHB and a 30% reduced activity of SSADH measured in cultured lymphoblasts compared to normal dogs. Intracranial MR imaging of the brain was not performed; on histopathology of the brain, there was a spongiform change in the cerebral cortex [56].

Saluki dogs have a more severe phenotype of SSADHD than people, but with very similar clinical signs of seizures, abnormal behavior, abnormalities of sleep and multifocal brain disease on neurological examination. Like people [42], affected Salukis have multifocal abnormalities in the brain on MR imaging, where affected dogs have identical bilaterally symmetrical T2-weighted hyperintensities in the same anatomical areas as people such as the basal nuclei, deep cerebellar nuclei, and brain stem. In people with SSADH deficiency, these abnormalities have considerable consistency, and are almost universal. However, in some people, there are reports of non-specific hyperintensities in subcortical white matter and the substantia nigra in the brainstem, as well as cerebellar atrophy [57,58]. These abnormalities are not considered specific for SSADHD, but considered to be imaging characteristics of cytotoxic edema, secondary to oxidative stress from the underlying SSADHD [43]. Differently to what is noted in people, dogs have hyperintensity of the deep cortical laminae of the grey matter of the cerebral cortex and atrophy of the cerebral cortex. On histopathology in dogs with SSADH deficiency, there is bilaterally symmetric multifocal spongiform change in the brainstem, deep cerebellar nuclei, but also severe in the thalamic nuclei and deep cortical grey matter. The brain lesions in affected Salukis on MRI (two dogs) and pathology (three dogs (same two dogs that had MRI plus one other) were identical. There is only one case report of the pathology of SSADHD in a person where the histopathology was not described and thus the comparative pathology between dogs and people is not possible [42].

Unlike in people with SSADHD, where GHB is elevated in the urine and therefore is an excellent biomarker, GHB is not elevated in the urine of affected dogs. GHB is, however, elevated in the serum, CSF and brain tissue of affected dogs. Since GHB is extensively reabsorbed from the proximal tubules in the kidney [59], it is possible that species differences in handling GHB resulted in extensive reabsorption of GHB and in a normal level of GHB in the affected dog's urine. It is also plausible that since urine GHB was evaluated in only two affected dogs, that evaluation of additional SSADHD-affected dogs may have yielded elevated levels of urine GHB. There may also have been loss of GHB in the samples during shipping to the lab or storage, as GHB is a volatile compound [40] and a reduction in GHB levels is reported with storage, which varied between 10% loss at just 3 days and in excess of 20% after

4 weeks of storage [60]. If the activity of the D-2-hydroxyglutarate dehydrogenase in liver (non-cofactor enzyme converting GHB to SSA and ketoglutarate to D-2-hydroxyglutarate) was very active in the liver of Salukis, this could be a plausible explanation for why GHB was not elevated in the urine; however, D-2-HG was not measured in affected or control dogs. Otherwise, aside from serum SSA that we were not able to measure in dogs with SSADHD, dogs are similar to human patients in regards to elevated body fluid levels of SSA, DHHA and GHB, and zero to low levels of brain SSA activity.

5. Conclusions

Saluki dogs with SSADHD have a disease phenotype resembling SSADHD in people, although it appears to be more severe clinically. On the other hand, it appears to be less severe than the phenotype described in in knockout mice, which is lethal [58]. Furthermore, GHB may not be an acceptable biomarker for the disease in dogs; alternatively, urine SSA or GHB in serum may be more appropriate biomarkers. Compared to mice models of human disease, dog models have naturally occurring disease, are more similar to humans in regards to size, and have more longevity than mice. Dogs are proven and valuable models of human disease, particularly in the field of lysosomal storage diseases [6]. This first *ALDH5A1*-related large animal model for SSADHD may provide an opportunity for evaluation of potential therapeutics for this rare orphan disease in people. Dogs may be a more appropriate disease model than the murine SSADH model, as dogs appear to have a more similar disease phenotype and similar MR imaging features to people. The identification of the pathogenic *ALDH5A1* variant will allow the screening of carriers to avoid producing further affected puppies and thereby contribute to maintaining breed health.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/9/1033/s1. Video S1: A Saluki puppy (dog number 5) showing an episode of abnormal vocalization, Video S2: A Saluki puppy (dog number 1) with bilateral thoracic limb hypermetria.

Author Contributions: Conceptualization, K.M.V., D.L.B., E.S., P.L.P., K.M.G., and C.D.; methodology, D.L.B. and E.S.; software, D.L.B.; validation, D.L.B., E.S., and Z.Z.; formal analysis, D.L.B., E.S., B.R., and A.L.; investigation, K.M.V., D.L.B., D.D.C., E.A.B., M.A., K.D.W., G.D.S., F.G., M.R.B., K.M., T.A.M., M.R.B., A.L., V.J., D.L.B., T.J.V.W., and F.K.; resources, D.L.B., E.S., and C.D.; data curation, D.L.B. and E.S.; writing—original draft preparation, K.M.V. and D.L.B.; writing—review and editing, D.L.B., P.J.D., C.D., K.D.W., E.S., P.L.P., K.M.G., G.D.S., F.K.,T.A.M., D.O., and A.L.; visualization, K.M.V., D.L.B., and A.L; supervision, D.L.B., C.T.B., and C.D.; project administration, K.M.V.; funding acquisition, D.L.B. and K.M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Center of Companion Animal Health, School of Veterinary Medicine, University of California Davis, grant numbers 2015-11F and 2016-18F and the Maxine Adler Endowed Chair Funds.

Acknowledgments: The authors acknowledge the support of Sharron Kinney, Lorrie Boldrick, Ember, Encore, Khrome, Kara, Yamal, Yari, Yanam and Kelly Kohen for this study. We acknowledge the Next-Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern for performing the whole-genome sequencing experiments and for providing the computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. El-Hattab, A.W. Inborn errors of metabolism. *Clin. Perinatol.* **2015**, *42*, 413–439. [CrossRef]
2. Seijo-Martinez, M.; Navarro, C.; Castro del Rio, M.; Vila, O.; Puig, M.; Ribes, A.; Butron, M. L-2-hydroxyglutaric aciduria: Clinical, neuroimaging, and neuropathological findings. *Arch. Neurol.* **2005**, *62*, 666–670. [CrossRef] [PubMed]
3. Jolly, R.D.; Walkley, S.U. Lysosomal storage diseases of animals: An essay in comparative pathology. *Vet. Pathol.* **1997**, *34*, 527–548. [CrossRef]
4. Mansour, T.A.; Woolard, K.D.; Vernau, K.L.; Ancona, D.M.; Thomasy, S.M.; Sebbag, L.; Moore, B.A.; Knipe, M.F.; Seada, H.A.; Cowan, T.M.; et al. Whole genome sequencing for mutation discovery in a single case of lysosomal storage disease (MPS type 1) in the dog. *Sci. Rep.* **2020**, *10*, 6558. [CrossRef] [PubMed]

5. Lucot, K.L.; Dickinson, P.J.; Finno, C.J.; Mansour, T.A.; Letko, A.; Minor, K.M.; Mickelson, J.R.; Drogemuller, C.; Brown, C.T.; Bannasch, D.L. A Missense Mutation in the Vacuolar Protein Sorting 11 (VPS11) Gene Is Associated with Neuroaxonal Dystrophy in Rottweiler Dogs. *G3 Genes Genomes Genet.* **2018**, *8*, 2773–2780. [CrossRef]
6. Haskins, M.E.; Giger, U.; Patterson, D.F. Animal models of lysosomal storage diseases: Their development and clinical relevance. In *Fabry Disease: Perspectives from 5 Years of FOS*; Mehta, A., Beck, M., Sunder-Plassmann, G., Eds.; Oxford PharmaGenesis: Oxford, UK, 2006.
7. Kamboj, M. Clinical approach to the diagnoses of inborn errors of metabolism. *Pediatr. Clin. N. Am.* **2008**, *55*, 1113–1127. [CrossRef]
8. Saudubray, J.M.; Garcia-Cazorla, A. An overview of inborn errors of metabolism affecting the brain: From neurodevelopment to neurodegenerative disorders. *Dialogues Clin. Neurosci.* **2018**, *20*, 301–325. [CrossRef]
9. Sewell, A.C.; Haskins, M.E.; Giger, U. Inherited metabolic disease in companion animals: Searching for nature's mistakes. *Vet. J.* **2007**, *174*, 252–259. [CrossRef]
10. Lavely, J.A. Pediatric seizure disorders in dogs and cats. *Vet. Clin. N. Am. Small Anim. Pract.* **2014**, *44*, 275–301. [CrossRef]
11. Berendt, M.; Farquhar, R.G.; Mandigers, P.J.; Pakozdy, A.; Bhatti, S.F.; De Risio, L.; Fischer, A.; Long, S.; Matiasek, K.; Munana, K.; et al. International veterinary epilepsy task force consensus report on epilepsy definition, classification and terminology in companion animals. *BMC Vet. Res.* **2015**, *11*, 182. [CrossRef]
12. De Risio, L.; Bhatti, S.; Munana, K.; Penderis, J.; Stein, V.; Tipold, A.; Berendt, M.; Farqhuar, R.; Fischer, A.; Long, S.; et al. International veterinary epilepsy task force consensus proposal: Diagnostic approach to epilepsy in dogs. *BMC Vet. Res.* **2015**, *11*, 148. [CrossRef]
13. Luttgen, P.; Storts, R. Central Nervous system spongiosus of Saluki Dogs. In Proceedings of the American College of Veterinary Internal Medicine Forum, San Diego, CA, USA, 21–24 May 1987.
14. Summers, B.; Cummings, J.F.; de Lahunta, A. *Veterinary Neuropathology*; Mosby: St. Louis, MO, USA, 2005.
15. Kanekar, S.; Gustas, C. Metabolic disorders of the brain: Part I. *Semin Ultrasound CT MR* **2011**, *32*, 590–614. [CrossRef]
16. Markovich, J.E.; Heinze, C.R.; Freeman, L.M. Thiamine deficiency in dogs and cats. *J. Am. Vet. Med. Assoc.* **2013**, *243*, 649–656. [CrossRef]
17. Brauer, C.; Jambroszyk, M.; Tipold, A. Metabolic and toxic causes of canine seizure disorders: A retrospective study of 96 cases. *Vet. J.* **2011**, *187*, 272–275. [CrossRef]
18. Peterson, M.E. Bromethalin. *Top Companion Anim. Med.* **2013**, *28*, 21–23. [CrossRef]
19. Tauro, A.; Beltran, E.; Cherubini, G.B.; Coelho, A.T.; Wessmann, A.; Driver, C.J.; Rusbridge, C.J. Metronidazole-induced neurotoxicity in 26 dogs. *Aust. Vet. J.* **2018**, *96*, 495–501. [CrossRef]
20. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J.; et al. PLINK(1.7): A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
21. Kierczak, M.; Jablonska, J.; Forsberg, S.K.; Bianchi, M.; Tengvall, K.; Pettersson, M.; Scholz, V.; Meadows, J.R.; Jern, P.; Carlborg, O.; et al. Cgmisc: Enhanced genome-wide association analyses and visualization. *Bioinformatics* **2015**, *31*, 3830–3831. [CrossRef]
22. Wickham. *Elegant Graphics for Data Analysis*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2009.
23. Mansour, T.A.; Lucot, K.; Konopelski, S.E.; Dickinson, P.J.; Sturges, B.K.; Vernau, K.L.; Choi, S.; Stern, J.A.; Thomasy, S.M.; Doring, S.; et al. Whole genome variant association across 100 dogs identifies a frame shift mutation in DISHEVELLED 2 which contributes to Robinow-like syndrome in Bulldogs and related screw tail dog breeds. *PLoS Genet.* **2018**, *14*, e1007850. [CrossRef]
24. McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G.R.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. *Genome Biol.* **2016**, *17*, 122. [CrossRef]
25. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248–249. [CrossRef]
26. Choi, Y.; Chan, A.P. PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **2015**, *31*, 2745–2747. [CrossRef]

27. Jagannathan, V.; Drögemüller, C.; Leeb, T.; Aguirre, G.; André, C.; Bannasch, D.; Becker, D.; Davis, B.; Ekenstedt, K.; Faller, K.; et al. A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. *Anim. Genet.* **2019**, *50*, 695–704. [CrossRef]
28. Minor, K.M.; Letko, A.; Becker, D.; Drogemuller, M.; Mandigers, P.J.J.; Bellekom, S.R.; Leegwater, P.A.J.; Stassen, Q.E.M.; Putschbach, K.; Fischer, A.; et al. Canine NAPEPLD-associated models of human myelin disorders. *Sci. Rep.* **2018**, *8*, 5818. [CrossRef]
29. Rozen, S.; Skaletsky, H. *Methods in Molecular Biology: Bioinformatics Methods and Protocols*; Primer3 on the WWW for general users and for biologist programmers; Humana Press: Totowa, NJ, USA, 2000; Volume 132.
30. Brinkhof, B.; Spee, B.; Rothuizen, J.; Penning, L.C. Development and evaluation of canine reference genes for accurate quantification of gene expression. *Anal. Biochem.* **2006**, *356*, 36–43. [CrossRef]
31. Gibson, K.M.; Aramaki, S.; Sweetman, L.; Nyhan, W.L.; DeVivo, D.C.; Hodson, A.K.; Jakobs, C. Stable isotope dilution analysis of 4-hydroxybutyric acid: An accurate method for quantification in physiological fluids and the prenatal diagnosis of 4-hydroxybutyric aciduria. *BioMed. Environ. Mass Spectrom.* **1990**, *19*, 89–93. [CrossRef]
32. Struys, E.A.; Jansen, E.E.; Gibson, K.M.; Jakobs, C. Determination of the GABA analogue succinic semialdehyde in urine and cerebrospinal fluid by dinitrophenylhydrazine derivatization and liquid chromatography-tandem mass spectrometry: Application to SSADH deficiency. *J. Inherit. Metab. Dis.* **2005**, *28*, 913–920. [CrossRef]
33. Gibson, K.M.; Lee, C.F.; Chambliss, K.L.; Kamali, V.; Francois, B.; Jaeken, J.; Jakobs, C. 4-Hydroxybutyric aciduria: Application of a fluorometric assay to the determination of succinic semialdehyde dehydrogenase activity in extracts of cultured human lymphoblasts. *Clin. Chim. Acta* **1991**, *196*, 219–221. [CrossRef]
34. Pearl, P.L.; Parviz, M.; Vogel, K.; Schreiber, J.; Theodore, W.H.; Gibson, K.M. Inherited disorders of gamma-aminobutyric acid metabolism and advances in ALDH5A1 mutation identification. *Dev. Med. Child Neurol.* **2015**, *57*, 611–617. [CrossRef]
35. Kim, K.J.; Pearl, P.L.; Jensen, K.; Snead, O.C.; Malaspina, P.; Jakobs, C.; Gibson, K.M. Succinic semialdehyde dehydrogenase: Biochemical-molecular-clinical disease mechanisms, redox regulation, and functional significance. *Antioxid. Redox Signal.* **2011**, *15*, 691–718. [CrossRef]
36. Liu, N.; Kong, X.D.; Kan, Q.C.; Shi, H.R.; Wu, Q.H.; Zhuo, Z.H.; Bai, Q.L.; Jiang, M. Mutation analysis and prenatal diagnosis in a Chinese family with succinic semialdehyde dehydrogenase and a systematic review of the literature of reported ALDH5A1 mutations. *J. Perinat. Med.* **2016**, *44*, 441–451. [CrossRef]
37. Parviz, M.; Vogel, K.; Gibson, K.M.; Pearl, P.L. Disorders of GABA metabolism: SSADH and GABA-transaminase deficiencies. *J. Pediatr. Epilepsy* **2014**, *3*, 217–227. [CrossRef]
38. Jakobs, C.; Bojasch, M.; Mönch, E.; Rating, D.; Siemes, H.; Hanefeld, F. Urinary excretion of gamma-hydroxybutyric acid in a patient with neurological abnormalities. The probability of a new inborn error of metabolism. *Clin. Chim. Acta* **1981**, *111*, 169–178. [CrossRef]
39. DiBacco, M.L.; Roullet, J.B.; Kapur, K.; Brown, M.N.; Walters, D.C.; Gibson, K.M.; Pearl, P.L. Age-related phenotype and biomarker changes in SSADH deficiency. *Ann. Clin. Transl. Neurol.* **2019**, *6*, 114–120. [CrossRef]
40. Pearl, P.L.; Gibson, K.M.; Acosta, M.T.; Vezina, L.G.; Theodore, W.H.; Rogawski, M.A.; Novotny, E.J.; Gropman, A.; Conry, J.A.; Berry, G.T.; et al. Clinical spectrum of succinic semialdehyde dehydrogenase deficiency. *Neurology* **2003**, *60*, 1413–1417. [CrossRef]
41. Gibson, K.M.; Christensen, E.; Jakobs, C.; Fowler, B.; Clarke, M.A.; Hammersen, G.; Raab, K.; Kobori, J.; Moosa, A.; Vollmer, B.; et al. The clinical phenotype of succinic semialdehyde dehydrogenase deficiency (4-hydroxybutyric aciduria): Case reports of 23 new patients. *Pediatrics* **1997**, *99*, 567–574. [CrossRef]
42. Knerr, I.; Gibson, K.M.; Murdoch, G.; Salomons, G.S.; Jakobs, C.; Combs, S.; Pearl, P.L. Neuropathology in succinic semialdehyde dehydrogenase deficiency. *Pediatr. Neurol.* **2010**, *42*, 255–258. [CrossRef]
43. Malaspina, P.; Roullet, J.B.; Pearl, P.L.; Ainslie, G.R.; Vogel, K.R.; Gibson, K.M. Succinic semialdehyde dehydrogenase deficiency (SSADHD): Pathophysiological complexity and multifactorial trait associations in a rare monogenic disorder of GABA metabolism. *Neurochem. Int.* **2016**, *99*, 72–84. [CrossRef]
44. Sgaravatti, A.M.; Sgarbi, M.B.; Testa, C.G.; Durigon, K.; Pederzolli, C.D.; Prestes, C.C.; Wyse, A.T.; Wannmacher, C.M.; Wajner, M.; Dutra-Filho, C.S. Gamma-hydroxybutyric acid induces oxidative stress in cerebral cortex of young rats. *Neurochem. Int.* **2007**, *50*, 564–570. [CrossRef]
45. Struys, E.A.; Verhoeven, N.M.; Jansen, E.E.; Ten Brink, H.J.; Gupta, M.; Burlingame, T.G.; Quang, L.S.; Maher, T.; Rinaldo, P.; Snead, O.C.; et al. Metabolism of gamma-hydroxybutyrate to d-2-hydroxyglutarate in mammals: Further evidence for d-2-hydroxyglutarate transhydrogenase. *Metabolism* **2006**, *55*, 353–358. [CrossRef]

46. Teter, C.J.; Guthrie, S.K. A comprehensive review of MDMA and GHB: Two common club drugs. *Pharmacotherapy* **2001**, *21*, 1486–1513. [CrossRef]
47. Kamal, R.M.; van Noorden, M.S.; Franzek, E.; Dijkstra, B.A.; Loonen, A.J.; De Jong, C.A. The Neurobiological Mechanisms of Gamma-Hydroxybutyrate Dependence and Withdrawal and Their Clinical Relevance: A Review. *Neuropsychobiology* **2016**, *73*, 65–80. [CrossRef]
48. Morris, M.E.; Hu, K.; Wang, Q. Renal clearance of gamma-hydroxybutyric acid in rats: Increasing renal elimination as a detoxification strategy. *J. Pharmacol. Exp. Ther.* **2005**, *313*, 1194–1202. [CrossRef]
49. Felmlee, M.A.; Wang, Q.; Cui, D.; Roiko, S.A.; Morris, M.E. Mechanistic toxicokinetic model for gamma-hydroxybutyric acid: Inhibition of active renal reabsorption as a potential therapeutic strategy. *AAPS J.* **2010**, *12*, 407–416. [CrossRef]
50. Shinka, T.; Ohfu, M.; Hirose, S.; Kuhara, T. Effect of valproic acid on the urinary metabolic profile of a patient with succinic semialdehyde dehydrogenase deficiency. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2003**, *792*, 99–106. [CrossRef]
51. Didiasova, M.; Banning, A.; Brennenstuhl, H.; Jung-Klawitter, S.; Cinquemani, C.; Opladen, T.; Tikkanen, R. Succinic Semialdehyde Dehydrogenase Deficiency: An Update. *Cells* **2020**, *9*, 477. [CrossRef]
52. Vogel, K.R.; Ainslie, G.R.; Walters, D.C.; McConnell, A.; Dhamne, S.C.; Rotenberg, A.; Roullet, J.B.; Gibson, K.M. Succinic semialdehyde dehydrogenase deficiency, a disorder of GABA metabolism: An update on pharmacological and enzyme-replacement therapeutic strategies. *J. Inherit. Metab. Dis.* **2018**, *41*, 699–708. [CrossRef] [PubMed]
53. Drasbek, K.R.; Vardya, I.; Delenclos, M.; Gibson, K.M.; Jensen, K. SSADH deficiency leads to elevated extracellular GABA levels and increased GABAergic neurotransmission in the mouse cerebral cortex. *J. Inherit. Metab. Dis.* **2008**, *31*, 662–668. [CrossRef]
54. Gibson, K.M.; Jakobs, C.; Pearl, P.L.; Snead, O.C. Murine succinate semialdehyde dehydrogenase (SSADH) deficiency, a heritable disorder of GABA metabolism with epileptic phenotype. *IUBMB Life* **2005**, *57*, 639–644. [CrossRef]
55. Gupta, M.; Hogema, B.M.; Grompe, M.; Bottiglieri, T.G.; Concas, A.; Biggio, G.; Sogliano, C.; Rigamonti, A.E.; Pearl, P.L.; Snead, O.C., 3rd; et al. Murine succinate semialdehyde dehydrogenase deficiency. *Ann. Neurol.* **2003**, *54* (Suppl. 6), S81–S90. [CrossRef]
56. Kelmer, E.; Gibson, K.M.; Jakobs, C.; Struys, E.; Shelton, G.D.; Aroch, I.; O'Brien, D.P. Severe Lactic Acidosis Associated with a Suspected Succinic Semialdehyde Dehydrogenas (SSADH) Deficiency in a Young Chihuahua Dog. *Isr. J. Vet. Med.* **2018**, *73*, 43–48.
57. Pearl, P.L.; Wiwattanadittakul, N.; Roullet, J.B.; Gibson, K.M. Succinic Semialdehyde Dehydrogenase Deficiency. In *GeneReviews (R)*; Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Stephens, K., Amemiya, A., Eds.; University of Washington: Seattle, WA, USA, 1993.
58. Pearl, P.L.; Gibson, K.M.; Cortez, M.A.; Wu, Y.; Carter Snead, O., 3rd; Knerr, I.; Forester, K.; Pettiford, J.M.; Jakobs, C.; Theodore, W.H. Succinic semialdehyde dehydrogenase deficiency: Lessons from mice and men. *J. Inherit. Metab. Dis.* **2009**, *32*, 343–352. [CrossRef] [PubMed]
59. Wang, Q.; Lu, Y.; Morris, M.E. Monocarboxylate transporter (MCT) mediates the transport of gamma-hydroxybutyrate in human kidney HK-2 cells. *Pharm. Res.* **2007**, *24*, 1067–1078. [CrossRef]
60. Busardo, F.P.; Zaami, S.; Baglio, G.; Indorato, F.; Montana, A.; Giarratana, N.; Kyriakou, C.; Marinelli, E.; Romano, G. Assessment of the stability of exogenous gamma hydroxybutyric acid (GHB) in stored blood and urine specimens. *Eur. Rev. Med. Pharmacol. Sci.* **2015**, *19*, 4187–4194.

SLC19A3 loss-of-function variant in Yorkshire terriers with Leigh-like subacute necrotizing encephalopathy

Journal:	Genes
Manuscript status:	published
Contributions:	Formal analysis, Visualization, Writing – original draft, Writing – review & editing

genes

MDPI

Case Report

SLC19A3 Loss-of-Function Variant in Yorkshire Terriers with Leigh-Like Subacute Necrotizing Encephalopathy

Michaela Drögemüller [1,†], Anna Letko [1,†], Kaspar Matiasek [2,†], Vidhya Jagannathan [1], Daniele Corlazzoli [3], Marco Rosati [2], Konrad Jurina [4], Susanne Medl [5], Thomas Gödde [6], Stefan Rupp [7], Andrea Fischer [8], Alejandro Luján Feliu-Pascual [9] and Cord Drögemüller [1,*]

1 Institute of Genetics, Vetsuisse Faculty, University of Bern, 3012 Bern, Switzerland; michaela.droegemueller@vetsuisse.unibe.ch (M.D.); anna.letko@vetsuisse.unibe.ch (A.L.); vidhya.jagannathan@vetsuisse.unibe.ch (V.J.)
2 Section of Clinical & Comparative Neuropathology, Centre for Clinical Veterinary Medicine, Ludwig Maximilians Universität Munich, 80539 Munich, Germany; kaspar.matiasek@neuropathologie.de (K.M.); marco.rosati@neuropathologie.de (M.R.)
3 Neurology & Neurosurgery Unit, Policlinico Veterinario Roma Sud, 00173 Roma, Italy; daniele.corlazzoli@me.com
4 Small Animal Hospital, Tierklinik Haar, 85540 Haar, Germany; datenschutz.haar@anicura.de
5 Small Animal Hospital, Anicura Kleintierklinik Babenhausen, 87727 Babenhausen, Germany; smedl@tierklinik-medl.de
6 Small Animal Referral Practice, Veterinary Health Centre, 83451 Piding, Germany; t.goedde@tierarzt-piding.com
7 Small Animal Hospital, Tierklinik Hofheim, 65719 Hofheim, Germany; s.rupp@tierklinik-hofheim.de
8 Section of Neurology, Centre for Clinical Veterinary Medicine, Ludwig-Maximilians-Universität, 80539 Munich, Germany; a.fischer@medizinische-kleintierklinik.de
9 Aúna Especialidades Veterinarias, 46980 Valencia, Spain; alf@aunaespecialidadesveterinarias.es
* Correspondence: cord.droegemueller@vetsuisse.unibe.ch
† These authors contributed equally to the work.

Received: 17 September 2020; Accepted: 15 October 2020; Published: 16 October 2020

Abstract: Sporadic occurrence of juvenile-onset necrotizing encephalopathy (SNE) has been previously reported in Yorkshire terriers. However, so far, no causative genetic variant has been found for this breed-specific form of suspected mitochondrial encephalomyopathy. Affected dogs showed gait abnormalities, central visual defects, and/or seizures. Histopathological analysis revealed the presence of major characteristics of human Leigh syndrome and SNE in Alaskan huskies. The aim of this study was to characterize the genetic etiology of SNE-affected purebred Yorkshire terriers. After SNP genotyping and subsequent homozygosity mapping, we identified a single loss-of-function variant by whole-genome sequencing in the canine *SLC19A3* gene situated in a 1.7 Mb region of homozygosity on chromosome 25. All ten cases were homozygous carriers of a mutant allele, an indel variant in exon 2, that is predicted to lead to a frameshift and to truncate about 86% of the wild type coding sequence. This study reports a most likely pathogenic variant in *SLC19A3* causing a form of SNE in Yorkshire terriers and enables selection against this fatal neurodegenerative recessive disorder. This is the second report of a pathogenic alteration of the *SLC19A3* gene in dogs with SNE.

Keywords: *Canis familiaris*; whole-genome sequencing; rare disease; precision medicine; neurometabolic disorder

1. Introduction

Subacute necrotizing encephalomyelopathy (SNE), also termed Leigh syndrome (LS; OMIM 256000) represents a devastating neurodegenerative disorder in people, characterized by a wide variety of clinical signs, ranging from severe neurologic problems to a near absence of abnormalities with the central nervous system most frequently affected [1]. Originally, Archibald Denis Leigh, a British neuropsychiatrist described the condition in 1951 [2]. SNE is characterized by focal and bilaterally symmetrical, necrotic lesions involving the thalamus, brainstem, and posterior columns of the spinal cord [3]. In SNE, various mutations in mitochondrial respiratory chain complexes lead to the disruption of ATP synthesis resulting in the characteristic pathology of SNE [4]. Mitochondrial encephalomyelopathies, such as SNE or LS, represent rare inherited neurometabolic disorders showing considerable genetic heterogeneity and associated pathogenic variants affecting over 85 different genes of the mitochondrial or nuclear genome [3]. Therefore, they represent mitochondrial disorders with the largest genetic heterogeneity [1].

As human SNE is rare and heterogeneous, studying domestic animal species showing resembling conditions might add to the understanding of such a complex group of disorders. Rare forms of SNE were described e.g., in cattle [5–7] and dogs (OMIA 001097-9615). The first report of this disorder was described in Alaskan huskies [8,9] and subsequently, a similar form of SNE was reported in Yorkshire terriers [10] and American Staffordshire bull terriers [11]. Neuropathologically, SNE in Yorkshire terriers is nearly identical to the Alaskan husky form and very similar to human Leigh syndrome [10]. An initial genetic investigation of SNE-affected Yorkshire terriers revealed no indication for disease-causing variants in the mitochondrial genome [10], whereas more recently in Alaskan huskies the pathogenesis of recessively inherited SNE was unraveled [12,13]. This breed-specific fatal brain disorder in Alaskan huskies is associated with a deleterious loss-of-function variant in *SLC19A3* encoding for a thiamine transporter 2 (THTR2) with a predominately central nervous system (CNS) distribution [12,13]. The *SLC19A3* gene product controls the uptake of thiamine in the CNS via expression of the thiamine transporter protein THTR2. Pathogenic variants are associated with thiamine metabolism dysfunction syndrome-2 in people (THMD2; OMIM 607483), also known as biotin-responsive basal ganglia disease (BBGD) or thiamine-responsive encephalopathy [14]. This *SLC19A3*–related condition is an autosomal recessive disorder with childhood-onset that presents as a subacute encephalopathy and progresses to severe cogwheel rigidity, dystonia, quadriparesis, and eventually death if left untreated (OMIM 606152). The *SLC19A3*–related SNE of Alaskan huskies was proposed as a possible large animal model that may allow prospective investigations into the mechanisms of *SLC19A3*-related syndromes and the potential role of thiamine and/or biotin as a therapeutic strategy [12,13].

To elucidate the disease mechanism underlying monogenic autosomal recessive inherited SNE in Yorkshire terriers, we applied homozygosity mapping and whole-genome sequencing revealing a most likely pathogenic variant in the canine *SLC19A3* gene.

2. Materials and Methods

2.1. Ethics Statement

All animal experiments were performed according to the local regulations. All animals in this study were examined with the consent of their owners. Sample collection was approved by the Cantonal Committee for Animal Experiments (Canton of Bern; permit 71/19).

2.2. Animals

In total, 172 blood samples of Yorkshire terriers were collected. Ten dogs were diagnosed with Leigh-like subacute necrotizing encephalopathy (SNE) according to Baiker et al. [10]. These affected dogs were unrelated, apart from two full siblings for which their sire and dam (obligate carriers) as well as a single normal littermate were also available. The remaining 159 dogs represented unrelated

purebred controls. Genomic DNA was isolated from EDTA blood samples using the Maxwell RSC Whole Blood DNA kit (Promega, Dübendorf, Switzerland).

2.3. Single Nucleotide Polymorphism Array Genotyping

Four selected SNE-affected Yorkshire terriers were genotyped on Illumina CanineHD BeadChip array (Illumina, San Diego, CA, USA). PLINK v1.9 [15] was used to perform the quality control filtering steps of the obtained genotyping data and the subsequent homozygosity analysis. Single nucleotide polymorphisms (SNP) with a call rate <90% were removed leaving 167,185 markers. All individuals had call rates >90%. Homozygosity analysis was carried out with PLINK v1.9 [15] to determine intervals of extended homozygous regions with alleles shared by all four affected dogs.

2.4. Whole-Genome Sequencing

Whole-genome sequence (WGS) data of a single affected dog was obtained at 19.7× coverage in order to identify the causative variant for SNE. The sequence data analysis and calling of single nucleotide variants and small indels (SNVs) including the prediction of functional effects were described before [16]. The dog reference genome assembly CanFam3.1 and NCBI annotation release 105 was used. Additionally, a publicly available control genomes cohort of 720 dogs from 130 various breeds, and 9 wolves [16] was used to filter variants private in the sequenced SNE-affected dog; this also included 60 unrelated Yorkshire terriers (Supplementary Table S1). The Integrative genomics viewer (IGV) software [17] was used for visual inspection and screening for structural variants in the associated regions.

2.5. Sanger Sequencing and Targeted Genotyping

Polymerase chain reaction (PCR) and Sanger sequencing were used to validate and characterize the *SLC19A3* indel variant (XM_022409850.1:c.205_210delins35) identified from whole-genome sequencing. PCR primers were designed using primer 3 [18]. PCR products from genomic DNA were amplified using AmpliTaqGold360 MasterMix (Thermo Fisher Scientific, Waltham, MA, USA) and the purified PCR amplicons were directly sequenced on an ABI3730 capillary sequencer (Thermo Fisher Scientific) using the following primers: GGCAGTCACCATCCCATAGA (forward) and GATATTGGGCAAGCCACCTA (reverse) generating 309 bp products. The sequence data were analyzed with Sequencher 5.1 software (GeneCodes, Ann Arbor, MI, USA). Diagnostic genotyping was performed by fragment length analysis using a different forward primer (ATCCCTTGCAGGATGATGAC) to produce amplicons of 218 bp or 247 bp representing the wild type or variant allele, respectively. The 29 bp size difference was visualized on a Fragment Analyzer capillary gel electrophoresis instrument (Advanced Analytical Technologies, Ames, IA, USA).

2.6. Availability of Data and Material

The whole-genome data of an SNE-affected Yorkshire terrier are freely available at the European Nucleotide Archive (ENA) under sample accession number SAMEA3928145. All accession numbers of the used control genomes are available in Supplementary Table S1.

All genome positions are reported with respect to dog reference genome assembly CanFam3.1 and NCBI annotation release 105. All references to the canine *SLC19A3* gene correspond to the accessions NC_006607.3 (NCBI accession), XM_022409850.1 (mRNA), and XP_022265558.1 (protein).

3. Results

3.1. Homozygosity Analysis

Based on the clinicopathological diagnosis of Leigh-like subacute necrotizing encephalopathy (SNE) in all examined Yorkshire terriers and the similarities to the recessively inherited conditions in SNE-affected Alaskan husky dogs and THMD2-affected humans, as well as the available pedigree

information of the two SNE-affected siblings, a recessive mode of inheritance was postulated. Therefore, homozygosity mapping assuming identity-by-descent (IBD) was used to determine critical genomic regions shared across four SNP array genotyped cases. This revealed five genome regions with a total of ~4.1 Mb located on five different dog chromosomes (Table 1), representing 0.17% of the canine reference sequence. Visual inspection of these regions in the WGS of the affected dog did not reveal any evidence for copy number variants or large structural rearrangements.

Table 1. Regions of shared homozygosity detected in four subacute necrotizing encephalomyelopathy (SNE)-affected Yorkshire terriers.

Chromosome	Position [1]		Length (kb)	Number of Annotated Protein-Coding Genes in the Region
	Start	End		
3	44,184,889	44,286,148	101.3	0
6	71,329,720	71,552,171	222.5	1
10	20,608,121	22,376,735	1768.6	22
25	39,477,619	41,191,570	1714.0	13
31	33,337,422	33,591,249	253.8	2

[1] in respect to dog reference genome assembly CanFam3.1.

3.2. Identification of the Causative Variant

Filtering the variants of a single affected Yorkshire terrier against 729 public control genomes [16], including 60 breed controls, for single-nucleotide variants (SNVs) and short indels present in the five identified IBD-regions resulted in only a single private protein-changing variant (Figure 1a). The indel affecting ~45 bp is located in exon 2 of the *thiamine transporter 2* (*LOC486151*) gene, also known as solute carrier family 19 members 3 (*SLC19A3*) gene (Figure 1b). PCR and subsequent Sanger sequencing confirmed the homozygous presence of this small structural variant in SNE-affected Yorkshire terriers and revealed the detailed features of the indel: a 35 bp insertion replacing 6 bp and thereby disturbing the correct reading frame (Figure 1c). There are 15 currently annotated transcript isoforms for the canine *SLC19A3*, which is in reverse complementary orientation with respect to the canine reference genome. While the canine SLC19A3 protein length is 495 amino acids, the human protein (NP_001358340.1) has 496 amino acids, from which 408 (82.3%) are identical between dog and human. The Yorkshire terrier variant leads to a frameshift and a premature stop codon (c.205_210delins35; p.Pro69Ilefs*45) truncating ~86% of the wild type coding sequence (Figure 1c).

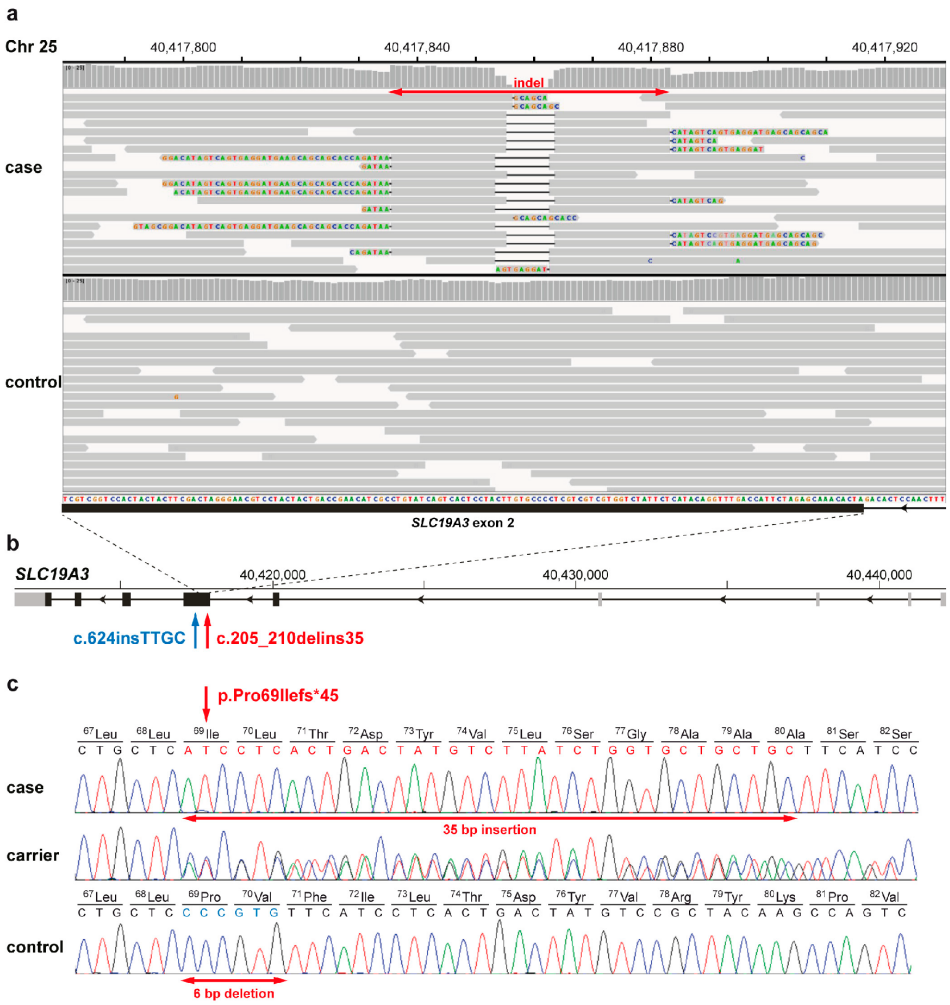

Figure 1. Subacute necrotizing encephalopathy (SNE)-associated *SLC19A3* loss-of-function variant in Yorkshire terriers. (**a**) IGV [17] screenshots of the genome region on canine chromosome 25 with the *SLC19A3*:c.205_210delins35 variant in an affected and a control Yorkshire terrier (NC_006607.3:40417780-40417930); The indel variant detected in the SNE-affected dog is indicated by a red arrow. (**b**) Schematic representation of the canine *SLC19A3* gene showing the location of both pathogenic variants in exon 2 (XM_022409850.1): the herein identified indel (red arrow) and the insertion previously described in encephalopathy-affected Alaskan huskies (blue arrow) [12]. Note that the number of 5′-untranslated exons (grey) varies between transcript isoforms, whereas the five protein-coding 3′-exons (black) are more conserved; (**c**) Sanger sequencing electropherograms illustrate sequences of a homozygous SNE-affected Yorkshire terrier, a heterozygous carrier, and a homozygous wild type dog. The red arrows indicate that the 35 bases shown in red are inserted, whereas the 6 bases in blue are deleted in the mutant allele. The predicted consequence of the shift in the reading frame altering the amino acid sequence of the SLC19A3 protein and leading to a premature stop is shown above.

3.3. Targeted Genotyping of the Variant

Genotyping by fragment size analysis of the 172 available Yorkshire terriers confirmed perfect segregation of the detected *SLC19A3* variant with the observed disease phenotype. Only the ten

SNE-affected dogs were homozygous for the variant allele (Table 2). Two obligate carriers and one tested normal littermate were heterozygous carriers of the variant, while 162 controls tested homozygous for the wild type allele (Table 2).

Table 2. Segregation of the *SLC19A3*: c.205_210delins35 genotypes with subacute necrotizing encephalopathy in Yorkshire terriers.

SNE Status	wt/wt	wt/var	var/var
Affected ($n = 10$)	0	0	10
Non-affected ($n = 222$) [1]	219	3 [2]	0

[1] including 60 dogs with WGS data [2] includes 2 obligate carriers and 1 normal littermate of the affected dogs.

4. Discussion

In this study, the obtained genetic results elucidate the underlying aetiology of the previous clinical and pathological characterization of a Leigh-like subacute necrotizing encephalopathy in the affected Yorkshire terriers, which resembles the human Leigh syndrome. The *SLC19A3* variant found by a combination of SNP genotyping-based homozygosity mapping and whole-genome sequencing, confirmed by Sanger sequencing, segregated perfectly in the investigated cohort of >200 unrelated Yorkshire terriers.

Numerous homozygous as well as compound heterozygous variants have been reported before in different regions of *SLC19A3* in human patients suffering from thiamine metabolism dysfunction syndrome-2 [19]. *SLC19A3* is a member of solute carrier family 19 and encodes thiamine transporter 2. Together with thiamine transporter 1, it is necessary for transport and homeostasis of thiamine that is important in brain development. [20]. *Slc19a3*-knockout mice showed progressive wasting and lethargy leading to a premature death as well as a significant decrease in thiamin uptake, even though there were no obvious histological changes in the brain [21].

The herein-described most likely pathogenic variant (XP_022265558.1:p.Pro69Ilefs*45) lies within the second of 12 transmembrane domains of the SLC19A3 protein and, therefore, affects ~86% of the wild type sequence. The *SLC19A3* gene probability of loss-of-function intolerance is pLI = 0.104 [22], which indicates variants in *SLC19A3* leading to a loss of gene function are most likely recessive, where loss of a single copy is often tolerated but the loss of both copies is not. The herein-described variant leads to an insertion of a premature termination in the second out of five coding exons, suggesting that any synthesized mRNA would likely be degraded through nonsense-mediated decay, unlikely to produce a fully functional protein. Heterozygous carriers did not show a visible clinical phenotype, as they can most likely compensate due to the presence of the normal protein, albeit at a decreased amount.

In conclusion, our results provide strong evidence for a breed-specific deleterious variant in *SLC19A3* as the most likely genetic cause of monogenic autosomal recessive Leigh-like subacute necrotizing encephalopathy in Yorkshire terriers, and they enable the development of a genetic test for veterinary diagnostic and breeding decisions. Finally, this presents the second, most likely breed-specific pathogenic variant in the canine *SLC19A3* gene in SNE-affected dogs.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/11/10/1215/s1, Table S1: Accession numbers of 721 dog and 9 wolf whole-genome genome sequences.

Author Contributions: Conceptualization, K.M. and C.D.; methodology, M.D.; data curation, V.J.; formal analysis, M.D. and A.L.; investigation, M.D.; resources, K.M., D.C., M.R., K.J., S.M., T.G., S.R., A.F. and A.L.F.-P.; writing—original draft preparation, A.L., M.D., K.M. and C.D.; writing—review and editing, M.D., A.L., K.M. and C.D.; visualization, A.L. and C.D.; software, V.J.; supervision, C.D.; funding acquisition, C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors are grateful to the owners of all dogs who provided samples and shared valuable information. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the WGS and providing high-performance computational infrastructure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Finsterer, J. Leigh and Leigh-Like Syndrome in Children and Adults. *Pediatr. Neurol.* **2008**, *39*, 223–235. [CrossRef] [PubMed]
2. Leigh, D. Subacute necrotizing encephalomyelopathy in an infant. *J. Neurol. Neurosurg. Psychiatry* **1951**, *14*, 216–221. [CrossRef] [PubMed]
3. Chang, X.; Wu, Y.; Zhou, J.; Meng, H.; Zhang, W.; Guo, J. A meta-analysis and systematic review of Leigh syndrome: Clinical manifestations, respiratory chain enzyme complex deficiency, and gene mutations. *Medicine* **2020**, *99*, e18634. [CrossRef] [PubMed]
4. Fecek, C.; Samanta, D. Subacute Necrotizing Encephalomyelopathy (Leigh Syndrome). Available online: http://www.ncbi.nlm.nih.gov/pubmed/32644590 (accessed on 14 September 2020).
5. Steffen, D.J.; Vestweber, J.G.; Cash, W.; El-Hamidi, M.; Leipold, H.W. Multifocal Subacute Necrotizing Encephalomyelopathy in Simmental Calves. *J. Vet. Diagn. Investig.* **1994**, *6*, 466–472. [CrossRef]
6. Desjardins, I.; Fecteau, G.; Hélie, P.; Desrochers, A. Multifocal subacute necrotizing encephalomyelopathy in a Simmental calf. *Can. Vet. J.* **2001**, *42*, 375–377.
7. Philbey, A.W.; Martel, K.S. A multifocal symmetrical necrotising encephalomyelopathy in Angus calves. *Aust. Vet. J.* **2003**, *81*, 226–229. [CrossRef] [PubMed]
8. Wakshlag, J.J.; de Lahunta, A.; Robinson, T.; Cooper, B.J.; Brenner, O.; O'Toole, T.D.; Olson, J.; Beckman, K.B.; Glass, E.; Reynolds, A.J. Subacute necrotising encephalopathy in an Alaskan husky. *J. Small Anim. Pract.* **1999**, *40*, 585–589. [CrossRef]
9. Brenner, O.; Wakshlag, J.J.; Summers, B.A.; de Lahunta, A. Alaskan Husky encephalopathy—A canine neurodegenerative disorder resembling subacute necrotizing encephalomyelopathy (Leigh syndrome). *Acta Neuropathol.* **2000**, *100*, 50–62. [CrossRef] [PubMed]
10. Baiker, K.; Hofmann, S.; Fischer, A.; Gödde, T.; Medl, S.; Schmahl, W.; Bauer, M.F.; Matiasek, K. Leigh-like subacute necrotising encephalopathy in Yorkshire Terriers: Neuropathological characterisation, respiratory chain activities and mitochondrial DNA. *Acta Neuropathol.* **2009**, *118*, 697–709. [CrossRef] [PubMed]
11. Collins, D.; Angles, J.M.; Christodoulou, J.; Spielman, D.; Lindsay, S.A.; Boyd, J.; Krockenberger, M.B. Severe Subacute Necrotizing Encephalopathy (Leigh-like Syndrome) in American Staffordshire Bull Terrier Dogs. *J. Comp. Pathol.* **2013**, *148*, 345–353. [CrossRef] [PubMed]
12. Vernau, K.M.; Runstadler, J.A.; Brown, E.A.; Cameron, J.M.; Huson, H.J.; Higgins, R.J.; Ackerley, C.; Sturges, B.K.; Dickinson, P.J.; Puschner, B.; et al. Genome-Wide Association Analysis Identifies a Mutation in the Thiamine Transporter 2 (SLC19A3) Gene Associated with Alaskan Husky Encephalopathy. *PLoS ONE* **2013**, *8*, e57195. [CrossRef] [PubMed]
13. Vernau, K.; Napoli, E.; Wong, S.; Ross-Inta, C.; Cameron, J.; Bannasch, D.; Bollen, A.; Dickinson, P.; Giulivi, C. Thiamine Deficiency-Mediated Brain Mitochondrial Pathology in Alaskan Huskies with Mutation in SLC19A3.1. *Brain Pathol.* **2015**, *25*, 441–453. [CrossRef] [PubMed]
14. Marcé-Grau, A.; Martí-Sánchez, L.; Baide-Mairena, H.; Ortigoza-Escobar, J.D.; Pérez-Dueñas, B. Genetic defects of thiamine transport and metabolism: A review of clinical phenotypes, genetics and functional studies. *J. Inherit. Metab. Dis.* **2019**, *42*, 581–597. [CrossRef] [PubMed]
15. Chang, C.C.; Chow, C.C.; Tellier, L.C.A.M.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 1–16. [CrossRef] [PubMed]
16. Jagannathan, V.; Drögemüller, C.; Leeb, T.; Aguirre, G.; André, C.; Bannasch, D.; Becker, D.; Davis, B.; Ekenstedt, K.; Faller, K.; et al. A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. *Anim. Genet.* **2019**, *50*, 695–704. [CrossRef] [PubMed]
17. Thorvaldsdóttir, H.; Robinson, J.T.; Mesirov, J.P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **2013**, *14*, 178–192. [CrossRef] [PubMed]
18. Rozen, S.; Skaletsky, H. Primer3 on the WWW for General Users and for Biologist Programmers. In *Bioinformatics Methods and Protocols*; Humana Press: Totowa, NJ, USA, 2000; Volume 132, pp. 365–386. ISBN 978-0-89603-732-8. [CrossRef]

19. Whitford, W.; Hawkins, I.; Glamuzina, E.; Wilson, F.; Marshall, A.; Ashton, F.; Love, D.R.; Taylor, J.; Hill, R.; Lehnert, K.; et al. Compound heterozygous SLC19A3 mutations further refine the critical promoter region for biotin-thiamine-responsive basal ganglia disease. *Mol. Case Stud.* **2017**, *3*, a001909. [CrossRef] [PubMed]
20. Eudy, J.D.; Spiegelstein, O.; Barber, R.C.; Wlodarczyk, B.J.; Talbot, J.; Finnell, R.H. Identification and characterization of the human and mouse SLC19A3 gene: A novel member of the reduced folate family of micronutrient transporter genes. *Mol. Genet. Metab.* **2000**, *71*, 581–590. [CrossRef] [PubMed]
21. Reidling, J.C.; Lambrecht, N.; Kassir, M.; Said, H.M. Impaired Intestinal Vitamin B1 (Thiamin) Uptake in Thiamin Transporter-2–Deficient Mice. *Gastroenterology* **2010**, *138*, 1802–1809. [CrossRef]
22. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]

Discussion and perspectives

In this thesis, I investigated the molecular basis of seven different rare forms of canine neurological diseases as well as the population diversity and disease prevalence of one particular breed, the Leonberger. I was able to use different methods of genetic analysis and in all of these studies, compelling candidate genetic variants were successfully identified. The studied disorders represent a wide phenotypic spectrum of nervous system abnormalities and fall into different groups of neurogenetic diseases affecting various parts of the nervous system and diverse biological processes (Figure 5).

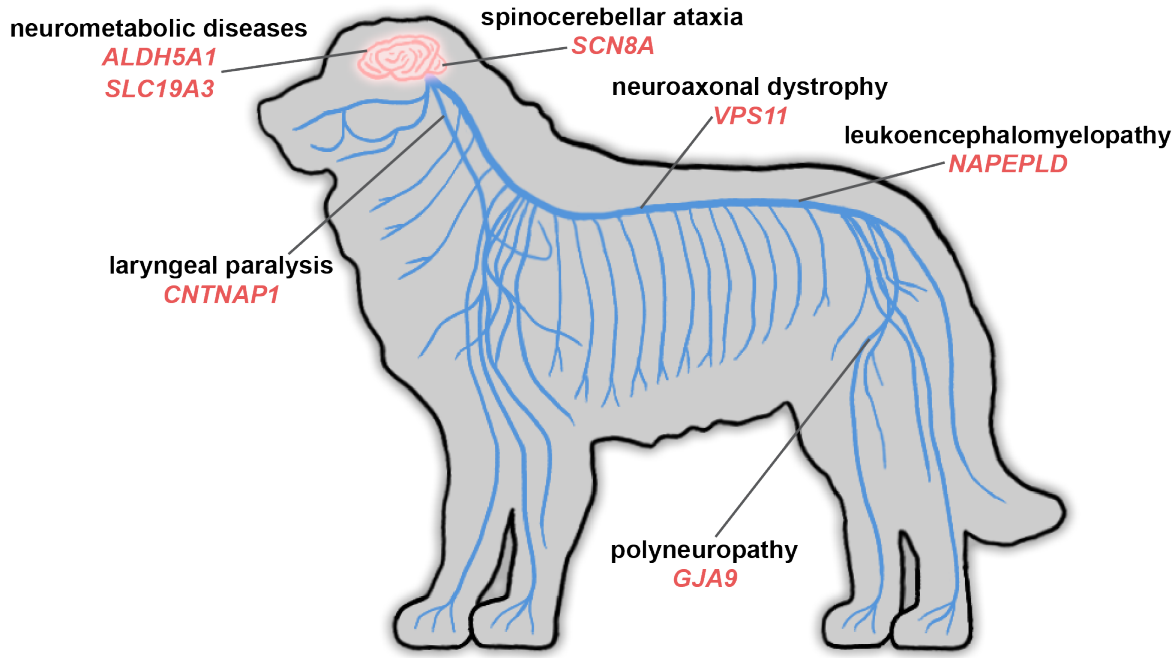

Figure 5. Simplified representation of the investigated canine neurological diseases showing the different parts of the nervous system they primarily affect, including the genes harboring the identified variants associated with each disorder (in red).

The dramatic inbreeding situation in some dog populations often leads to breed-specific increased prevalence of inherited diseases. In the course of this thesis, I intensively studied the Leonberger breed as extensive genealogical, SNP array, as well as WGS data was available for more than 140,000, 1,200, and 39 dogs, respectively. In similar studies of within-breed diversity performed for different breeds, comparably smaller SNP array datasets have been previously explored, e.g. 34 German shorthaired pointers [108], 48 Braque Français [109], 188 Bullmastiff dogs [110], or 255 Border collies [111]. Also WGS data of a limited number of dogs, e.g. 15 Standard poodles [112] or 6 Norwegian lundehund dogs [113], have been used before to characterize the ROH and explore fixed or enriched variants present in the breeds. My study of the Leonberger breed provided the first comprehensive characterization of the worldwide population and confirmed the assumed reduced genetic diversity after a historical bottleneck, as well as recent inbreeding despite the ever-growing

number of dogs. Analysis of pedigree and medical records of thousands of purebred Leonbergers revealed a serious popular sire effect, high relatedness among the dogs, a slow continuous decrease of average life expectancy across time (from 9.4 years in 1989 to 7.7 years in 2004), and a high prevalence of various forms of cancer, orthopedic, and neurological disorders. Using SNP array data I found no obvious subpopulations in the dataset collected worldwide from 28 countries, and investigated the ROH to estimate the average genomic inbreeding coefficient (F_{ROH}) of 0.28. This was expected in an inbred modern dog breed and falls within the range reported in breeds, such as Rottweiler (F_{ROH}=0.29), German shepherd (F_{ROH}=0.31), or Bernese mountain dog (F_{ROH}=0.30) [114]. However, many breeds show comparably lower values, such as Czechoslovakian wolfdog (F_{ROH}=0.17) [115], Braque Français (F_{ROH}=0.18), Poodle (F_{ROH}=0.18), Border collie (F_{ROH}=0.15), or Jack Russell terrier (F_{ROH}=0.08) [114]. On the other hand, the Norwegian lundehund breed that is highly inbred and known to be at risk for a breed-specific multifactorial life-threatening syndrome has an inbreeding coefficient as high as 0.87 [116]. Furthermore, the WGS data provided evidence for breed-specific non-coding variants as well as private (unique to Leonbergers) and enriched (rarely seen in other breeds) potentially pathogenic variants in 38 Leonbergers affected by neurological disease. In addition, only two haplotypes were detected in their mitochondrial genomes. The overall high relatedness underlines the elevated risk for the occurrence of recessive Mendelian diseases. On the other hand, fixation or enrichment of many uncharacterized alleles in various genes due to the excessive usage of a few popular sires might explain the overall general predisposition towards more complex neurological disorders [117]. Besides the technical limitations of using only short-read sequencing methods, late-onset forms of disorders, and possible complex underlying genetics complicate the discovery of disease-associated variants and illustrate the limitation of current genomic approaches. In the light of the low diversity and always new emerging diseases in the Leonberger breed, outcrossing might be a prospective solution to increase heterozygosity and improve the overall health and well-being of the dogs in the long term. It has to be, however, performed continuously and ideally complemented with changes in population structure and breeding organization to have a lasting effect [118].

Studies of inherited neuropathies are challenging due to their complexity and heterogeneity of clinicopathological signs, as well as other features such as varying age of onset. I described two novel forms of canine polyneuropathy and/or laryngeal paralysis (LPPN) during this thesis. The affected dogs showed diverse phenotypes of clinical signs and severity. As known from human medicine, the nomenclature of CMT forms is complicated and has been regularly revised due to the continuous identification of new genetic causes [21]. In dogs, the situation is comparably complicated and a classification based solely on clinicopathological observations is usually not precise enough [26,31]. The success of finding the causal variant

by genomic analysis is dependent on the available sample size, especially in the presence of significant disease heterogeneity. In two studies of canine forms of polyneuropathy, different strategies of GWAS combined with WGS were applied. First, we assumed a breed-specific variant for an adult-onset form of polyneuropathy observed in Leonbergers. Using SNP array genotypes of more than 200 cases and controls we fine-mapped the associated ~7.7 Mb-sized genomic region in which private case-specific variants were then filtered in WGS data. This approach revealed a private homozygous 2 bp deletion leading to a frameshift in the *GJA9* gene encoding a connexin protein involved in the formation of gap junctions. This likely pathogenic small indel was present in about a quarter of the studied cases illustrating the challenge of genetic heterogeneity. Following segregation analysis by genotyping thousands of Leonbergers suggested a dominant mode of inheritance with incomplete penetrance as some heterozygous dogs did not show any clinical signs at the age of 8 years. Other gap junction proteins have been associated with diverse nervous system disorders in humans, including a form of CMT [119,120]. Thus, our study adds *GJA9* for the first time to the list of candidate genes for human neuropathy patients with unsolved etiology.

In the subsequently performed study aiming to detect the causes of further forms of LPPN, I found an autosomal recessive missense variant in the *CNTNAP1* gene in two closely related dog breeds, the Leonberger and Saint Bernard, showing early signs of laryngeal paralysis as the main disease sign. Using strict inclusion criteria for the cases and controls, across-breed GWAS with imputed SNP array data of more than 500 dogs was successful in mapping the disease-associated genomic region to about 1 Mb. Exploring this region in the WGS data of 5 cases from both breeds for shared variants identified a single protein-changing missense variant in the *CNTNAP1* gene. *CNTNAP1* encodes a contactin-associated protein important for the organization of myelinated axons and variants in this gene have been previously implicated in human childhood-onset CMT, and hypomyelinating neuropathy, including vocal cord palsy and severe respiratory distress [121,122]. The canine LPPN form seems therefore clinically equivalent to the human disease. Only 4% of the analyzed affected Leonbergers and 42% of Saint Bernards were homozygous mutant for this variant illustrating once more the challenge of large genetic heterogeneity of different forms of canine LPPN. Interestingly, when searching the Dog biomedical variant database [20] of >700 publicly available dogs of 131 different breeds and wolves, I found the *CNTNAP1* variant in several other unrelated dog breeds, indicating it likely predates modern breed establishment. Using targeted genotyping by Sanger sequencing of almost 5,000 animals, the variant occurred in 25 different dog breeds with mostly unknown precise health history. However, we were able to evaluate available LPPN phenotypes in one breed, the Labrador retriever, where homozygosity for the described variant explained a portion of cases with a younger age of onset. Whether this variant is disease-causing also in smaller dog breeds segregating for this variant needs to be evaluated

in the future when detailed disease phenotypes are available. Furthermore, the breed differences might be due to different genetic background and/or differences in size and correspondingly shorter laryngeal nerves. Such relationship was previously suggested by correlation between height and laryngeal neuropathy in horses [123].

Leukoencephalomyelopathy (LEMP) in Leonberger and Rottweiler dogs is classified as a myelin disorder. I used similar methodological approaches in this study of a limited number of affected dogs from the two breeds as in the previously described polyneuropathies. Within-breed GWAS was performed and subsequent WGS and targeted genotyping by Sanger sequencing revealed a private deleterious variant in the same gene in each breed. Both variants, a missense and a frameshift, occurred in the *NAPEPLD* gene encoding an enzyme of the endocannabinoid system showing evidence for obvious allelic heterogeneity of this disorder in dogs. Similar heterogeneity was previously found e.g. in *RAB3GAP1*-related polyneuropathy that is associated with a large insertion in Alaskan husky [124], and a frameshift variant in Black Russian terrier and Rottweiler [125] (OMIA 001970-9615). The first description of *NAPEPLD*-associated inherited defects in the endocannabinoid system might be important for future investigation of myelin disorders, as the neuromodulatory system has important functions in central nervous system development, synaptic plasticity, and the response to endogenous and environmental insults [126]. Interestingly, when screening other related breeds for the *NAPEPLD* frameshift variant, I found it was present heterozygously in several unaffected Great Danes indicating an older origin of this mutation compared to the missense variant found in Leonbergers. However, until now no homozygous mutant dog was seen in this third breed. The identification of different independently occurring variants in the same gene and diverse dog breeds affected by one disease with heterogeneous pathological features implicates the NAPEPLD enzyme as important in myelin homeostasis. Therefore, this study provides *NAPEPLD* as a new functional candidate gene for LEMP, a canine model for the corresponding human myelin disorders, to establish molecular diagnosis and potential pre-clinical trials in dogs, and highlights dog as a model to study the functions and molecular pathology of the *NAPEPLD* gene.

Another herein described disease is the neuroaxonal dystrophy seen in young Rottweilers, which has clinicopathological features in common with the human neuroaxonal dystrophy. DNA samples from seven affected and almost 300 normal Rottweilers were available. Interestingly, three of the cases examined during the study were not closely related, while the other four were halfsiblings born around 1980 [127]. This suggested a relatively older mutation event and silent transmission of the recessive disease-causing allele through many generations. Therefore, homozygosity mapping was performed to narrow down the associated genomic region detected by GWAS. By WGS of two cases and comparison of their data with ~100 controls of 25 different breeds, only a single case-specific homozygous protein-changing

variant in the *VPS11* gene was found. Further targeted genotyping of more than 400 dogs from 20 breeds confirmed the suspected breed-specificity of the *VPS11* missense variant and an estimated allele frequency of ~2% in the Rottweiler population. *VPS11* encodes a protein with a known function in vesicle transport to vacuoles and an essential role in the endosomal-autophagy-lysosomal pathways [128]. In humans, *VPS11* mutations are associated with an infantile-onset neurological syndrome characterized as hypomyelinating leukodystrophy and lysosomal storage disease [129]. The human and canine phenotypes have distinct differences, clearly pointing out the spectrum of phenotypic heterogeneity caused by variants within the same gene disrupting an important biological system. Therefore, *VPS11* represents a potential candidate for unexplained forms of human neuroaxonal dystrophy patients.

Spinocerebellar ataxia in Alpine dachsbracke dogs is an example of a simple monogenic recessive disease. Four puppies from two litters showed signs of cerebellar dysfunction, while all parents of the affected dogs were phenotypically normal. By pedigree analysis of the two ataxia-affected litters, I revealed an obvious recent inbreeding loop. Although only a very limited number of cases and their relatives was available, the outcome of the genetic analyses was unambiguous. Assuming the recessive mode of inheritance, homozygosity and linkage analyses were performed to map the ~5.5 Mb-sized disease-associated locus, followed by WGS filtering for homozygous SNVs unique to the sequenced case. This led to the identification of a single non-synonymous variant in the *SCN8A* gene, which segregated perfectly with the phenotype in the pedigree; after further targeted genotyping of ~200 unrelated Alpine dachsbracke dogs, the variant occurred homozygously only in the original four studied cases. SCN8A encodes a voltage-gated sodium channel [130] and variants in this gene were previously reported in human patients exhibiting ataxia, cerebellar atrophy, cognitive disability, as well as early-onset epileptic encephalopathies [131,132]. The developed DNA test enables informed breeding decisions and provides an important tool in the precise diagnosis of canine ataxia as it is a relatively non-specific condition.

Finally, I investigated two canine neurometabolic disorders, namely succinic semialdehyde dehydrogenase deficiency (SSADHD) in Saluki dogs and subacute necrotizing encephalopathy in Yorkshire terriers. Even though both diseases fall into the same category of inborn errors of metabolism, they represent very different clinicopathological phenotypes, also reflected in the diverse underlying genetics. Consequently, different samples were available, and various methods for genomic analysis were used. In the SSADHD-affected Saluki dogs, independent studies were initially conducted in Europe and the USA in different Saluki litters showing the identical rare disease phenotype. Through international collaboration the available data was then combined, strengthening the power of detection of the disease-associated variant. All seven cases had striking similarities to SSADHD in humans, although the canine phenotype was more severe. By in-depth pedigree analysis, I showed the presence

of a common male ancestor connecting the American and European dog families and suspected an autosomal recessive mode of inheritance. GWAS with only a limited number of dogs and subsequent filtering of WGS data of cases from geographically different locations revealed a breed-specific missense variant in the *ALDH5A1* gene. This gene encodes a mitochondrial enzyme SSADH involved in the catabolism of major inhibitory neurotransmitter in the CNS [133]. Disruption of this metabolic pathway leads to an accumulation of potentially toxic metabolites in brain, and was previously described in human neurological disease with wide phenotypic heterogeneity [134]. Elevated levels of such metabolites were also detected in the Saluki cases and can be used as biomarkers in urine or serum for canine SSADHD. As there is no effective treatment, the dog provides a potential model for evaluation of therapeutic approaches that have so far not been tested in SSADHD, such as pharmacological treatment by targeting neurotransmitter receptors, enzyme-replacement therapy, or gene therapy. This large-animal model could be especially important because of early lethality in the existing *Aldh5a*$^{-/-}$ knockout mouse model, which makes any treatment trials problematic [135].

The second metabolic disorder occurred in ten mostly unrelated Yorkshire terriers that were diagnosed with subacute necrotizing encephalopathy (SNE) according to the previous neuropathological description of the condition, which closely resembled a form of SNE in a different dog breed [136] as well as Leigh syndrome (LS) in humans [137]. Human LS is an early-onset, often fatal progressive mitochondrial respiratory chain disease affecting the CNS with considerable genetic heterogeneity. Apart from the mutations of respiratory chain protein complexes, a biochemical defect in thiamine metabolism was also suggested in LS. Pathogenic variants in over 60 different genes of the mitochondrial or nuclear genome still do not explain all described human cases [138]. In the previous study of SNE in Yorkshire terriers, Baiker *et al*. used targeted sequencing and Southern blot analysis to exclude mitochondrial tRNA mutations and large genetic rearrangements as the possible underlying cause [137]. In our study, I used homozygosity analysis combined with WGS comparison of the SNE-affected dog with 60 unrelated Yorkshire terriers and more than 600 other breed controls to identify a small structural variant in *SLC19A3* gene. Sanger sequencing of the variant confirmed its homozygous presence in all SNE-affected dogs and revealed the detailed features of this indel that affects ~45 nucleotides and disturbs the reading frame. Notably, the initial automatic variant calling reported a simple 8 bp deletion highlighting the need for experimental validation of the WGS output in the lab. The *SLC19A3* gene encodes a thiamine transporter 2 important in brain development. Its disruption by homozygous or compound heterozygous variants was previously seen in human thiamine metabolism dysfunction syndrome, also known as thiamine-responsive encephalopathy [139,140]. A different small indel in the same gene was also reported in Alaskan huskies with SNE [141]. The identification of a second, independent variant in Yorkshire terriers, therefore, adds evidence to the causal relationship between

SLC19A3 and this form of neurometabolic disorder. The availability of naturally occurring dogs with a Leigh-like disease may provide a feasible approach for an investigation of mitochondrial disease mechanisms, and serve as a model for further comparative and translational research.

In summary, the eight developed DNA tests (Table 6) provide breeders a tool to improve animal health and welfare. Although directly eliminating all carriers from the breeds would quickly decrease the mutant allele frequency, it might have negative consequences on genetic diversity of the closed populations. Therefore, genetic testing of breeding dogs to avoid carrier matings and slowly decreasing the allele frequency is usually recommended [117]. This strategy has been successfully adopted in recent years as shown by the decline in the number of affected dogs as well as the disease-associated variant frequencies within several breeds following the commercial availability of DNA tests [142]. Notably, the various sizes of the mapped critical regions together with the variant occurrence in other breeds support the different assumptions about the age of mutation events, as e.g. the *CNTNAP1*-associated region was almost eight times shorter compared to the *GJA9* haplotype. Furthermore, recent inbreeding loops result in relatively larger associated regions as seen in the ataxia-affected Alpine dachsbracke dogs as opposed to a probably older and more widespread variants in the SNE-affected Yorkshire terriers or SSADHD-affected Saluki dogs (Table 6).

Table 6. A list of likely causative variants in the seven investigated phenotypes showing the size of mapped associated genomic regions, the total number of genotyped dogs in each study, and the estimated variant allele frequencies in the respective cohorts of genotyped dogs.

Breed	MOI[1]	Critical region (Mb)	Associated gene	Variant[2]	Genotyped dogs	Variant allele frequency (%)
Polyneuropathy and/or laryngeal paralysis						
Leonberger	AD	7.68	*GJA9*	FS	7,455	3.2
Leonberger					2,738	6.6
Saint Bernard	AR	0.98	*CNTNAP1*	MS	305	13.9
Labrador retriever					1,524	5.2
Myelin disorder						
Leonberger		0.95		MS	7,086	8.5
Rottweiler	AR	0.05	*NAPEPLD*	FS	233	5.4
Great Dane		0.05		FS	262	4.6
Neuroaxonal dystrophy						
Rottweiler	AR	4.46	*VPS11*	MS	288	4.7
Spinocerebellar ataxia						
Alpine dachsbracke	AR	5.53	*SCN8A*	MS	285	4.9
Neurometabolic disorder						
Saluki	AR	2.68	*ALDH5A1*	MS	72	21.5
Yorkshire terrier	AR	1.71	*SLC19A3*	FS	232	5.0

[1] MOI = mode of inheritance as follows: AR = autosomal recessive, AD = autosomal dominant.
[2] Type of variant as follows: MS = missense, FS = frameshift. Multiple entries correspond to different variant types found in respective breeds.

In theory, access to a single affected animal is sufficient to find the genetic cause of a simple recessive condition if a candidate gene is affected by a variant that is detectable by standard bioinformatics pipeline (i.e. SNV or small indel), and has a relatively easily predictable effect on the encoded protein. Many more related and unrelated case and control animals are needed for genetic analysis in situations of complex phenotypes with more complicated modes of inheritance, presence of phenocopies, and large structural, silent, or non-coding variants with difficult effect predictions. For example, a synonymous variant in an exon of the bovine *MFN2* gene has been perfectly associated with degenerative axonopathy in Tyrolean grey cattle [143]. The impact of such variant was unclear until further transcript analysis by Northern blot and RT-PCR was performed, which revealed partial intron retention leading to a premature stop codon [143]. Furthermore, the effect and impact prediction of non-coding variants has become one of the main challenges of the NGS era. To correctly interpret and finally prove the causality of an identified variant, several conditions should be met [144,145]. With appropriate sample sizes and properly described phenotypes, sufficient proof of causality is usually achieved with genomic data for highly penetrant variants following Mendelian inheritance. In complex traits, robust evidence of pathogenicity is more challenging and often requires additional functional validation studies with sophisticated tools such as expression studies in different model systems, or CRISPR/Cas9 gene editing [146]. For example, successful preservation of myelin was previously shown in a mouse model of the *PMP22*-associated form of CMT after targeted editing by CRISPR/Cas9, and the approach partially rescued the axonal demyelination even after the onset of disease [147].

Due to the small number of human patients with rare neurogenetic diseases, and the large lists of variants identified from WGS, it might be difficult to pinpoint the pathogenic variant. Thus, the availability of clinically relevant, reproducible, and representative animal models is crucial [148]. Naturally occurring inherited disorders in companion animals, such as dogs and cats, are often orthologous to those found in humans and, in comparison with mouse models, provide a more relevant large-animal system enabling the development of e.g. accurate biodistribution studies due to increased brain and body size, or long-term efficacy and safety studies due to longer lifespan [148]. Several dog models of inherited musculoskeletal and neurodegenerative disorders, such as various myelin disorders or neurometabolic diseases, have been already described and proven invaluable for developing therapies for humans as well as for dogs in the frame of precision medicine [44]. The usefulness of naturally occurring canine models is evident from many examples, such as enzyme replacement therapy and adeno-associated viral gene therapy tested in the Dachshund model of *CLN2*-related NCL [149,150]. Both treatment studies showed a delay in the onset of neurological deficits and an increased lifespan in the dogs, and led to the initiation of human clinical trials [149,150]. Globoid cell leukodystrophy (GLD) is an example of myelin disorder affecting both the CNS

and PNS that often results in early death. As the current and only available treatment is hematopoietic stem cell transplantation in pre-symptomatic patients, the evaluation of different approaches in canine pre-clinical trials is important [44]. Bradbury *et al.* showed that adeno-associated viral gene therapy in the naturally occurring canine model of *GALC*-associated GLD improved myelination in both the CNS and PNS, delayed the onset of clinical signs, and doubled the lifespan of GLD-affected dogs. These results may offer dramatic improvements for treatment of human patients [151]. The herein described spontaneous canine models of neurometabolic disorders related to *SLC19A3* and *ALDH5A1* defects may serve similarly in the future.

Furthermore, the assignment of novel functions to disease-associated genes in dog studies leads to new functional candidates for similar human conditions. Previously, for example, *SERPINH1*-related osteogenesis imperfecta in Dachshunds [66] or *ASPRV1*-related skin disorder in German shepherds [152] have been characterized before the first variants in these genes were discovered in human patients with unexplained forms of the corresponding disorders [153,154], and thus provided additional evidence to the causal relationship. The herein described association of *GJA9* with neuropathy or *NAPEPLD* with myelin disorders provides a basis for such discoveries in human patients.

In conclusion, this thesis shows the great potential of WGS-based precision diagnosis in veterinary medicine using large datasets produced with affordable NGS techniques and thus confirms numerous similar reports in companion animals [91,155–157]. The typical structure of dog populations allows for efficient use of high-density SNP array data in various genetic mapping methods. With a combination of these approaches, I described the genetic etiology of seven canine forms of rare inherited neurological diseases. My results provide an immediate translation into breeding practice as genetic tests are adopted by breeders of the respective dog populations to avoid producing affected dogs and eradicate the described conditions. Beyond aspects of dog breeding and veterinary medicine, the disease association of pathogenic variants in genes that are not yet known to be involved in corresponding human genetic disorders provides valuable insights into the molecular mechanisms and biological functions, and contributes novel candidate genes for human patients with unsolved genetic etiology. Further research into different forms of canine neurological diseases, with an in-depth description of the pathology, may help the development of therapeutic interventions benefiting both human and canine patients.

Acknowledgments

~Discovery is easy. Understanding is the hard part.~

This thesis was carried out at the Institute of Genetics at the University of Bern under the supervision of Cord Drögemüller. I am very grateful to Cord for welcoming me to the group briefly for my master thesis research, and later inviting me back to Bern, and giving me the opportunity to continue pursuing my interest in animal genetics during this massive PhD project. It has been a long journey with a steep learning curve, but in all respects very much worthwhile. Thank you for both professional and personal support and advice.

All this work would not be possible without the help and support of Katie Minor and Jim Mickelson from the Canine Genetics Lab at the University of Minnesota. Thank you Katie and Mick for always being there to discuss, suggest, edit, proofread, motivate, and so much more.

A big thank you goes to Nathalie Besuchet-Schmutz in our lab for the tremendous help with genotyping and always knowing where everything is. Equally big thank you to Vidhya Jagannathan behind the computer for showing me the magic of bioinformatics and always being helpful even with a full table of work to do.

I would like to thank the whole genetics team for creating the best working environment one could ever wish for. Thank you to Tosso Leeb for friendly discussions, always new challenges, and overall support. Furthermore, this includes all the past and present group members, thanks to fellow PhD students for fun discussions about common problems, all doctoral students for providing essential veterinary advice, master students for allowing me to develop my mentoring skills, as well as the secretaries for always knowing the answer when filling out difficult forms.

Thank you to my thesis committee, Rémy Bruggmann for being my co-supervisor, André Schaller for being my mentor, and Göran Andersson for agreeing to act as my external co-referee.

Thank you to all dog owners, breeders, breeder clubs, and unions, as well as veterinarians for sending us samples, pedigrees, and health reports about their dogs, without this valuable information and continuing support there would be no project. I would also like to acknowledge all other collaborators in the different research projects. It is a pleasure to work together with great scientists from around the world.

My warmest thank you to my family for the enormous support during this journey and a very special thank you to Ivan Letko, my husband, cook, manager, illustrator, fellow gamer, emotional support, and above all my love, for being on the journey with me.

Curriculum vitae

Removed due to data privacy reasons.

List of publications

Published articles

Letko, A.; Schauer A.M.; Derks M.F.L.; Grau-Roma L.; Drögemüller, C.; Grahofer A. Phenotypic and Genomic Analysis of Cystic Hygroma in Pigs. *Genes (Basel).* **2021**, *12*, 207, doi:10.3390/genes12020207.

Letko, A.; Minor, K.M.; Friedenberg, S.G.; Shelton, G.D.; Salvador, J.P.; Mandigers, P.J.J.; Leegwater, P.A.J.; Winkler, P.A.; Petersen-Jones, S.M.; Stanley, B.J.; Ekenstedt, K.J.; Johnson, G.S.; Hansen, L.; Jagannathan, V.; Mickelson, J.R.; Drögemüller, C. A *CNTNAP1* Missense Variant Is Associated with Canine Laryngeal Paralysis and Polyneuropathy. *Genes (Basel).* **2020**, *11*, 1426, doi:10.3390/genes11121426.

Letko, A.; Strugnell, B.; Häfliger, I.M.; Paris, J.M.; Waine, K.; Drögemüller, C.; Scholes, S. Compound heterozygous *PLA2G6* loss-of-function variants in Swaledale sheep with neuroaxonal dystrophy. *Mol. Genet. Genom.* **2020**, doi:10.1007/s00438-020-01742-1.

Jacinto, J.G.P.; Häfliger, I.M.; Veiga, I.M.B.; **Letko, A.**; Benazzi, C.; Bolcato, M.; Drögemüller, C. A Heterozygous Missense Variant in the *COL5A2* in Holstein Cattle Resembling the Classical Ehlers–Danlos Syndrome. *Animals (Basel).* **2020**, *10*, 2002, doi:10.3390/ani10112002.

Drögemüller, M.; **Letko, A.**; Matiasek, K.; Jagannathan, V.; Corlazzoli, D.; Rosati, M.; Jurina, K.; Medl, S.; Gödde, T.; Rupp, S.; Fischer, A.; Luján Feliu-Pascual, A.; Drögemüller, C. *SLC19A3* Loss-of-Function Variant in Yorkshire Terriers with Leigh-Like Subacute Necrotizing Encephalopathy. *Genes (Basel).* **2020**, *11*, 1215, doi:10.3390/genes11101215.

Letko, A.; Minor, K.M.; Jagannathan, V.; Seefried, F.R.; Mickelson, J.R.; Oliehoek, P.; Drögemüller, C. Genomic diversity and population structure of the Leonberger dog breed. *Genet. Sel. Evol.* **2020**, *52*, 61, doi:10.1186/s12711-020-00581-3.

Brunetti, B.; Muscatello, L. V.; **Letko, A.**; Papa, V.; Cenacchi, G.; Grillini, M.; Murgiano, L.; Jagannathan, V.; Drögemüller, C. X-Linked Duchenne-Type Muscular Dystrophy in Jack Russell Terrier Associated with a Partial Deletion of the Canine *DMD* Gene. *Genes (Basel).* **2020**, *11*, 1175, doi:10.3390/genes11101175.

Letko, A.; Dijkman, R.; Strugnell, B.; Häfliger, I.M.; Paris, J.M.; Henderson, K.; Geraghty, T.; Orr, H.; Scholes, S.; Drögemüller, C. Deleterious *AGXT* Missense Variant Associated with Type 1 Primary Hyperoxaluria (PH1) in Zwartbles Sheep. *Genes (Basel).* **2020**, *11*, 1147, doi:10.3390/genes11101147.

Jacinto, J.G.P.; Häfliger, I.M.; **Letko, A.**; Drögemüller, C.; Agerholm, J.S. A large deletion in the *COL2A1* gene expands the spectrum of pathogenic variants causing bulldog calf syndrome in cattle. *Acta Vet. Scand.* **2020**, *62*, 49, doi:10.1186/s13028-020-00548-w.

Vernau, K.M.; Struys, E.; **Letko, A.**; Woolard, K.D.; Aguilar, M.; Brown, E.A.; Cissell, D.D.; Dickinson, P.J.; Shelton, G.D.; Broome, M.R.; Gibson, K.M.; Pearl, P.L.; König, F.; Van Winkle, T.J.; O'Brien, D.; Roos, B.; Matiasek, K.; Jagannathan, V.; Drögemüller, C.; Mansour, T.A.; Brown, C.T.; Bannasch, D.L. A Missense Variant in *ALDH5A1* Associated with Canine Succinic Semialdehyde Dehydrogenase Deficiency (SSADHD) in the Saluki Dog. *Genes (Basel).* **2020**, *11*, 1033, doi:10.3390/genes11091033.

Paris, J.M.; **Letko, A.**; Häfliger, I.M.; Ammann, P.; Drögemüller, C. Ear type in sheep is associated with the *MSRB3* locus. *Anim. Genet.* **2020**, *51*, 968–972, doi:10.1111/age.12994.

Batcher, K.; Dickinson, P.; Maciejczyk, K.; Brzeski, K.; Rasouliha, S.H.; **Letko, A.**; Drögemüller, C.; Leeb, T.; Bannasch, D. Multiple *FGF4* Retrocopies Recently Derived within Canids. *Genes (Basel).* **2020**, *11*, 839, doi:10.3390/genes11080839.

Häfliger, I.M.; **Letko, A.**; Murgiano, L.; Drögemüller, C. De novo stop-lost germline mutation in *FGFR3* causes severe chondrodysplasia in the progeny of a Holstein bull. *Anim. Genet.* **2020**,

51, 466–469, doi:10.1111/age.12934.

Simon, R.; Lischer, H.E.L.; Pieńkowska-Schelling, A.; Keller, I.; Häfliger, I.M.; **Letko, A.**; Schelling, C.; Lühken, G.; Drögemüller, C. New genomic features of the polled intersex syndrome variant in goats unraveled by long-read whole-genome sequencing. *Anim. Genet.* **2020**, *51*, 439–448, doi:10.1111/age.12918.

Letko, A.; Leuthard, F.; Jagannathan, V.; Corlazzoli, D.; Matiasek, K.; Schweizer, D.; Hytönen, M.K.; Lohi, H.; Leeb, T.; Drögemüller, C. Whole Genome Sequencing Indicates Heterogeneity of Hyperostotic Disorders in Dogs. *Genes (Basel).* **2020**, *11*, 163, doi:10.3390/genes11020163.

Leeb, T.; Leuthard, F.; Jagannathan, V.; Kiener, S.; **Letko, A.**; Roosje, P.; Welle, M.M.; Gailbreath, K.L.; Cannon, A.; Linek, M.; Banovic, F.; Olivry, T.; White, S.D.; Batcher, K.; Bannasch, D.; Minor, K.M.; Mickelson, J.R.; Hytönen, M.K.; Lohi, H.; Mauldin, E.A.; Casal, M.L. A Missense Variant Affecting the C-Terminal Tail of UNC93B1 in Dogs with Exfoliative Cutaneous Lupus Erythematosus (ECLE). *Genes (Basel).* **2020**, *11*, 159, doi:10.3390/genes11020159.

Paris, J.M.; **Letko, A.**; Häfliger, I.M.; Švara, T.; Gombač, M.; Klinc, P.; Škibin, A.; Pogorevc, E.; Drögemüller, C. A de novo variant in *OTX2* in a lamb with otocephaly. *Acta Vet. Scand.* **2020**, *62*, 5, doi:10.1186/s13028-020-0503-z.

Hirter, N.; **Letko, A.**; Häfliger, I.M.; Becker, D.; Greber, D.; Drögemüller, C. A genome-wide significant association on chromosome 15 for congenital entropion in Swiss White Alpine sheep. *Anim. Genet.* **2020**, *51*, 278–283, doi:10.1111/age.12903.

Letko, A.; Ammann, B.; Jagannathan, V.; Henkel, J.; Leuthard, F.; Schelling, C.; Carneiro, M.; Drögemüller, C.; Leeb, T. A deletion spanning the promoter and first exon of the hair cycle-specific *ASIP* transcript isoform in black and tan rabbits. *Anim. Genet.* **2020**, *51*, 137–140, doi:10.1111/age.12881.

Paris, J.M.; **Letko, A.**; Häfliger, I.M.; Ammann, P.; Flury, C.; Drögemüller, C. Identification of two *TYRP1* loss-of-function alleles in Valais Red sheep. *Anim. Genet.* **2019**, *50*, 778–782, doi:10.1111/age.12863.

Letko, A.; Zdora, I.; Hitzler, V.; Jagannathan, V.; Beineke, A.; Möhrke, C.; Drögemüller, C. A de novo in-frame duplication in the *COL1A2* gene in a Lagotto Romagnolo dog with osteogenesis imperfecta. *Anim. Genet.* **2019**, *50*, 786–787, doi:10.1111/age.12843.

Küttel, L.; **Letko, A.**; Häfliger, I.M.; Signer-Hasler, H.; Joller, S.; Hirsbrunner, G.; Mészáros, G.; Sölkner, J.; Flury, C.; Leeb, T.; Drögemüller, C. A complex structural variant at the *KIT* locus in cattle with the Pinzgauer spotting pattern. *Anim. Genet.* **2019**, *50*, 423–429, doi:10.1111/age.12821.

Hédan, B.; Cadieu, E.; Botherel, N.; Dufaure de Citres, C.; **Letko, A.**; Rimbault, M.; Drögemüller, C.; Jagannathan, V.; Derrien, T.; Schmutz, S.; Leeb, T.; André, C. Identification of a Missense Variant in *MFSD12* Involved in Dilution of Phaeomelanin Leading to White or Cream Coat Color in Dogs. *Genes (Basel).* **2019**, *10*, 386, doi:10.3390/genes10050386.

Letko, A.; Dietschi, E.; Nieburg, M.; Jagannathan, V.; Gurtner, C.; Oevermann, A.; Drögemüller, C. A Missense Variant in *SCN8A* in Alpine Dachsbracke Dogs Affected by Spinocerebellar Ataxia. *Genes (Basel).* **2019**, *10*, 362, doi:10.3390/genes10050362.

Grahofer, A.; **Letko, A.**; Häfliger, I.M.; Jagannathan, V.; Ducos, A.; Richard, O.; Peter, V.; Nathues, H.; Drögemüller, C. Chromosomal imbalance in pigs showing a syndromic form of cleft palate. *BMC Genomics* **2019**, *20*, 349, doi:10.1186/s12864-019-5711-4.

Lucot, K.L.; Dickinson, P.J.; Finno, C.J.; Mansour, T.A.; **Letko, A.**; Minor, K.M.; Mickelson, J.R.; Drögemüller, C.; Brown, C.T.; Bannasch, D.L. A Missense Mutation in the Vacuolar Protein Sorting 11 (*VPS11*) Gene Is Associated with Neuroaxonal Dystrophy in Rottweiler Dogs. *G3 (Bethesda)* **2018**, *8*, 2773–2780, doi:10.1534/g3.118.200376.

Dürig, N.; **Letko, A.**; Lepori, V.; Hadji Rasouliha, S.; Loechel, R.; Kehl, A.; Hytönen, M.K.; Lohi, H.; Mauri, N.; Dietrich, J.; Wiedmer, M.; Drögemüller, M.; Jagannathan, V.; Schmutz, S.M.; Leeb, T. Two *MC1R* loss-of-function alleles in cream-coloured Australian Cattle Dogs and white Huskies. *Anim. Genet.* **2018**, *49*, 284–290, doi:10.1111/age.12660.

Minor, K.M.; **Letko, A.**; Becker, D.; Drögemüller, M.; Mandigers, P.J.J.; Bellekom, S.R.; Leegwater, P.A.J.; Stassen, Q.E.M.; Putschbach, K.; Fischer, A.; Flegel, T.; Matiasek, K.; Ekenstedt, K.J.; Furrow, E.; Patterson, E.E.; Platt, S.R.; Kelly, P.A.; Cassidy, J.P.; Shelton, G.D.; Lucot, K.; Bannasch, D.L.; Martineau, H.; Muir, C.F.; Priestnall, S.L.; Henke, D.; Oevermann, A.; Jagannathan, V.; Mickelson, J.R.; Drögemüller, C. Canine *NAPEPLD*-associated models of human myelin disorders. *Sci. Rep.* **2018**, *8*, 5818, doi:10.1038/s41598-018-23938-7.

Letko, A.; Drögemüller, C. Two brown coat colour-associated *TYRP1* variants (b^c and b^d) occur in Leonberger dogs. *Anim. Genet.* **2017**, *48*, 732–733, doi:10.1111/age.12612.

Becker, D.; Minor, K.M.; **Letko, A.**; Ekenstedt, K.J.; Jagannathan, V.; Leeb, T.; Shelton, G.D.; Mickelson, J.R.; Drögemüller, C. A *GJA9* frameshift variant is associated with polyneuropathy in Leonberger dogs. *BMC Genomics* **2017**, *18*, 662, doi:10.1186/s12864-017-4081-z.

Unpublished articles

Bannasch, D.L.; Kaelin, C.B.; **Letko, A.**; Loechel, R.; Hug, P.; Jagannathan, V.; Henkel, J.; Roosje, P.; Hytönen, M.K.; Lohi, H.; Arumilli, M.; DoGA consortium; Minor, K.M.; Mickelson, J.R.; Drögemüller, C.; Barsh, G.S.; Leeb, T. Dog color patterns explained by modular promoters of ancient canid origin. *Nat. Ecol. Evol*. **2021.** *In revision* (https://www.biorxiv.org/content/10.1101/2020.12.21.423812v1).

Conference abstracts

Letko, A.; Grahofer, A.; Häfliger I.M.; Jagannathan, V.; Ducos, A.; Richard, O.; Peter, V.; Nathues, H.; Drögemüller, C. Array genotyping and/or whole genome sequencing facilitates detection of structural variants and chromosomal imbalance in pigs. Conference of the International Society for Animal Genetics, Lleida, Spain, Jul 7-12 **2019.**

Letko, A.; Minor, K.M.; Jagannathan, V.; Seefried, F.R.; Oliehoek, P.; Mickelson, J.R.; Drögemüller, C. Leonberger dogs at risk: predisposition for diseases and inbreeding. 10th International Conference on Canine and Feline Genetics and Genomics, Bern, Switzerland, May 26-29 **2019.**

Letko, A.; Grahofer, A.; Häfliger I.M.; Jagannathan, V.; Ducos, A.; Richard, O.; Peter, V.; Nathues, H.; Drögemüller, C. Array genotyping and/or whole genome sequencing facilitates detection of structural variants and chromosomal imbalance in pigs. Symposium of the Swiss Association of Animal Sciences, Lindau, Switzerland, Apr 16 **2019.**

Letko, A.; Minor, K.M.; Mickelson, J.R.; Seefried, F.R.; Drögemüller, C. Population structure of Leonberger dogs. Companion Animal Genetic Health conference, Edinburgh, UK, May 14-15 **2018.** *Published in*: Selected canine abstracts from the Companion Animal Genetic Health conference 2018 (CAGH 2018): Canine Genetics and Epidemiology. *Canine Genet. Epidemiol.* **2018**, *5*, 7, doi:10.1186/s40575-018-0062-z.

Letko, A.; Minor, K.M.; Mickelson, J.R.; Drögemüller, C. Rare forms of canine neurological disorders in Leonberger breed. The 4th International Workshop of Veterinary Neuroscience, Bern, Switzerland, Feb 16-17 **2018.**

Letko, A.; Minor, K.M.; Becker, D.; Ekenstedt, K.J.; Mickelson, J.R.; Drögemüller, C. Update on polyneuropathy in Leonberger dogs. 9th International Conference on Canine and Feline Genetics and Genomics, Saint Paul, Minnesota, USA, May 21-24 **2017.**

References

1. Ostrander, E.; Ruvinsky, A. *The genetics of the dog*; Ostrander, E.A., Ruvinsky, A., Eds.; 2nd ed.; CABI: Wallingford, **2012**; ISBN 9781845939403.
2. Mellersh, C. Give a dog a genome. *Vet. J.* **2008**, *178*, 46–52, doi:10.1016/j.tvjl.2007.06.029.
3. Frantz, L.A.F.; Mullin, V.E.; Pionnier-Capitan, M.; Lebrasseur, O.; Ollivier, M.; Perri, A.; Linderholm, A.; Mattiangeli, V.; Teasdale, M.D.; Dimopoulos, E.A.; *et al*. Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science* **2016**, *352*, 1228–1231, doi:10.1126/science.aaf3161.
4. Larson, G.; Karlsson, E.K.; Perri, A.; Webster, M.T.; Ho, S.Y.W.; Peters, J.; Stahl, P.W.; Piper, P.J.; Lingaas, F.; Fredholm, M.; *et al*. Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc. Natl. Acad. Sci.* **2012**, *109*, 8878–8883, doi:10.1073/pnas.1203005109.
5. Perri, A. A wolf in dog's clothing: Initial dog domestication and Pleistocene wolf variation. *J. Archaeol. Sci.* **2016**, *68*, 1–4, doi:10.1016/j.jas.2016.02.003.
6. Nielsen, R. Molecular Signatures of Natural Selection. *Annu. Rev. Genet.* **2005**, *39*, 197–218, doi:10.1146/annurev.genet.39.073003.112420.
7. Vonholdt, B.M.; Pollinger, J.P.; Lohmueller, K.E.; Han, E.; Parker, H.G.; Quignon, P.; Degenhardt, J.D.; Boyko, A.R.; Earl, D.A.; Auton, A.; *et al*. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **2010**, *464*, 898–902, doi:10.1038/nature08837.
8. Ostrander, E.A.; Wayne, R.K.; Freedman, A.H.; Davis, B.W. Demographic history, selection and functional diversity of the canine genome. *Nat. Rev. Genet.* **2017**, *18*, 705–720, doi:10.1038/nrg.2017.67.
9. Ostrander, E.A.; Wayne, R.K. The canine genome. *Genome Res.* **2005**, *15*, 1706–1716, doi:10.1101/gr.3736605.
10. Fédération Cynologique Internationale Available online: http://www.fci.be/en/ (accessed on Feb 23, 2021).
11. Andersson, L. Domestic animals as models for biomedical research. *Ups. J. Med. Sci.* **2016**, *121*, 1–11, doi:10.3109/03009734.2015.1091522.
12. Hytönen, M.K.; Lohi, H. Canine models of human rare disorders. *Rare Dis.* **2016**, *4*, 1–6, doi:10.1080/21675511.2016.1241362.
13. Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) Available online: https://omim.org/ (accessed on Mar 18, 2020).
14. Online Mendelian Inheritance in Animals, OMIA. Sydney School of Veterinary Science Available online: https://omia.org/ (accessed on Mar 18, 2020).
15. Gough, A.; Thomas, A.; O'Neill, D. *Breed predispositions to disease in dogs and cats: Third Edition*; **2018**; ISBN 9781119225584.
16. Clamp, M.; Fry, B.; Kamal, M.; Xie, X.; Cuff, J.; Lin, M.F.; Kellis, M.; Lindblad-Toh, K.; Lander, E.S. Distinguishing protein-coding and noncoding genes in the human genome. *Proc. Natl. Acad. Sci.* **2007**, *104*, 19428–19433, doi:10.1073/pnas.0709013104.
17. Lindblad-Toh, K.; Wade, C.M.; Mikkelsen, T.S.; Karlsson, E.K.; Jaffe, D.B.; Kamal, M.; Clamp, M.; Chang, J.L.; Kulbokas, E.J.; Zody, M.C.; *et al*. Genomic sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **2005**, *438*, 803–819, doi:10.1038/nature04338.

18. Shaffer, L.G. Special issue on canine genetics: animal models for human disease and gene therapies, new discoveries for canine inherited diseases, and standards and guidelines for clinical genetic testing for domestic dogs. *Hum. Genet.* **2019**, *138*, 437–440, doi:10.1007/s00439-019-02025-5.

19. Ashley, E.A. Towards precision medicine. *Nat. Rev. Genet.* **2016**, *17*, 507–522, doi:10.1038/nrg.2016.86.

20. Jagannathan, V.; Drögemüller, C.; Leeb, T.; Dog Biomedical Variant Database Consortium, (DBVDC). A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. *Anim. Genet.* **2019**, *50*, 695–704, doi:10.1111/age.12834.

21. Vallat, J.-M.; Goizet, C.; Tazir, M.; Couratier, P.; Magy, L.; Mathis, S. Classifications of neurogenetic diseases: An increasingly complex problem. *Rev. Neurol. (Paris).* **2016**, *172*, 339–349, doi:10.1016/j.neurol.2016.04.005.

22. Sisó, S.; Hanzlíček, D.; Fluehmann, G.; Kathmann, I.; Tomek, A.; Papa, V.; Vandevelde, M. Neurodegenerative diseases in domestic animals: A comparative review. *Vet. J.* **2006**, *171*, 20–38, doi:10.1016/j.tvjl.2004.08.015.

23. Katona, I.; Weis, J. Diseases of the peripheral nerves. In *Handbook of Clinical Neurology*; Elsevier B.V., **2018**; Vol. 145, pp. 453–474; ISBN 9780128023952.

24. Willemsen, M.A.; Harting, I.; Wevers, R.A. Neurometabolic disorders. *Neurol. Clin. Pract.* **2016**, *6*, 348–357, doi:10.1212/CPJ.0000000000000266.

25. Braund, K.G.; Neuropathic Disorders. In *Clinical Neurology in Small Animals - Localization, Diagnosis and Treatment*; Vite, C.H., Ed.; International Veterinary Information Service: Ithaca, New York, USA, **2003**.

26. Correard, S.; Plassais, J.; Lagoutte, L.; Botherel, N.; Thibaud, J.-L.; Hédan, B.; Richard, L.; Lia, A.-S.; Delague, V.; Mège, C.; *et al*. Canine neuropathies: powerful spontaneous models for human hereditary sensory neuropathies. *Hum. Genet.* **2019**, *138*, 455–466, doi:10.1007/s00439-019-02003-x.

27. Fridman, V.; Reilly, M. Inherited Neuropathies. *Semin. Neurol.* **2015**, *35*, 407–423, doi:10.1055/s-0035-1558981.

28. Rudnik-Schöneborn, S.; Auer-Grumbach, M.; Senderek, J. Hereditary Neuropathies: Update 2017. *Neuropediatrics* **2017**, *48*, 282–293, doi:10.1055/s-0037-1603518.

29. Reilly, M.M.; Murphy, S.M.; Laurá, M. Charcot-Marie-Tooth disease. *J. Peripher. Nerv. Syst.* **2011**, *16*, 1–14, doi:10.1111/j.1529-8027.2011.00324.x.

30. Morena, J.; Gupta, A.; Hoyle, J.C. Charcot-Marie-Tooth: From Molecules to Therapy. *Int. J. Mol. Sci.* **2019**, *20*, 3419, doi:10.3390/ijms20143419.

31. Granger, N. Canine inherited motor and sensory neuropathies: An updated classification in 22 breeds and comparison to Charcot-Marie-Tooth disease. *Vet. J.* **2011**, *188*, 274–285, doi:10.1016/j.tvjl.2010.06.003.

32. Kitshoff, A.M.; Van Goethem, B.; Stegen, L.; Vandekerckhov, P.; De Rooster, H. Laryngeal paralysis in dogs: An update on recent knowledge. *J. S. Afr. Vet. Assoc.* **2013**, *84*, 1–9, doi:10.4102/jsava.v84i1.909.

33. Monnet, E. Surgical Treatment of Laryngeal Paralysis. *Vet. Clin. North Am. - Small Anim. Pract.* **2016**, *46*, 709–717, doi:10.1016/j.cvsm.2016.02.003.

34. Coates, J.R.; O'Brien, D.P. Inherited peripheral neuropathies in dogs and cats. *Vet. Clin. North Am. Small Anim. Pract.* **2004**, *34*, 1361–1401, doi:10.1016/j.cvsm.2004.05.011.

35. Waxman, S.G.; Ritchie, J.M. Molecular dissection of the myelinated axon. *Ann. Neurol.* **1993**, *33*, 121–136, doi:10.1002/ana.410330202.

36. Duncan, I.D.; Radcliff, A.B. Inherited and acquired disorders of myelin: The underlying myelin pathology. *Exp. Neurol.* **2016**, *283*, 452–475, doi:10.1016/j.expneurol.2016.04.002.

37. Kaye, E.M. Update on genetic disorders affecting white matter. *Pediatr. Neurol.* **2001**, *24*, 11–24, doi:10.1016/S0887-8994(00)00232-0.

38. Vanderver, A.; Prust, M.; Tonduti, D.; Mochel, F.; Hussey, H.M.; Helman, G.; Garbern, J.; Eichler, F.; Labauge, P.; Aubourg, P.; *et al*. Case definition and classification of leukodystrophies and leukoencephalopathies. *Mol. Genet. Metab.* **2015**, *114*, 494–500, doi:10.1016/j.ymgme.2015.01.006.

39. Salsano, E. Leukodystrophy or genetic leukoencephalopathy? Nature does not make leaps. *Mol. Genet. Metab.* **2015**, *114*, 491–493, doi:10.1016/j.ymgme.2015.02.005.

40. Schiffmann, R.; Van Der Knaap, M.S. Invited Article: An MRI-based approach to the diagnosis of white matter disorders. *Neurology* **2009**, *72*, 750–759, doi:10.1212/01.wnl.0000343049.00540.c8.

41. Wang, X.; He, F.; Yin, F.; Chen, C.; Wu, L.; Yang, L.; Peng, J. The use of targeted genomic capture and massively parallel sequencing in diagnosis of Chinese Leukoencephalopathies. *Sci. Rep.* **2016**, *6*, 35936, doi:10.1038/srep35936.

42. Robert, N. Neurologic Disorders in Cheetahs and Snow Leopards. In *Zoo and Wild Animal Medicine*; Fowler, M.E., Miller, R.E., Eds.; Elsevier, **2008**; ISBN 9781416040477.

43. Andersen, H.A.; Palludan, B. Leucodystrophy in mink. *Acta Neuropathol.* **1968**, *11*, 347–360, doi:10.1007/BF00686731.

44. Story, B.D.; Miller, M.E.; Bradbury, A.M.; Million, E.D.; Duan, D.; Taghian, T.; Faissler, D.; Fernau, D.; Beecy, S.J.; Gray-Edwards, H.L. Canine Models of Inherited Musculoskeletal and Neurodegenerative Diseases. *Front. Vet. Sci.* **2020**, *7*, 80, doi:10.3389/fvets.2020.00080.

45. Coates, J.R.; Wininger, F.A. Canine degenerative myelopathy. *Vet. Clin. North Am. - Small Anim. Pract.* **2010**, *40*, 929–950.

46. Davies, D.R.; Irwin, P.J. Degenerative neurological and neuromuscular disease in young rottweilers. *J. Small Anim. Pract.* **2003**, *44*, 388–394, doi:10.1111/j.1748-5827.2003.tb00173.x.

47. Oevermann, A.; Bley, T.; Konar, M.; Lang, J.; Vandevelde, M. A Novel Leukoencephalomyelopathy of Leonberger Dogs. *J. Vet. Intern. Med.* **2008**, 22, 467–471, doi:10.1111/j.1939-1676.2008.0068.x.

48. Hayflick, S.J.; Kurian, M.A.; Hogarth, P. Neurodegeneration with brain iron accumulation. *Handb. Clin. Neurol.* **2018**, *147*, 293–305, doi:10.1016/B978-0-444-63233-3.00019-1.

49. Gregory, A.; Polster, B.J.; Hayflick, S.J. Clinical and genetic delineation of neurodegeneration with brain iron accumulation. *J. Med. Genet.* **2008**, *46*, 73–80, doi:10.1136/jmg.2008.061929.

50. Carrilho, I.; Santos, M.; Guimarães, A.; Teixeira, J.; Chorão, R.; Martins, M.; Dias, C.; Gregory, A.; Westaway, S.; Nguyen, T.; *et al*. Infantile neuroaxonal dystrophy: What's most important for the diagnosis? *Eur. J. Paediatr. Neurol.* **2008**, *12*, 491–500, doi:10.1016/j.ejpn.2008.01.005.

51. Letko, A.; Strugnell, B.; Häfliger, I.M.; Paris, J.M.; Waine, K.; Drögemüller, C.; Scholes, S. Compound heterozygous *PLA2G6* loss-of-function variants in Swaledale sheep with neuroaxonal dystrophy. *Mol. Genet. Genom.* **2021**, *296*, 235–242, doi:10.1007/s00438-020-01742-1.

52. Fyfe, J.C.; Al-Tamimi, R.A.; Liu, J.; Schäffer, A.A.; Agarwala, R.; Henthorn, P.S. A novel mitofusin 2 mutation causes canine fetal-onset neuroaxonal dystrophy. *Neurogenetics* **2011**, *12*, 223–232, doi:10.1007/s10048-011-0285-6.

53. Hahn, K.; Rohdin, C.; Jagannathan, V.; Wohlsein, P.; Baumgärtner, W.; Seehusen, F.; Spitzbarth, I.; Grandon, R.; Drögemüller, C.; Jäderlund, K.H. *TECPR2* associated neuroaxonal dystrophy in Spanish water dogs. *PLoS One* **2015**, *10*, 1–18, doi:10.1371/journal.pone.0141824.

54. Lucot, K.L.; Dickinson, P.J.; Finno, C.J.; Mansour, T.A.; Letko, A.; Minor, K.M.; Mickelson, J.R.; Drögemüller, C.; Brown, C.T.; Bannasch, D.L. A Missense Mutation in the *Vacuolar Protein Sorting 11* (*VPS11*) Gene Is Associated with Neuroaxonal Dystrophy in Rottweiler Dogs. *G3 (Bethesda)* **2018**, *8*, 2773–2780, doi:10.1534/g3.118.200376.

55. Synofzik, M.; Schüle, R. Overcoming the divide between ataxias and spastic paraplegias: Shared phenotypes, genes, and pathways. *Mov. Disord.* **2017**, *32*, 332–345, doi:10.1002/mds.26944.

56. Manto, M.; Marmolino, D. Cerebellar ataxias. *Curr. Opin. Neurol.* **2009**, *22*, 419–429, doi:10.1097/WCO.0b013e32832b9897.

57. Urkasemsin, G.; Olby, N.J. Canine Hereditary Ataxia. *Vet. Clin. North Am. Small Anim. Pract.* **2014**, *44*, 1075–1089, doi:10.1016/j.cvsm.2014.07.005.

58. El-Hattab, A.W. Inborn Errors of Metabolism. *Clin. Perinatol.* **2015**, *42*, 413–439, doi:10.1016/j.clp.2015.02.010.

59. Kamboj, M. Clinical Approach to the Diagnoses of Inborn Errors of Metabolism. *Pediatr. Clin. North Am.* **2008**, *55*, 1113–1127, doi:10.1016/j.pcl.2008.07.004.

60. Saudubray, J.-M.; Garcia-Cazorla, A. An overview of inborn errors of metabolism affecting the brain: from neurodevelopment to neurodegenerative disorders. *Dialogues Clin. Neurosci.* **2018**, *20*, 301–326, doi:10.31887/DCNS.2018.20.4/jmsaudubray.

61. Sewell, A.C.; Haskins, M.E.; Giger, U. Inherited metabolic disease in companion animals: Searching for nature's mistakes. *Vet. J.* **2007**, *174*, 252–259, doi:10.1016/j.tvjl.2006.08.017.

62. Leeb, T. Animal DNA diagnostics – Personal genomics for our pets and livestock is at the Horizon. *Mol. Cell. Probes* **2012**, *26*, 223, doi:10.1016/j.mcp.2012.09.001.

63. Mellersh, C.S.; Langston, A.A.; Acland, G.M.; Fleming, M.A.; Ray, K.; Wiegand, N.A.; Francisco, L. V.; Gibbs, M.; Aguirre, G.D.; Ostrander, E.A. A Linkage Map of the Canine Genome. *Genomics* **1997**, *46*, 326–336, doi:10.1006/geno.1997.5098.

64. Karlsson, E.K.; Lindblad-Toh, K. Leader of the pack: gene mapping in dogs and other model organisms. *Nat. Rev. Genet.* **2008**, *9*, 713–725, doi:10.1038/nrg2382.

65. Leegwater, P.A.; van Hagen, M.A.; van Oost, B.A. Localization of White Spotting Locus in Boxer Dogs on CFA20 by Genome-Wide Linkage Analysis with 1500 SNPs. *J. Hered.* **2007**, *98*, 549–552, doi:10.1093/jhered/esm022.

66. Drögemüller, C.; Becker, D.; Brunner, A.; Haase, B.; Kircher, P.; Seeliger, F.; Fehr, M.; Baumann, U.; Lindblad-Toh, K.; Leeb, T. A Missense Mutation in the *SERPINH1* Gene in Dachshunds with Osteogenesis Imperfecta. *PLoS Genet.* **2009**, *5*, e1000579, doi:10.1371/journal.pgen.1000579.

67. Karlsson, E.K.; Baranowska, I.; Wade, C.M.; Salmon Hillbertz, N.H.C.; Zody, M.C.; Anderson, N.; Biagi, T.M.; Patterson, N.; Pielberg, G.R.; Kulbokas, E.J.; *et al*. Efficient mapping of mendelian traits in dogs through genome-wide association. *Nat. Genet.* **2007**, *39*, 1321–1328, doi:10.1038/ng.2007.10.

68. Drögemüller, C.; Karlsson, E.K.; Hytönen, M.K.; Perloski, M.; Dolf, G.; Sainio, K.; Lohi,

H.; Lindblad-Toh, K.; Leeb, T. A Mutation in Hairless Dogs Implicates *FOXI3* in Ectodermal Development. *Science* **2008**, *321*, 1462–1462, doi:10.1126/science.1162525.

69. McGinn, S.; Gut, I.G. DNA sequencing – spanning the generations. *N. Biotechnol.* **2013**, *30*, 366–372, doi:10.1016/j.nbt.2012.11.012.

70. Drögemüller, M.; Jagannathan, V.; Howard, J.; Bruggmann, R.; Drögemüller, C.; Ruetten, M.; Leeb, T.; Kook, P.H. A frameshift mutation in the *cubilin* gene (*CUBN*) in Beagles with Imerslund-Gräsbeck syndrome (selective cobalamin malabsorption). *Anim. Genet.* **2014**, *45*, 148–150, doi:10.1111/age.12094.

71. Guevar, J.; Olby, N.J.; Meurs, K.M.; Yost, O.; Friedenberg, S.G. Deafness and vestibular dysfunction in a Doberman Pinscher puppy associated with a mutation in the *PTPRQ* gene. *J. Vet. Intern. Med.* **2018**, *32*, 665–669, doi:10.1111/jvim.15060.

72. Lepori, V.; Mühlhause, F.; Sewell, A.C.; Jagannathan, V.; Janzen, N.; Rosati, M.; Alves de Sousa, F.M.M.; Tschopp, A.; Schüpbach, G.; Matiasek, K.; *et al*. A Nonsense Variant in the *ACADVL* Gene in German Hunting Terriers with Exercise Induced Metabolic Myopathy. *G3 (Bethesda)* **2018**, *8*, 1545–1554, doi:10.1534/g3.118.200084.

73. Drögemüller, M.; Jagannathan, V.; Becker, D.; Drögemüller, C.; Schelling, C.; Plassais, J.; Kaerle, C.; Dufaure de Citres, C.; Thomas, A.; Müller, E.J.; *et al*. A Mutation in the *FAM83G* Gene in Dogs with Hereditary Footpad Hyperkeratosis (HFH). *PLoS Genet.* **2014**, *10*, e1004370, doi:10.1371/journal.pgen.1004370.

74. Gerber, M.; Fischer, A.; Jagannathan, V.; Drögemüller, M.; Drögemüller, C.; Schmidt, M.J.; Bernardino, F.; Manz, E.; Matiasek, K.; Rentmeister, K.; *et al*. A Deletion in the *VLDLR* Gene in Eurasier Dogs with Cerebellar Hypoplasia Resembling a Dandy-Walker-Like Malformation (DWLM). *PLoS One* **2015**, *10*, e0108917, doi:10.1371/journal.pone.0108917.

75. Wielaender, F.; Sarviaho, R.; James, F.; Hytönen, M.K.; Cortez, M.A.; Kluger, G.; Koskinen, L.L.E.; Arumilli, M.; Kornberg, M.; Bathen-Noethen, A.; *et al*. Generalized myoclonic epilepsy with photosensitivity in juvenile dogs caused by a defective DIRAS family GTPase 1. *Proc. Natl. Acad. Sci.* **2017**, *114*, 2669–2674, doi:10.1073/pnas.1614478114.

76. Letko, A.; Zdora, I.; Hitzler, V.; Jagannathan, V.; Beineke, A.; Möhrke, C.; Drögemüller, C. A de novo in-frame duplication in the *COL1A2* gene in a Lagotto Romagnolo dog with osteogenesis imperfecta. *Anim. Genet.* **2019**, *50*, 786–787, doi:10.1111/age.12843.

77. Murgiano, L.; Becker, D.; Torjman, D.; Niggel, J.K.; Milano, A.; Cullen, C.; Feng, R.; Wang, F.; Jagannathan, V.; Pearce-Kelling, S.; *et al*. Complex Structural *PPT1* Variant Associated with Non-syndromic Canine Retinal Degeneration. *G3 (Bethesda)* **2018**, *9*, g3.200859.2018, doi:10.1534/g3.118.200859.

78. Mauri, N.; Kleiter, M.; Dietschi, E.; Leschnik, M.; Högler, S.; Wiedmer, M.; Dietrich, J.; Henke, D.; Steffen, F.; Schuller, S.; *et al*. A SINE Insertion in *ATP1B2* in Belgian Shepherd Dogs Affected by Spongy Degeneration with Cerebellar Ataxia (SDCA2). *G3 (Bethesda)* **2017**, *7*, 2729–2737, doi:10.1534/g3.117.043018.

79. Hirz, M.; Drögemüller, M.; Schänzer, A.; Jagannathan, V.; Dietschi, E.; Goebel, H.H.; Hecht, W.; Laubner, S.; Schmidt, M.J.; Steffen, F.; *et al*. Neuronal ceroid lipofuscinosis (NCL) is caused by the entire deletion of *CLN8* in the Alpenländische Dachsbracke dog. *Mol. Genet. Metab.* **2017**, *120*, 269–277, doi:10.1016/j.ymgme.2016.12.007.

80. Lequarré, A.S.; Andersson, L.; André, C.; Fredholm, M.; Hitte, C.; Leeb, T.; Lohi, H.; Lindblad-Toh, K.; Georges, M. LUPA: A European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. *Vet. J.* **2011**, *189*, 155–159.

81. Friedenberg, S.G.; Meurs, K.M. Genotype imputation in the domestic dog. *Mamm. Genome* **2016**, *27*, 485–494, doi:10.1007/s00335-016-9636-9.

82. Hayward, J.J.; White, M.E.; Boyle, M.; Shannon, L.M.; Casal, M.L.; Castelhano, M.G.; Center, S.A.; Meyers-Wallen, V.N.; Simpson, K.W.; Sutter, N.B.; *et al*. Imputation of canine genotype array data using 365 whole-genome sequences improves power of genome-wide association studies. *PLOS Genet.* **2019**, *15*, e1008003, doi:10.1371/journal.pgen.1008003.

83. Teare, M.D.; Santibanez Koref, M.F. Linkage analysis and the study of Mendelian disease in the era of whole exome and genome sequencing. *Brief. Funct. Genomics* **2014**, *13*, 378–383, doi:10.1093/bfgp/elu024.

84. Ott, J.; Wang, J.; Leal, S.M. Genetic linkage analysis in the age of whole-genome sequencing. *Nat. Rev. Genet.* **2015**, *16*, 275–284.

85. Sams, A.J.; Boyko, A.R. Fine-Scale Resolution of Runs of Homozygosity Reveal Patterns of Inbreeding and Substantial Overlap with Recessive Disease Genotypes in Domestic Dogs. *G3 (Bethesda)* **2019**, *9*, 117–123, doi:10.1534/g3.118.200836.

86. Vahidnezhad, H.; Youssefian, L.; Jazayeri, A.; Uitto, J. Research Techniques Made Simple: Genome-Wide Homozygosity/Autozygosity Mapping Is a Powerful Tool for Identifying Candidate Genes in Autosomal Recessive Genetic Diseases. *J. Invest. Dermatol.* **2018**, *138*, 1893–1900, doi:10.1016/j.jid.2018.06.170.

87. Curik, I.; Ferenčaković, M.; Sölkner, J. Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest. Sci.* **2014**, *166*, 26–34, doi:10.1016/j.livsci.2014.05.034.

88. Gondro, C.; Porto-Neto, L.R.; Lee, S.H. Genome-Wide Association Studies and Genomic Prediction. In *Methods in Molecular Biology*; Gondro, C., van der Werf, J., Hayes, B., Eds.; Humana Press: Totowa, NJ, **2013**; Vol. 1019; ISBN 978-1-62703-446-3.

89. Ekenstedt, K.J.; Becker, D.; Minor, K.M.; Shelton, G.D.; Patterson, E.E.; Bley, T.; Oevermann, A.; Bilzer, T.; Leeb, T.; Drögemüller, C.; *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, doi:10.1371/journal.pgen.1004635.

90. Price, A.L.; Zaitlen, N.A.; Reich, D.; Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **2010**, *11*, 459–463, doi:10.1038/nrg2813.

91. Mealey, K.L.; Martinez, S.E.; Villarino, N.F.; Court, M.H. Personalized medicine: going to the dogs? *Hum. Genet.* **2019**, *138*, 467–481, doi:10.1007/s00439-019-02020-w.

92. Gauthier, J.; Vincent, A.T.; Charette, S.J.; Derome, N. A brief history of bioinformatics. *Brief. Bioinform.* **2019**, *20*, 1981–1996, doi:10.1093/bib/bby063.

93. Kumar, K.R.; Cowley, M.J.; Davis, R.L. Next-Generation Sequencing and Emerging Technologies. *Semin. Thromb. Hemost.* **2019**, *45*, 661–673, doi:10.1055/s-0039-1688446.

94. Koboldt, D.C.; Steinberg, K.M.; Larson, D.E.; Wilson, R.K.; Mardis, E.R. The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell* **2013**, *155*, 27–38, doi:10.1016/j.cell.2013.09.006.

95. Depristo, M.A.; Banks, E.; Poplin, R.; Garimella, K. V.; Maguire, J.R.; Hartl, C.; Philippakis, A.A.; Del Angel, G.; Rivas, M.A.; Hanna, M.; *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **2011**, *43*, 491–501, doi:10.1038/ng.806.

96. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.;

Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; *et al*. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303, doi:10.1101/gr.107524.110.

97. Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; Kaul, R.; *et al*. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74, doi:10.1038/nature11247.

98. Plassais, J.; Kim, J.; Davis, B.W.; Karyadi, D.M.; Hogan, A.N.; Harris, A.C.; Decker, B.; Parker, H.G.; Ostrander, E.A. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat. Commun.* **2019**, *10*, 1–14, doi:10.1038/s41467-019-09373-w.

99. Ostrander, E.A.; Wang, G.D.; Larson, G.; Vonholdt, B.M.; Davis, B.W.; Jagannathan, V.; Hitte, C.; Wayne, R.K.; Zhang, Y.P.; André, C.; *et al*. Dog10K: An international sequencing effort to advance studies of canine domestication, phenotypes and health. *Natl. Sci. Rev.* 2019, *6*, 810–824.

100. Dog Genome Assembled. National Human Genome Research Institute. Available online: https://www.genome.gov/12511476/2004-advisory-dog-genome-assembled (accessed on Mar 11, 2021).

101. Hoeppner, M.P.; Lundquist, A.; Pirun, M.; Meadows, J.R.S.; Zamani, N.; Johnson, J.; Sundström, G.; Cook, A.; FitzGerald, M.G.; Swofford, R.; *et al*. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One* **2014**, *9*, 91172, doi:10.1371/journal.pone.0091172.

102. *Canis lupus familiaris* Annotation Release 106. National Center for Biotechnology Information. Available online: https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Canis_lupus_familiaris/106/ (accessed on Mar 4, 2021).

103. Dog10K_Boxer_Tasha - Genome - Assembly - NCBI. Dog Genome Sequencing Consortium. Available online: https://www.ncbi.nlm.nih.gov/assembly/GCF_000002285.5/ (accessed on Mar 4, 2021).

104. Halo, J. V.; Pendleton, A.L.; Shen, F.; Doucet, A.J.; Derrien, T.; Hitte, C.; Kirby, L.E.; Myers, B.; Sliwerska, E.; Emery, S.; *et al*. Long-read assembly of a Great Dane genome highlights the contribution of GC-rich sequence and mobile elements to canine genomes. *Proc. Natl. Acad. Sci.* **2021**, *118*, e2016274118, doi:10.1073/pnas.2016274118.

105. Wang, C.; Wallerman, O.; Arendt, M.-L.; Sundström, E.; Karlsson, Å.; Nordin, J.; Mäkeläinen, S.; Pielberg, G.R.; Hanson, J.; Ohlsson, Å.; *et al*. A novel canine reference genome resolves genomic architecture and uncovers transcript complexity. *Commun. Biol.* **2021**, *4*, 185, doi:10.1038/s42003-021-01698-x.

106. Edwards, R.J.; Field, M.A.; Ferguson, J.M.; Dudchenko, O.; Keilwagen, J.; Rosen, B.D.; Johnson, G.S.; Rice, E.; Hillier, L.D.; Hammond, J.M.; *et al*. Chromosome-length genome assembly and structural variations of the primal Basenji dog (*Canis lupus familiaris*) genome. *bioRxiv* **2020**, 2020.11.11.379073, doi:10.1101/2020.11.11.379073.

107. ROS_Cfam_1.0 - Genome - Assembly - NCBI. The Roslin Institute. Available online: https://www.ncbi.nlm.nih.gov/assembly/GCF_014441545.1/ (accessed on Mar 4, 2021).

108. Boccardo, A.; Marelli, S.P.; Pravettoni, D.; Bagnato, A.; Busca, G.A.; Strillacci, M.G. The German Shorthair Pointer Dog Breed (*Canis lupus familiaris*): Genomic Inbreeding and Variability. *Animals (Basel).* **2020**, *10*, 498, doi:10.3390/ani10030498.

109. Mastrangelo, S.; Filippo, B.; Auzino, B.; Marco, R.; Spaterna, A.; Ciampolini, R.

Genome-wide diversity and runs of homozygosity in the “Braque Français, type Pyrénées” dog breed. *BMC Res. Notes* **2018**, *11*, 11–16, doi:10.1186/s13104-017-3112-9.

110. Mortlock, S.A.; Khatkar, M.S.; Williamson, P. Comparative analysis of genome diversity in bullmastiff dogs. *PLoS One* **2016**, *11*, 1–17, doi:10.1371/journal.pone.0147941.

111. Soh, P.X.Y.; Hsu, W.T.; Khatkar, M.S.; Williamson, P. Evaluation of genetic diversity and management of disease in Border Collie dogs. *Sci. Rep.* **2021**, *11*, 6243, doi:10.1038/s41598-021-85262-x.

112. Friedenberg, S.G.; Meurs, K.M.; Mackay, T.F.C. Evaluation of artificial selection in Standard Poodles using whole-genome sequencing. *Mamm. Genome* **2016**, *27*, 599–609, doi:10.1007/s00335-016-9660-9.

113. Metzger, J.; Pfahler, S.; Distl, O. Variant detection and runs of homozygosity in next generation sequencing data elucidate the genetic background of Lundehund syndrome. *BMC Genomics* **2016**, *17*, doi:10.1186/s12864-016-2844-6.

114. Mastrangelo, S.; Biscarini, F.; Tolone, M.; Auzino, B.; Ragatzu, M.; Spaterna, A.; Ciampolini, R. Genomic characterization of the Braque Français type Pyrénées dog and relationship with other breeds. *PLoS One* **2018**, *13*, e0208548, doi:10.1371/journal.pone.0208548.

115. Caniglia, R.; Fabbri, E.; Hulva, P.; Bolfíková, B.Č.; Jindřichová, M.; Stronen, A.V.; Dykyy, I.; Camatta, A.; Carnier, P.; Randi, E.; *et al*. Wolf outside, dog inside? The genomic make-up of the Czechoslovakian Wolfdog. *BMC Genomics* **2018**, *19*, 533, doi:10.1186/s12864-018-4916-2.

116. Pfahler, S.; Distl, O. Effective Population Size, Extended Linkage Disequilibrium and Signatures of Selection in the Rare Dog Breed Lundehund. *PLoS One* **2015**, *10*, e0122680, doi:10.1371/journal.pone.0122680.

117. O’Brien, D.P.; Leeb, T. DNA Testing in Neurologic Diseases. *J. Vet. Intern. Med.* **2014**, *28*, 1186–1198, doi:10.1111/jvim.12383.

118. Windig, J.J.; Doekes, H.P. Limits to genetic rescue by outcross in pedigree dogs. *J. Anim. Breed. Genet.* **2018**, *135*, 238–248, doi:10.1111/jbg.12330.

119. Kleopa, K.A.; Sargiannidou, I. Connexins, gap junctions and peripheral neuropathy. *Neurosci. Lett.* **2015**, *596*, 27–32.

120. Abrams, C.K.; Scherer, S.S. Gap junctions in inherited human disorders of the central nervous system. *Biochim. Biophys. Acta - Biomembr.* **2012**, *1818*, 2030–2047.

121. Freed, A.S.; Weiss, M.D.; Malouf, E.A.; Hisama, F.M. *CNTNAP1* mutations in an adult with Charcot Marie Tooth disease. *Muscle and Nerve* **2019**, *60*, E28–E30, doi:10.1002/mus.26658.

122. Low, K.; Stals, K.; Caswell, R.; Clayton-Smith, J.; Donaldson, A.; Foulds, N.; Splitt, M.; Norman, A.; Urankar, K.; Vijayakumar, K.; *et al*. *CNTNAP1*: Extending the phenotype of congenital hypomyelinating neuropathy in 6 further patients. *Neuromuscul. Disord.* **2017**, *27*, S148, doi:10.1016/j.nmd.2017.06.201.

123. Boyko, A.R.; Brooks, S.A.; Behan-Braman, A.; Castelhano, M.; Corey, E.; Oliveira, K.C.; Swinburne, J.E.; Todhunter, R.J.; Zhang, Z.; Ainsworth, D.M.; *et al*. Genomic analysis establishes correlation between growth and laryngeal neuropathy in Thoroughbreds. *BMC Genomics* **2014**, *15*, 1–9, doi:10.1186/1471-2164-15-259.

124. Wiedmer, M.; Oevermann, A.; Borer-Germann, S.E.; Gorgas, D.; Shelton, G.D.; Drögemüller, M.; Jagannathan, V.; Henke, D.; Leeb, T. A *RAB3GAP1* SINE Insertion in Alaskan Huskies with Polyneuropathy, Ocular Abnormalities, and Neuronal Vacuolation (POANV) Resembling Human Warburg Micro Syndrome 1 (WARBM1).

G3 (Bethesda) **2016**, *6*, 255–262, doi:10.1534/g3.115.022707.

125. Mhlanga-Mutangadura, T.; Johnson, G.S.; Ashwini, A.; Shelton, G.D.; Wennogle, S.A.; Johnson, G.C.; Kuroki, K.; O'Brien, D.P. A Homozygous *RAB3GAP1*:c.743delC Mutation in Rottweilers with Neuronal Vacuolation and Spinocerebellar Degeneration. *J. Vet. Intern. Med.* **2016**, *30*, 813–818, doi:10.1111/jvim.13921.

126. Lu, H.C.; MacKie, K. An introduction to the endogenous cannabinoid system. *Biol. Psychiatry* 2016, *79*, 516–525.

127. Cork, L.C.; Troncoso, J.C.; Price, D.L.; Stanley, E.F.; Griffin, J.W. Canine Neuroaxonal Dystrophy. *J. Neuropathol. Exp. Neurol.* **1983**, *42*, 286–296, doi:10.1097/00005072-198305000-00006.

128. Plemel, R.L.; Lobingier, B.T.; Brett, C.L.; Angers, C.G.; Nickerson, D.P.; Paulsel, A.; Sprague, D.; Merz, A.J. Subunit organization and Rab interactions of Vps-C protein complexes that control endolysosomal membrane traffic. *Mol. Biol. Cell* **2011**, *22*, 1353–1363, doi:10.1091/mbc.E10-03-0260.

129. Hörtnagel, K.; Krägeloh-Mann, I.; Bornemann, A.; Döcker, M.; Biskup, S.; Mayrhofer, H.; Battke, F.; du Bois, G.; Harzer, K. The second report of a new hypomyelinating disease due to a defect in the *VPS11* gene discloses a massive lysosomal involvement. *J. Inherit. Metab. Dis.* **2016**, *39*, 849–857, doi:10.1007/s10545-016-9961-x.

130. Catterall, W.A. Sodium Channels. *Encycl. Life Sci.* **2006**, 65–71, doi:10.1038/npg.els.0000127.

131. Kearney, J.A.; Buchner, D.A.; De Haan, G.; Adamska, M.; Levin, S.I.; Furay, A.R.; Albin, R.L.; Jones, J.M.; Montal, M.; Stevens, M.J.; *et al*. Molecular and pathological effects of a modifier gene on deficiency of the sodium channel Scn8a (Na(v)1.6). *Hum. Mol. Genet.* **2002**, *11*, 2765–2775, doi:10.1093/hmg/11.22.2765.

132. Wagnon, J.L.; Mencacci, N.E.; Barker, B.S.; Wengert, E.R.; Bhatia, K.P.; Balint, B.; Carecchio, M.; Wood, N.W.; Patel, M.K.; Meisler, M.H. Partial loss-of-function of sodium channel SCN8A in familial isolated myoclonus. *Hum. Mutat.* **2018**, *39*, 965–969, doi:10.1002/humu.23547.

133. Pearl, P.L.; Parviz, M.; Vogel, K.; Schreiber, J.; Theodore, W.H.; Gibson, K.M. Inherited disorders of gamma-aminobutyric acid metabolism and advances in *ALDH5A1* mutation identification. *Dev. Med. Child Neurol.* **2015**, *57*, 611–617, doi:10.1111/dmcn.12668.

134. DiBacco, M.L.; Pop, A.; Salomons, G.S.; Hanson, E.; Roullet, J.-B.; Gibson, K.M.; Pearl, P.L. Novel *ALDH5A1* variants and genotype. *Neurology* **2020**, *95*, e2675–e2682, doi:10.1212/WNL.0000000000010730.

135. Didiasova, M.; Banning, A.; Brennenstuhl, H.; Jung-Klawitter, S.; Cinquemani, C.; Opladen, T.; Tikkanen, R. Succinic Semialdehyde Dehydrogenase Deficiency: An Update. *Cells* **2020**, *9*, 477, doi:10.3390/cells9020477.

136. Wakshlag, J.J.; Lahunta, A. de; Robinson, T.; Cooper, B.J.; Brenner, O.; O'Toole, T.D.; Olson, J.; Beckman, K.B.; Glass, E.; Reynolds, A.J. Subacute necrotising encephalopathy in an Alaskan husky. *J. Small Anim. Pract.* **1999**, *40*, 585–589, doi:10.1111/j.1748-5827.1999.tb03028.x.

137. Baiker, K.; Hofmann, S.; Fischer, A.; Gödde, T.; Medl, S.; Schmahl, W.; Bauer, M.F.; Matiasek, K. Leigh-like subacute necrotising encephalopathy in Yorkshire Terriers: neuropathological characterisation, respiratory chain activities and mitochondrial DNA. *Acta Neuropathol.* **2009**, *118*, 697–709, doi:10.1007/s00401-009-0548-6.

138. Gerards, M.; Sallevelt, S.C.E.H.; Smeets, H.J.M. Leigh syndrome: Resolving the clinical and genetic heterogeneity paves the way for treatment options. *Mol. Genet.*

Metab. **2016**, *117*, 300–312, doi:10.1016/j.ymgme.2015.12.004.

139. Gerards, M.; Kamps, R.; van Oevelen, J.; Boesten, I.; Jongen, E.; de Koning, B.; Scholte, H.R.; de Angst, I.; Schoonderwoerd, K.; Sefiani, A.; *et al*. Exome sequencing reveals a novel Moroccan founder mutation in *SLC19A3* as a new cause of early-childhood fatal Leigh syndrome. *Brain* **2013**, *136*, 882–890, doi:10.1093/brain/awt013.

140. Whitford, W.; Hawkins, I.; Glamuzina, E.; Wilson, F.; Marshall, A.; Ashton, F.; Love, D.R.; Taylor, J.; Hill, R.; Lehnert, K.; *et al*. Compound heterozygous *SLC19A3* mutations further refine the critical promoter region for biotin-thiamine-responsive basal ganglia disease. *Mol. Case Stud.* **2017**, *3*, a001909, doi:10.1101/mcs.a001909.

141. Vernau, K.M.; Runstadler, J.A.; Brown, E.A.; Cameron, J.M.; Huson, H.J.; Higgins, R.J.; Ackerley, C.; Sturges, B.K.; Dickinson, P.J.; Puschner, B.; *et al*. Genome-Wide Association Analysis Identifies a Mutation in the Thiamine Transporter 2 (*SLC19A3*) Gene Associated with Alaskan Husky Encephalopathy. *PLoS One* **2013**, *8*, e57195, doi:10.1371/journal.pone.0057195.

142. Lewis, T.W.; Mellersh, C.S. Changes in mutation frequency of eight Mendelian inherited disorders in eight pedigree dog populations following introduction of a commercial DNA test. *PLoS One* **2019**, *14*, 1–21, doi:10.1371/journal.pone.0209864.

143. Drögemüller, C.; Reichart, U.; Seuberlich, T.; Oevermann, A.; Baumgartner, M.; Kühni Boghenbor, K.; Stoffel, M.H.; Syring, C.; Meylan, M.; Müller, S.; *et al*. An Unusual Splice Defect in the *Mitofusin 2* Gene (*MFN2*) Is Associated with Degenerative Axonopathy in Tyrolean Grey Cattle. *PLoS One* **2011**, *6*, e18931, doi:10.1371/journal.pone.0018931.

144. Marian, A.J. Causality in Genetics. *Circ. Res.* **2014**, *114*, e18–e21, doi:10.1161/CIRCRESAHA.114.302904.

145. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, W.W.; Hegde, M.; Lyon, E.; Spector, E.; *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **2015**, *17*, 405–423, doi:10.1038/gim.2015.30.

146. Rodenburg, R.J. The functional genomics laboratory: functional validation of genetic variants. *J. Inherit. Metab. Dis.* **2018**, *41*, 297–307, doi:10.1007/s10545-018-0146-7.

147. Lee, J.-S.; Lee, J.Y.; Song, D.W.; Bae, H.S.; Doo, H.M.; Yu, H.S.; Lee, K.J.; Kim, H.K.; Hwang, H.; Kwak, G.; *et al*. Targeted *PMP22* TATA-box editing by CRISPR/Cas9 reduces demyelinating neuropathy of Charcot-Marie-Tooth disease type 1A in mice. *Nucleic Acids Res.* **2019**, *48*, 130–140, doi:10.1093/nar/gkz1070.

148. Gurda, B.L.; Bradbury, A.M.; Vite, C.H. Canine and Feline Models of Human Genetic Diseases and Their Contributions to Advancing Clinical Therapies . *Yale J. Biol. Med.* **2017**, *90*, 417–431.

149. Katz, M.L.; Coates, J.R.; Sibigtroth, C.M.; Taylor, J.D.; Carpentier, M.; Young, W.M.; Wininger, F.A.; Kennedy, D.; Vuillemenot, B.R.; O'Neill, C.A. Enzyme replacement therapy attenuates disease progression in a canine model of late-infantile neuronal ceroid lipofuscinosis (CLN2 disease). *J. Neurosci. Res.* **2014**, *92*, 1591–1598, doi:10.1002/jnr.23423.

150. Katz, M.L.; Tecedor, L.; Chen, Y.; Williamson, B.G.; Lysenko, E.; Wininger, F.A.; Young, W.M.; Johnson, G.C.; Whiting, R.E.H.; Coates, J.R.; *et al*. AAV gene transfer delays disease onset in a *TPP1*-deficient canine model of the late infantile form of Batten disease. *Sci. Transl. Med.* **2015**, *7*, 313ra180-313ra180, doi:10.1126/scitranslmed.aac6191.

151. Bradbury, A.M.; Rafi, M.A.; Bagel, J.H.; Brisson, B.K.; Marshall, M.S.; Pesayco Salvador, J.; Jiang, X.; Swain, G.P.; Prociuk, M.L.; ODonnell, P.A.; *et al*. AAVrh10

Gene Therapy Ameliorates Central and Peripheral Nervous System Disease in Canine Globoid Cell Leukodystrophy (Krabbe Disease). *Hum. Gene Ther.* **2018**, *29*, 785–801, doi:10.1089/hum.2017.151.

152. Bauer, A.; Waluk, D.P.; Galichet, A.; Timm, K.; Jagannathan, V.; Sayar, B.S.; Wiener, D.J.; Dietschi, E.; Muller, E.J.; Roosje, P.; *et al*. A de novo variant in the *ASPRV1* gene in a dog with ichthyosis. *PLoS Genet.* **2017**, *13*, e1006651, doi:10.1371/journal.pgen.1006651.

153. Christiansen, H.E.; Schwarze, U.; Pyott, S.M.; AlSwaid, A.; Al Balwi, M.; Alrasheed, S.; Pepin, M.G.; Weis, M.A.; Eyre, D.R.; Byers, P.H. Homozygosity for a Missense Mutation in *SERPINH1*, which Encodes the Collagen Chaperone Protein HSP47, Results in Severe Recessive Osteogenesis Imperfecta. *Am. J. Hum. Genet.* **2010**, *86*, 389–398, doi:10.1016/j.ajhg.2010.01.034.

154. Boyden, L.M.; Zhou, J.; Hu, R.; Zaki, T.; Loring, E.; Scott, J.; Traupe, H.; Paller, A.S.; Lifton, R.P.; Choate, K.A. Mutations in *ASPRV1* Cause Dominantly Inherited Ichthyosis. *Am. J. Hum. Genet.* **2020**, *107*, 158–163, doi:10.1016/j.ajhg.2020.05.013.

155. Mauler, D.A.; Gandolfi, B.; Reinero, C.R.; Spooner, J.L.; Lyons, L.A.; 99 Lives Consortium Precision Medicine in Cats: Novel Niemann-Pick Type C1 Diagnosed by Whole-Genome Sequencing. *J. Vet. Intern. Med.* **2017**, 539–544, doi:10.1111/jvim.14599.

156. Brunetti, B.; Muscatello, L. V.; Letko, A.; Papa, V.; Cenacchi, G.; Grillini, M.; Murgiano, L.; Jagannathan, V.; Drögemüller, C. X-Linked Duchenne-Type Muscular Dystrophy in Jack Russell Terrier Associated with a Partial Deletion of the Canine *DMD* Gene. *Genes (Basel).* **2020**, *11*, 1175, doi:10.3390/genes11101175.

157. Margolis, C.A.; Schneider, P.; Huttner, K.; Kirby, N.; Houser, T.P.; Wildman, L.; Grove, G.L.; Schneider, H.; Casal, M.L. Prenatal Treatment of X-Linked Hypohidrotic Ectodermal Dysplasia using Recombinant Ectodysplasin in a Canine Model. *J. Pharmacol. Exp. Ther.* **2019**, *370*, 806–813, doi:10.1124/jpet.118.256040.

Declaration of Originality

Last name, first name: Letko, Anna

Matriculation number: 16-129-553

I hereby declare that this thesis represents my original work and that I have used no other sources except as noted by citations.

All data, tables, figures and text citations which have been reproduced from any other source, including the internet, have been explicitly acknowledged as such.

I am aware that in case of non-compliance, the Senate is entitled to withdraw the doctorate degree awarded to me on the basis of the present thesis, in accordance with the "Statut der Universität Bern (Universitätsstatut; UniSt)", Art. 69, of 7 June 2011.

Place, date

Bern, 22.03.2021

Signature

Anna Letko Digitally signed by Anna Letko
Date: 2021.03.22 13:52:22 +01'00'