Bioinformatic analysis of a potential hereditary disease in New Zealand police dogs

A thesis presented

in partial fulfilment of the requirements

for the degree of

Master of Science in Animal Science

Massey University, Palmerston North,

New Zealand.

Olivia Breanna Doughty

2025

Abstract

Dogs have been used as working animals for centuries, one more recent example being the German Shepherd Dog (GSD) being trained and utilised as police dogs. Problems can arise from diseases associated to specific breeds, as well as the increased risks of inbreeding within already limited populations. This study aimed to investigate the genetic cause of an unknown illness seen in four dogs from a litter of seven GSD puppies bred for the New Zealand police dog section.

Four nine-month-old GSDs presented with clinical symptoms of severe exercise intolerance, and extremely elevated muscle enzymes (creatine kinase, alanine aminotransferase and aspartate transferase) post-exercise. A muscle biopsy of an affected dog gave insufficient evidence for conclusive diagnosis, leading to subsequent whole genome sequencing. Various bioinformatic filters were applied to the sequence data from all seven puppies, and the dam and sire, to determine the mode of inheritance for the affected phenotype. Sex-linked, autosomal dominant and mitochondrial inheritance were ruled out due to a lack of evidence, which suggests that autosomal recessive inheritance was likely.

The recessive filtering marked variants where the affected animals had a homozygous affected haplotype, and the parents were both heterozygous. Unaffected animals could be either heterozygous or homozygous unaffected. There were 94 variants across 30 different genes estimated to have either moderate or high impact. Further filters reduced this to four candidate variants, in the genes *BOP1*, *IQANK1*, *THEMIS2*, and *ZNF517*. None of the genes identified were directly associated with exercise tolerance in the affected phenotype, therefore it remains unclear whether one or a combination of these may be the cause of the phenomenon presented in affected animals.

Variants were almost exclusively found across a 21 million base pair region on chromosome 2 and a 10 million base pair region on chromosome 13. This is a possible indication of a larger structural issue within the affected animals; however, it was not within the scope of this study to investigate this further.

With inconclusive evidence, no single variant can be nominated as the likely causal variant, but considering the welfare implications of the affected condition, alongside their inability to continue their training as police dogs, all nine animals should be removed from further breeding to prevent the spread of the deleterious alleles.

Acknowledgements

I would firstly like to thank my supervisor Dr Nick Sneddon. I am incredibly grateful to have had your endless support, feedback and encouragement throughout my postgraduate study, your help has been instrumental in me being able to finish writing my thesis. I also want to thank my co-supervisors Prof Matt Littlejohn and Dr Thomas Lopdell. All three of you brought limitless ideas, suggestions and guidance and made this project possible.

Thank you to Dr Hayley Hunt for being my veterinary liaison and sharing your expertise to help me interpret the clinical results. And to the rest of the genetic R&D team at LIC for welcoming me in temporarily and helping me troubleshoot my never-ending battle coding my data analysis.

My thanks also go to Inspector John Bleackley and Dr Vicki Ercig at the New Zealand Police Dog Section for sharing the data and DNA of the dogs within this study, as well as information about the breeding and training procedures.

To my parents and family, for their unwavering support and encouragement, and for nodding along when I would ramble on about something completely foreign to them. And finally, to my horse, Evie, and her surprise son Moonshine, for both simultaneously testing and maintaining my patience and sanity.

Table of Contents

Abstract ..... **iii**
Acknowledgements ..... **v**
Table of Contents ..... **vii**
List of Tables ..... **ix**
List of Figures ..... **x**
List of Abbreviations ..... **xii**
Chapter 1 – Introduction **1**
Chapter 2 – Literature Review **3**
2.1 – German Shepherd Dog Breed History 3
2.2 – Police Dogs in New Zealand 4
2.2.1 – Police Dog History *4*
2.2.2 – Police Dog Training *5*
2.2.3 – Police Dog Breeding *8*
2.3 – Genetics 8
2.3.1 – Hereditary Conditions *8*
2.3.2 – Dominant versus Recessive Autosomal Conditions *10*
2.3.3 – Sex Linked Conditions *11*
2.3.4 – Polygenic/Multifactorial Conditions *13*
2.3.5 – Mitochondrial Conditions *14*
2.3.6 – Risks of Inbreeding *15*
2.4 – Known Hereditary Conditions in German Shepherd Dogs 15
2.4.1 – Degenerative Myelopathy *17*
2.4.2 – Haemophilia A *18*
2.4.3 – Hereditary Multifocal Renal Cystadenocarcinoma and Nodular Dermatofibrosis *19*
2.4.4 – Megaoesophagus *19*
2.4.5 – Musculoskeletal Disorders *20*
2.4.6 – Pancreatic Acinar Atrophy *21*
2.4.7 – Pituitary Dwarfism *22*
2.4.8 – von Willebrand Disease *22*
2.4.9 – Previously Undescribed Genetic Mutations in German Shepherd Dogs *23*
2.4.9.1 – BOP1 ..... *23*
2.4.9.2 – IQANK1 ..... *24*
2.4.9.3 – THEMIS2 ..... *24*
2.4.9.4 – ZNF517 ..... *24*

2.5 – Gene Characterisation and Discovery 25
Chapter 3 – Methods 27
3.1 – Animal Background 27
3.1.1 – Inbreeding Coefficients 29
3.1.2 – Blood Tests 30
3.1.3 – Muscle Biopsy 30
3.2 – Genetic Testing 31
3.2.1 – DNA Sequencing 31
3.2.2 – DNA Processing 31
Chapter 4 – Results 33
4.1 – Clinical Results 33
4.1.1 – Blood Tests 33
4.1.2 – Histology and Electron Microscopy 33
4.2 – Genetic Findings 36
4.2.1 – Pedigree Analysis 36
4.2.2 – DNA Quality 38
4.2.3 – Variant Calling 39
4.2.4 – Chromosomal Abnormalities 43
4.2.5 – Variants of Interest 45
4.2.6 – Population Filtering and RNA Coverage 46
Chapter 5 – Discussion 49
5.1 – Clinical Findings 49
5.1.1 – Mitochondrial Myopathies 49
5.2 – Chromosomal Issues 50
5.3 – Candidate Variants 50
5.3.1 – BOP1 50
5.3.2 – IQANK1 51
5.3.3 – THEMIS2 51
5.3.4 – ZNF517 51
5.4 – Implications of Inbreeding for Animals 51
5.5 – Limitations on Study 52
Chapter 6 – Conclusion 53
Chapter 7 – References 55
Chapter 9 – Appendix 73

List of Tables

Table 1 – Police patrol dog qualification marking guide, based on the 'Schutzhund' (Salt & Salt, 1972)..... 7

Table 2 – Heritable diseases commonly seen in German Shepherd Dogs and their mode of inheritance. Adapted from Wahl et al. (2008)..... 16

Table 3 – Comparison of heritability scores of some different complex musculoskeletal diseases from various studies on populations of German Shepherd Dogs, arthritis and spinal disease have been excluded from this table due to low to no papers presenting genetic parameters..... 21

Table 4 – Blood serum levels (IU/L) of creatine kinase (CK), aspartate transferase (AST) and alanine aminotransferase (ALT) of the affected and unaffected dogs before or after exercise. 33

Table 5 – Tabulated inbreeding coefficient of animals from the dam and sires' pedigree in Figure 9. Values on the diagonal represent the inbreeding coefficient of the individual(s) plus one, meaning that numbers greater than one indicate a level of inbreeding. Off-diagonal values describe the degree of relatedness between individual animals. 37

Table 6 – Average quality, coverage depth and error rate of the individual BAM files produced for each animal..... 39

Table 7 – Number of variants (n) found for each filtering method by the type of mutation and subsequent impact. 40

Table 8 – Number and proportion of variant effects produced through recessive variant filtering by variant type. Includes modifier variants. 42

Table 9 – Number and proportion of variant effects produced through recessive variant filtering by region variant occurred in. Includes modifier variants..... 42

Table 10 – Frequency and rate at which variants occur in each autosome and non-nuclear DNA. 44

Table 11 – Genes remaining with variants present within an exon after all filtering steps, including their location and impact on the gene and change to amino acids. 47

List of Figures

Figure 1 – Punnett squares displaying different simple Mendelian crosses with different colours of the offspring based on genotype to show different methods of expression. For a simple recessive condition, the red and orange genotypes will show the same phenotype, different to the recessive (yellow) phenotype. A codominant trait will have three distinct phenotypes coordinated to the three genotypes as shown in red, orange and yellow. **A** shows a homozygous dominant, heterozygous cross, **B** shows a heterozygous or 'test' cross and **C** shows a heterozygous, homozygous recessive cross. 35

Figure 2 – Family trees displaying inheritance patterns of dominant sex-linked conditions based on various combinations of parental genotypes. Circles denote female animals; squares denote males with a dash over affected chromosomes. Black shapes depict affected animals; white shapes are unaffected. 12

Figure 3 – Family trees displaying inheritance patterns of recessive sex-linked conditions based on various combinations of parental genotypes. Circles denote female animals; squares denote males with a dash over affected chromosomes. Black shapes depict affected animals, hatched shapes indicate carrier animals, white shapes are unaffected. 13

Figure 4 – Segregation of mutant and wild-type mitochondrial DNA at differing levels to illustrate the variance caused by heteroplasmy. Adapted from Tarnopolsky and Raha (2005). 14

Figure 5 – 'W' litter to parents of unknown genotype. A square indicates a male animal and a circle a female. Filled in shapes indicate affected dogs. Strikethrough animals are deceased. 27

Figure 6 – The sire's offspring to different bitches with the litters date of birth and numbers of each sex born. The letter within the diamond indicates the name applied to the litter in the naming scheme. 29

Figure 7 – Photomicrographs of skeletal muscle from affected 1, stained with heamatoxylin and eosin A 40x magnification, B 200x magnification, C 400x magnification and D 400x magnification of acute degeneration. 34

Figure 8 – Electron microscopy of skeletal muscle from a normal dog (A and B) and affected 1 (C – F)..... 35

Figure 9 – Dam and sire's family tree with common ancestors 2 and 3. A square indicates a male animal and a circle a female. Filled in shapes indicate affected genotype dogs. Strikethrough animals are deceased. 36

Figure 10 – Distribution of recessive variants across chromosome 2. 41

Figure 11 – Distribution of recessive variants across chromosome 13. 41

Figure 12 – Number of variants (log_{10}) produced through filtering for recessive traits where affected animals are homozygous for loci that parents are both heterozygous. 45

Figure 13 – Integrative genomics viewer (IGV) screenshot of RNA coverage aligned to the CanFam4 reference genome. Includes areas of coverage within the exon (red line, right) and areas without RNA coverage (intron, red line, left). Areas with no RNA coverage are not expressed within the tissue. 46

List of Abbreviations

ALT – Alanine aminotransferase

AST – Aspartate transferase

CK – Creatine kinase

DNA – Deoxyribonucleic acid

GSD – German Shepherd Dog

IGV – Integrative genomics viewer

INDEL – Insertion deletion

mtDNA – Mitochondrial DNA

nDNA – Nuclear DNA

RNA – Ribonucleic acid

SNP – Single nucleotide polymorphism

VCF – Variant calling file

VEP – Variant effect predictor

VWD – von Willebrand disease

VWF – von Willebrand factor

WGS – Whole genome sequencing

Chapter 1 – Introduction

Working dogs have been a key component of New Zealand's police force since the establishment of the Trentham breeding and training centre in the 1950's. Breeding populations of primarily German Shepherd Dogs (GSD) and Labradors have been maintained to produce both patrol and detector dogs for the country (*Police Dog Section*, n.d.; *Police Dog Trust | New Zealand*, n.d.; Redwood, 1980; Salt & Salt, 1972). Since its founding of 16 dogs initially imported from Britain, the police dog squad has expanded to over 120 active patrol dogs and handlers spread across the country (Pers comm, Bleackley, 2025; Salt & Salt, 1972).

Patrol dogs undergo extensive specialised training with initial courses to qualify a dog and their handler taking roughly one and a half to two years and costing $58,000, including the cost of breeding, training, equipment and staff salaries (Pers comm, Bleackley, 2025). Beyond this, dogs and handlers must pass annual assessments to continue active duty, alongside any further specific training in areas such as the Armed Offenders Squad, Special Tactics Group or search and rescue for dogs suited to such jobs (New Zealand Police, 2022; Salt & Salt, 1972).

Given the highly physically demanding work performed by police dogs, as well as the time and monetary investment, it is vital to produce physically and mentally capable animals that can maintain a long working life. Assessing the genetic makeup of the dogs is an important step in breeding healthy offspring, particularly in a breed that is prone to heritable diseases such as GSD (Padgett, 1998; Wahl et al., 2008). Breeding animals are chosen not only for competency in work and having the correct temperament, but also for their health scores, including PennHIP, New Zealand Veterinary Association elbow scores and lumbosacral radiographs (Pers comm, Erceg, 2025). These assessments reduce the chance of producing dogs with degenerative musculoskeletal diseases such as hip and elbow dysplasia. Animals and genetic material are frequently imported from Europe to maintain genetic diversity amongst breeding stock and to reduce the risk of inbreeding (Pers comm, Erceg, 2025).

When heritable health issues arise, it is important that affected animals are identified and removed from the breeding pool to prevent further spread of detrimental traits due to the potential time, financial and welfare strains. This can be difficult in instances of common recessive disease, such as progressive retinal atrophy and type II and III von Willebrand disease. In such cases it is important to avoid carrier-carrier matings, which can be determined through genotyping animals prior to breeding. Thanks to the evolution of genome sequencing technology and the growing accessibility of single nucleotide polymorphism (SNP) chip

genotyping and polymerase chain reaction testing, it is relatively quick and easy to assess if an animal carries a gene known to cause heritable issues and can therefore be removed or selected against in the breeding population (Baker et al., 2019; Ng & Kirkness, 2010).

Chapter 2 – Literature Review

2.1 – German Shepherd Dog Breed History

Since the domestication of dogs began over 14,000 years ago (Vila et al., 1999), humans have employed them for many types of jobs, including managing and shepherding livestock. Humans began to selectively breed dogs for their aptitude in these skills to the point at which breeds for specific tasks (herding, guarding, hunting, tracking) began to develop. Within the herding dog breeds, there are those bred for their aptitude in moving sheep across various rugged terrain, including the German Shepherd Dog (GSD), although its modern usage has changed from these agricultural origins. In the mid 1800's, there were two somewhat distinct types of GSD's based on their region of origin, Thuringia and Württemberg. The Thuringia type were more wolf-like, with short grey fur, pricked ears, a curled tail, a smaller, wirier frame and more aggressive in nature (Goldbecker & Hart, 1984; Strickland, 1998; Tenner, 2017; Willis, 1992; Wootton, 1988). The Württemberg, however, was larger, with strong feet, floppy ears, long hair, a straight tail, an even temper and better movement (Goldbecker & Hart, 1984; Strickland, 1998; Tenner, 2017; Willis, 1992; Wootton, 1988).

In the late 1800's Max von Stephanitz began to develop and consolidate the two shepherd type dogs into a distinctive utility breed, the GSD (Tenner, 2017; Willis, 1992; Wootton, 1988). In 1899 he founded the *Verein für Deutsche Schäferhunde* (SV), or German Shepherd Association, alongside the first stud book and breed specific shows (Goldbecker & Hart, 1984; Strickland, 1998; Tenner, 2017; Willis, 1992; Wootton, 1988). Furthermore, von Stephanitz produced a 776-page breed standard, the *'Wortbild'* (Von Stephanitz & Schwabacher, 1923), for dogs to be judged on by himself at these shows (Tenner, 2017). This allowed him to choose champion dogs with the knowledge that they would become incredibly on the next generation of GSDs, reducing divergence and shaping the breed very rapidly (Tenner, 2017; Willis, 1992). This process was not without its drawbacks, as an examination of early GSD pedigrees showed a high level of inbreeding within the champion dogs. These dogs often sired hundreds of pups (Wootton, 1988), thereby compounding issues of hereditary diseases.

The breed was intended to be a working dog, with many being utilised as sheepdogs before their adoption into the German police force for the deterrence and apprehension of criminals (Strickland, 1998; Wootton, 1988). During World War I, they were widely used by the military for jobs such as messengers, finding wounded soldiers and as sentry guard dogs (Willis, 1992). Their utility noted by American soldiers, resulting in dogs being taken to America to produce

puppies for civilian adoption (Goldbecker & Hart, 1984; Willis, 1992). This led to a rapid growth in the popularity of the breed, largely due to Strongheart and Rin Tin Tin, GSDs featured widely in American film in the 1920's (Goldbecker & Hart, 1984; Strickland, 1998; Tenner, 2017). Increased demand for GSD puppies caused a spike of indiscriminate breeding, with many dogs displaying poor disposition and conformation. Consequently, there was a downshift in popularity due to the resulting bad publicity (Goldbecker & Hart, 1984; Redwood, 1980; Strickland, 1998) and fear that the breed contained 'wolf blood' (Redwood, 1980). This fear may have been furthered by the name 'Alsatian wolfdog' as they widely referred to in Britain following World War II as an aversion to owning a German breed of dog (Wootton, 1988).

Dedicated breeders continued to produce dogs fitting breed standards which also included an ability to work, to maintain the utility within the standard of the breed as von Stephanitz had desired (Goldbecker & Hart, 1984; Redwood, 1980). The breed standards resulted in the *'Schutzhund'* (protection dog), qualification based on skills of obedience, agility, tracking and protection (Goldbecker & Hart, 1984; Wootton, 1988). Following World War II, the popularity of the breed increased once more, as knowledge of their aptitude as a working dog spread, promoting their use by military and police forces (Redwood, 1980; Wootton, 1988). One such group was the Surrey Police Dog School in England, who later assisted in setting up New Zealand's police dog department through the contribution of expertise and breeding animals (Redwood, 1980; Salt & Salt, 1972).

2.2 – Police Dogs in New Zealand

2.2.1 – Police Dog History

The Surrey Constabulary Police Dog School was established in 1949 by Sergeant Frank Riley, who saw the utility of the GSD or Alsatian, as they were referred to in Britian following World War II (Goldbecker & Hart, 1984; Redwood, 1980). His own dog Miska was crowned the champion police dog in the United Kingdom in 1955 (Redwood, 1980; Salt & Salt, 1972). The same year, New Zealand's Prime Minister, Sir Sydney Holland, visited the British police dog school and recruited Sergeant Riley to establish a police dog department in New Zealand (Redwood, 1980; Salt & Salt, 1972). Sergeant Riley arrived in New Zealand with his fully trained police dog (Miska), a nine-month-old dog, (Dante), two breeding bitches (Karen and Silver), and their 12 two-month-old puppies. The Police Dog Training Centre was launched in

Trentham, Upper Hutt in 1956 (*Police Dog Section*, n.d.; *Police Dog Trust | New Zealand*, n.d.; Salt & Salt, 1972).

Further dogs were imported from Britain due to import bans directly from Germany following World War II (Redwood, 1980). There were also heavy restrictions on imports from Australia, due to the fear that GSDs may have been crossbred with Dingoes (Salt & Salt, 1972). By the 1960's, the police dog unit had 17 trained dogs alongside breeding stock. At this point New Zealand began to supply the Pacific islands and French territories including New Guinea, Fiji and Samoa with animals for police work (*Police Dog Section*, n.d.; *Police Dog Trust | New Zealand*, n.d.; Salt & Salt, 1972), which continues to this day (Pers comm, Bleackley, 2025).

The New Zealand Police Dog Unit has further expanded with a new, purpose-built training complex in 1996 (*Police Dog Section*, n.d.; *Police Dog Trust | New Zealand*, n.d.). Alongside the upgraded facilities, training was further developed with a greater range of qualifications for patrol dogs including drug detection (1976), explosives detection (1977), the Armed Offenders Squad (1992), accelerant detection (1997), and search and research (1998) (*Police Dog Section*, n.d.; *Police Dog Trust | New Zealand*, n.d.).

2.2.2 – Police Dog Training

Training a police dog is a continual process, beginning at eight weeks of age when puppies are sent to foster homes to be socialised and exposed to various environments and circumstances. At approximately 10 months of age, they can be placed with a handler and begin training (Pers comm, Bleackley 2025; Kyono, 2002). Dogs may be deemed unsuitable for work at any point in training by the facility trainer, however, most ineligible dogs are rehomed before they are placed with a handler, due to the roughly $58,000 that accrues across breeding, initial development/socialisation and training (Pers comm, Bleackley, 2025; Kyono, 2002). Unsuitable animals are typically removed due to fearfulness, lacking the drive or aptitude for patrol dog tasks, or for health issues that may impair their ability to work (Kyono, 2002).

Coursework for new patrol dogs is split into four qualifying courses: puppy, initial, intermediate and final courses (New Zealand Police, 2022). Each of these steps builds on skills such as obedience, agility, tracking and 'man work' (Kyono, 2002). Following the final course, all patrol dogs must pass an assessment to be qualified as an operational dog, and this test is repeated annually to maintain aptitude in all dogs (New Zealand Police, 2022). This assessment was originally based on the Schutzhund assessments outlined in the *Wortbild* written by von Stephanitz, in which pairs of dogs and handlers had to reach a minimum qualifying mark across

16 tasks, as outlined in Table 1 from Salt and Salt (1972). More recently, the assessment has been adjusted to account for the competency of both handler and dog for various applicable skills in more realistic working circumstances than the arbitrary outline of the *Schutzhund* (Pers comm, Bleackley, 2025).

Alongside the basic training courses that all patrol dogs undertake, there are several additional, more specialised courses including the Armed Offenders Squad and Special Tactics Group for more tactically capable dogs, or search and rescue. Detection-based courses (e.g. narcotics, firearms, cash, cadaver) are also offered at the training centre for dogs bred for detection, typically Labrador's and Springer Spaniels (Pers comm, Bleackley, 2025; New Zealand Police, 2022). Across all these initial, advanced, and refresher courses available at the Trentham training centre, there are typically 25 to 45 dogs in formal training at any one time (Pers comm, Bleackley, 2025).

Table 1 – Police patrol dog qualification marking guide, based on the 'Schutzhund' (Salt & Salt, 1972).

	Marks	Group Total	Minimum Group Qualifying Marks
Group I – Control			
1. Heel free	5		
2. Directional control, sending dog away in direction indicated by Judge, not less than 50 yards, stopping and redirecting not less than 20 yards, and then recall on Judge's order	15	35	25
3. Sit two minutes. Handlers out of sight	5		
4. Down ten minutes. Handlers out of sight	10		
Group II – Agility			
5. Retrieving dumbbell over 6'6" standard obstacle	10	20	14
6. Clear jump over 3'6"	5		
7. Clear long jump 10 feet	5		
Group III – Track			
8. Leash track. There shall be no re-casts on the direction of the Judge. Track shall not be less than 7/8th mile long and at least 3 hours old. Three different kinds of article will be distributed evenly along track. Similar articles to be used by each track layer and easily recognised by him		115	82
For track	100		
For recognition of 3 articles 5 points each	15		
Group IV – Search			
9. Controlled search of area of foiled ground to find and retrieve four strange articles handled and placed by some person other than the handler. The area should be approx. 25 yards square, and time allowed 5 minutes	35	35	25
Group V – Patrol			
10. Quartering the ground for hidden person, baying, not biting when found and refusing food from hidden person	35		
11. Test of courage against attack by criminal with a stick	10		
12. Test of courage against attack by gunfire	10		
13. Defending handler when attacked, and immediate release when attack ceases or when commanded to leave by handler to follow a search and escort	20	140	98
14. Guarding criminal in handler's absence	20		
15. Recall from criminal running away	30		
16. Pursuit and detention of criminal running away until arrival of handler	15		
	Totals	345	244

2.2.3 – Police Dog Breeding

The New Zealand Police Dog Breeding Centre produces all their own working dogs, both GSDs for patrol dogs and Labradors for detector dogs from a wide pool of breeding stock (Pers comm, Bleackley, 2025). Approximately 100 puppies are produced annually from a selection of ~40 breeding bitches and over 100 stud dogs (including imported and frozen semen) (Pers comm, Erceg, 2025). Approximately 20% of these offspring are Labrador puppies, and almost all of the remaining puppies are GSD's (Pers comm, Bleackley, 2025).

Dogs that successfully make it through training go on to become or replace one of the approximately 120 pairs of patrol dogs and handlers spread out across New Zealand. In addition, a portion of the dogs trained are sent overseas to supply the Pacific islands (Pers comm, Bleackley, 2025). The health and longevity of these dogs is vital to consider as they are highly valuable animals (Pers comm, Bleackley, 2025). Ideally, police dogs will work until they are eight years of age, however, they may be retired early due to behavioural or health issues. Worth et al. (2013) reported that the average age of retirement was six and a half years.

2.3 – Genetics

2.3.1 – Hereditary Conditions

The genome of *Canis lupus familiaris* (domestic dog) contains 38 pairs of autosomes alongside the sex-determining X and Y chromosomes. The genome contains approximately 2.5 gigabases (*Genome Assembly (UU_Cfam_GSD_1.0)*, 2010) that include 21.2 thousand coding genes and 16.5 thousand non-coding genes, ("NCBI *Canis lupus familiaris* Annotation," 2021). A locus (location on the chromosome or within the genome) contains one copy (or allele) for a specific gene on each of the paired chromosomes which are inherited from the two parents. This combination of alleles provides the animal's genotype, leading ultimately to the physical traits of the animal, i.e., the phenotype (Mahdieh & Rabbani, 2013).

Selective breeding is the process by which animals that perform well for a particular skill or favoured trait are chosen to produce offspring that may perform better (Oberbauer & Sampson, 1999). Where New Zealand police dogs are concerned, traits related to working such as a high drive, strong obedience, and agility are favoured in breeding dogs, whereas physical traits such as coat colour are not considered. Additionally, deleterious traits, such as a predisposition to

certain diseases, which impede the dog's ability to have a long working career and impair welfare are selected against and removed from the breeding population where possible.

The phenotype conferred by any gene is dependent on the specific genotype, which is represented by a combination, or the additive effect of, dominant and recessive alleles based on what is inherited from the parents (Mahdieh & Rabbani, 2013). A dominant trait requires only one allele to be present in the genotype to be expressed in the phenotype. Dominant traits are typically denoted with a capital letter (e.g. 'AA' or 'Aa') in traditional punnet square examples (Figure 1). A recessive phenotype requires both alleles to be expressed and is denoted with a lower-case letter (e.g. 'aa') (Padgett, 1998). Where an animal has two copies of the same allele, i.e., both dominant or both recessive, the genotype is called homozygous (e.g. 'AA' or 'aa'), whereas a genotype comprising one copy each of the dominant and recessive alleles each is called heterozygous (e.g. 'Aa') (Padgett, 1998).

Canine nuclear DNA (nDNA) is roughly 3000 megabases long and encodes approximately 35,000 – 40,000 known genes across 38 autosomes plus X and Y chromosomes (Sargan et al., 2001). During meiosis there are many opportunities for genetic material to be damaged, causing mutations which create genetic variants that can alter the outcome of a trait, such as the expression of a disease (M. Brooks & Sargan, 2001; Padgett, 1998). There are over 200 diseases with an identified genetic cause in canines, the majority of which are caused by simple autosomal inheritance of a single variant (Brooks & Sargan, 2001). However, genetic anomalies can also occur on a chromosomal or polygenic/multifactorial level (Mahdieh & Rabbani, 2013). Chromosomal abnormalities such as aneuploidies, where there is an extra chromosome (trisomy) or are missing a chromosome (monosomy) are most often described in the sex chromosomes (X/Y). These abnormalities depend on the combination of sex chromosomes present and have different physical and endocrinological consequences (Breen et al., 2001).

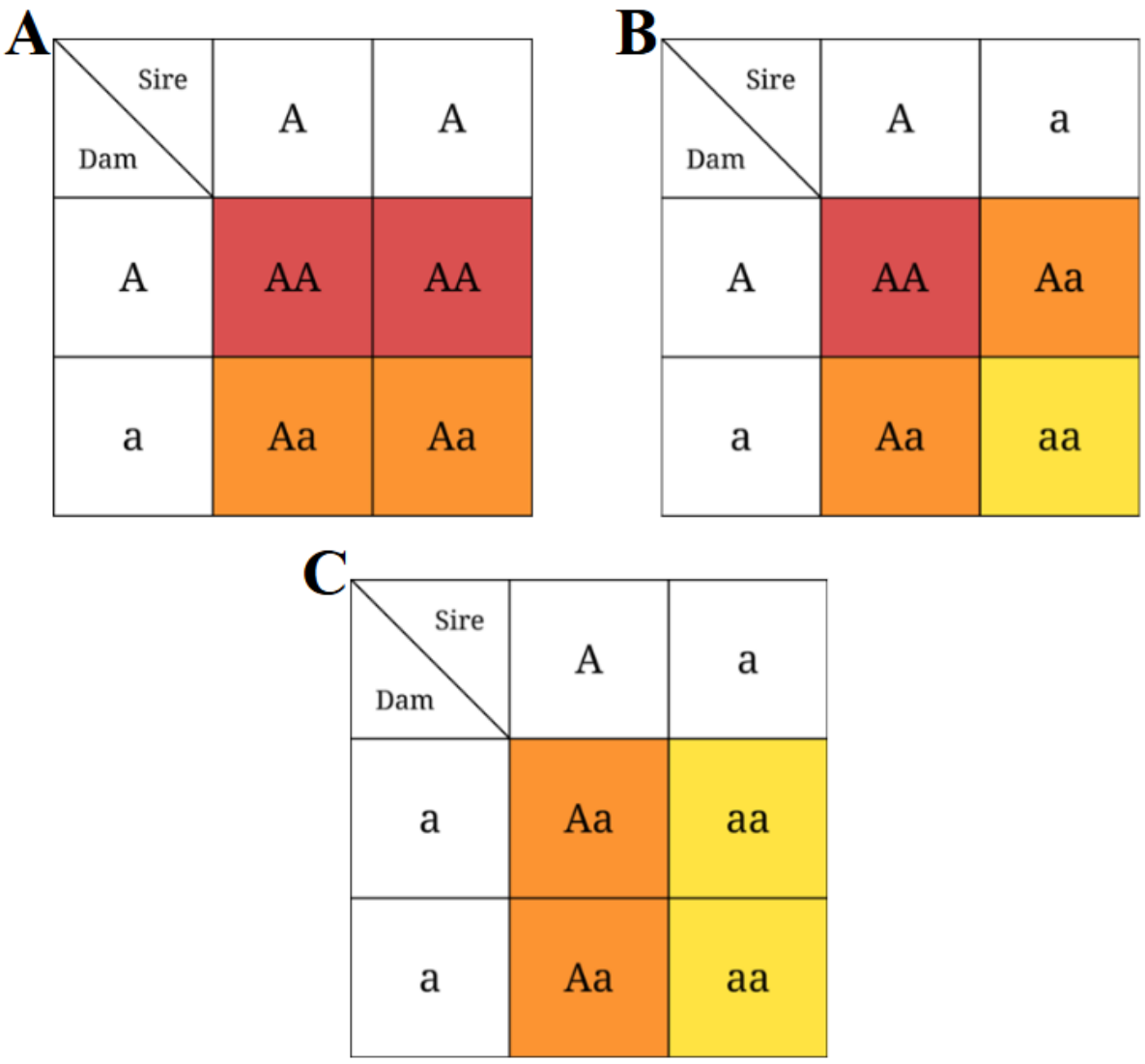

Figure 1 – Punnett squares displaying different simple Mendelian crosses with different colours of the offspring based on genotype to show different methods of expression. For a simple recessive condition, the red and orange genotypes will show the same phenotype, different to the recessive (yellow) phenotype. A codominant trait will have three distinct phenotypes coordinated to the three genotypes as shown in red, orange and yellow. **A** shows a homozygous dominant, heterozygous cross, **B** shows a heterozygous or 'test' cross and **C** shows a heterozygous, homozygous recessive cross.

2.3.2 – Dominant versus Recessive Autosomal Conditions

For a simple, or single gene, autosomal recessive trait such as pituitary dwarfism, (see section 2.4.7), an animal must have a homozygous recessive genotype to display the recessive phenotype. In contrast, an animal with a heterozygous genotype is considered to be a carrier for the trait and a homozygous dominant animal does not carry the recessive gene (Padgett, 1998). However, for an autosomal dominant trait, such as renal cystadenocarcinoma and

nodular dermatofibrosis (see section 2.4.3), to be displayed in the phenotype, only one copy of the dominant allele is required. Therefore, animals with either the homozygous dominant or the heterozygous genotype will have the affected phenotype, and only animals with a homozygous recessive genotype will be clear of the disease (Padgett, 1998). When breeding for a particular trait, the genotype of each parent and whether the gene is dominant or recessive will impact the possible proportions of genotypes and phenotypes seen in the offspring (Figure 1).

There are several ways that these simple inheritance patterns can be disrupted, such as multiple interacting traits (epistasis), incomplete dominance, co-dominance and lethal traits (Mahdieh & Rabbani, 2013). When a trait is affected by multiple variants at the same time, the number of possible genotypes can exponentially increase. The range of possibilities for the resulting phenotype will also increase, especially when considering epistatic interactions between the variants. One such phenotype is coat colour, which in dogs is controlled by at least 21 loci which contribute 52 different alleles (Sponenberg & Rothschild, 2001).

Dominance can also be expressed in different ways such as the co-dominance seen in blood types, or coat colour, where all three possible genotypes (AA, Aa and aa), will all cause different phenotypes (Figure 1, colour coding). Additionally, there can also be incomplete dominance/penetrance, where the related phenotype can be expressed at varying levels both between and with different genotypes, such as Type I von Willebrand's disease (see section 2.4.8). Lethal traits are often a rarer form of heritable disease as a carrier (heterozygous) is typically asymptomatic and an affected offspring (homozygous) will be miscarried, stillborn or die soon after birth due to the condition, for example, lethal lung disease in the Airedale Terrier (Dillard et al., 2020).

2.3.3 – Sex Linked Conditions

Sex linked conditions, although rare, are found almost exclusively on the X chromosome can be either recessive or dominant (Padgett, 1998). Due to males only possessing one X chromosome, they are more likely to be affected by a sex-linked condition (Mahdieh & Rabbani, 2013; Padgett, 1998). This can cause inheritance patterns to differ between sexes as daughters can only receive an affected chromosome from their sires (Figure 2A and 2C) (Lynch & Walsh, 1998; Mahdieh & Rabbani, 2013; Padgett, 1998). Comparatively, an affected dam will produce 50% affected offspring, regardless of sex (Figure 2B). The X chromosome is

related to several blood disorders including haemophilia A and von Willebrand disease (see sections 2.4.2 and 2.4.8 respectively).

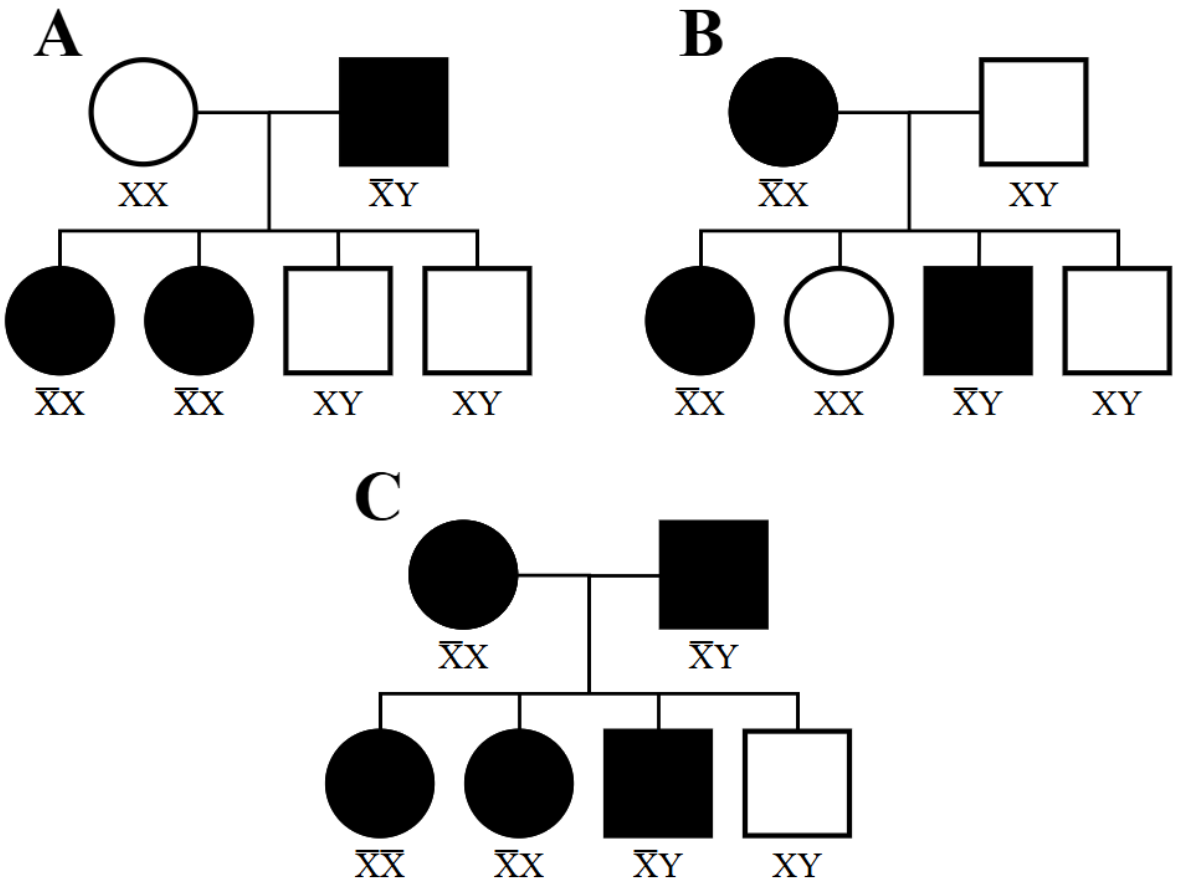

Figure 2 – Family trees displaying inheritance patterns of dominant sex-linked conditions based on various combinations of parental genotypes. Circles denote female animals; squares denote males with a dash over affected chromosomes. Black shapes depict affected animals; white shapes are unaffected.

For a recessive sex-linked condition, a male will be affected whenever he carries a single copy and will pass that affected allele onto all of his daughters (Figure 3A). Sons, however, are not impacted by their sire's X genotype, as they instead receive the sires Y chromosome. A carrier female when bred to an unaffected sire will, on average, have half of her daughters and half her sons be affected (Figure 3B). An affected female, however, is homozygous, so will produce only affected sons and carrier daughters to an unaffected sire (Figure 3C). The daughters of a carrier dam mated to an affected sire will be evenly split between affected and carriers, and sons will be split evenly between affected and unaffected (Figure 3D). Two affected parents will only produce affected offspring (Figure 3E). The same patterns of inheritance are also

followed for a dominant X linked condition, but carriers of the trait will become affected (Padgett, 1998).

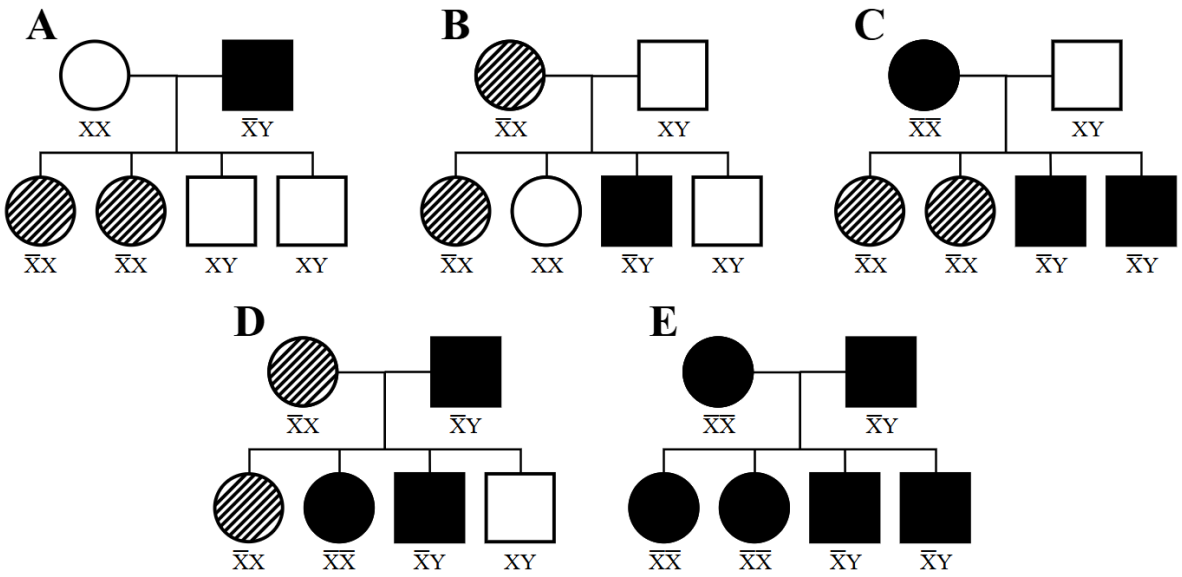

Figure 3 – Family trees displaying inheritance patterns of recessive sex-linked conditions based on various combinations of parental genotypes. Circles denote female animals; squares denote males with a dash over affected chromosomes. Black shapes depict affected animals, hatched shapes indicate carrier animals, white shapes are unaffected.

2.3.4 – Polygenic/Multifactorial Conditions

Compared to Mendelian traits, polygenic, or multifactorial traits, display incredibly complex modes of inheritance that are poorly understood and difficult to quantify, particularly disease related traits (Leeb et al., 2022; Padgett, 1998). Polygenic traits are the cumulative result of many different alleles at various loci and may be spread across different chromosomes with extrinsic factors such as environment, nutrition and exercise impacting the expression of said traits (Leeb et al., 2022; Momen & Muir, 2025; Padgett, 1998).

Assessment for complex traits, especially deleterious traits such as hip dysplasia, is mostly based on phenotypic evidence (Leeb et al., 2022; Momen & Muir, 2025; Padgett, 1998); however, new methods similar to those used to calculate breeding values are being developed to produce a ‘polygenic risk score’ (Momen & Muir, 2025). While extrinsic factors can influence the expression of the trait, purebred animals are more frequently affected by polygenic diseases. This is due to the founder effect during breed creation and high levels of inbreeding compared to mixed breed dogs, making it easier for the deleterious alleles to accumulate (Leeb et al., 2022; Momen & Muir, 2025).

2.3.5 – Mitochondrial Conditions

Mitochondria are an almost exclusively maternally inherited organelle encoded for by both nDNA as well as its own mitochondrial DNA (mtDNA). Due to the mitochondria's vital role in metabolic processes such as energy production and storage, many mitochondrial diseases are energy dependent conditions (Gomes, 2021; Paciello et al., 2003).

Mitochondrial DNA is a single circular chromosome that is 0.017 megabases long that is only responsible for 37 genes (Tkaczyk-Wlizło et al., 2022). The canine mtDNA contains no introns or protective histones, leaving it open to constant oxidative stress, leading to a mutation rate ten times greater than that of nDNA (Paciello et al., 2003).

Given the high rate of polymorphisms, mitochondrial genomes within a given cell can differ, leading to the phenomenon known as 'heteroplasmy' (Tarnopolsky & Raha, 2005; Tkaczyk-Wlizło et al., 2022). Heteroplasmy is where the ratio of mutant to wild-type mitochondria can range in a ratio from 99:1 to 1:99 (Figure 4; Tarnopolsky & Raha, 2005). Due to this effect, there can be many phenotypic variances between different tissues or organs as well as animals (Tarnopolsky & Raha, 2005). Identification and diagnosis of mitochondrial conditions very difficult and therefore poorly understood (Tkaczyk-Wlizło et al., 2022).

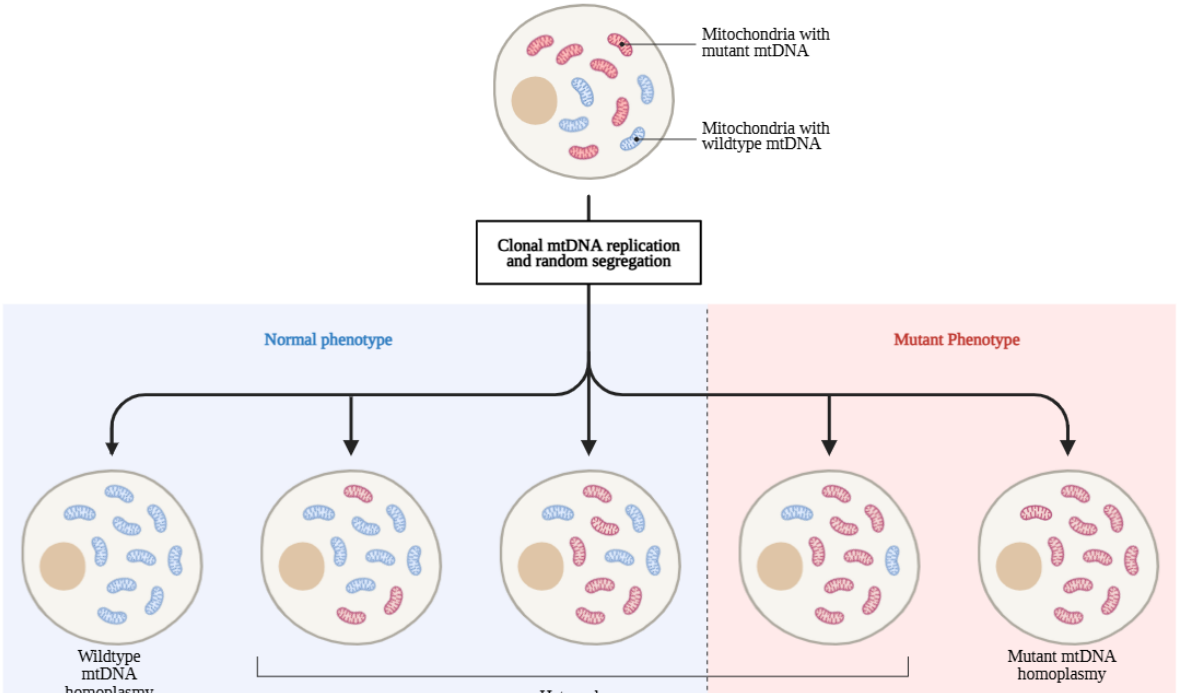

Figure 4 – Segregation of mutant and wild-type mitochondrial DNA at differing levels to illustrate the variance caused by heteroplasmy. Adapted from Tarnopolsky and Raha (2005).

2.3.6 – Risks of Inbreeding

Inbreeding is a phenomenon that occurs when related animals' mate. In populations of a finite size, for example endangered animals where populations are smaller or fragmented, inbreeding may be impossible to avoid but may also occur with the preferential selection of mates (Edmands, 2006; Kristensen & Sørensen, 2005). Domestic animals such as dogs and cats and livestock can be impacted by inbreeding due to heavy selection pressures, where related animals are bred together for preferential traits (Kristensen & Sørensen, 2005; Weigel, 2001). Problems occur with inbreeding due to the redistribution of genotypes and the reduction in heterozygosity of traits. This leads to the 'fixation' of genes in a homozygous form, as alleles are lost directly or indirectly through selection (Edmands, 2006; Kristensen & Sørensen, 2005; Weigel, 2001). The deleterious effects of inbreeding, known as inbreeding depression, have been described for centuries where the mean value of a trait is reduced (Kristensen & Sørensen, 2005). Inbreeding depression is typically described by two hypotheses, partial dominance and overdominance, although these two phenomena are not mutually exclusive (Kristensen & Sørensen, 2005). Partial dominance explains inbreeding depression as caused by the accumulation of deleterious recessive alleles in homozygous form, leading to their increased expression. Overdominance describes animals where the heterozygote state is superior compared to homozygote animals, due to the hypothesis that heterozygotes have a greater capacity to survive in variable environments (Edmands, 2006; Kristensen & Sørensen, 2005; Weigel, 2001). Under both hypotheses, there are serious concerns regarding declining animal health, growth, production and fertility with increased inbreeding. These lead to significant economic impacts on livestock production (Weigel, 2001) or a decreased capacity for work alongside the same, or increased, financial requirements. This is particularly important for working dogs which have a significant cost to produce.

2.4 – Known Hereditary Conditions in German Shepherd Dogs

Considering the time and monetary investment put into the GSDs used by the police, it is important to produce genetically sound animals. There is a high prevalence of hereditary diseases within the breed that could impair the ability of these animals to work. There are many hereditary conditions seen in the GSD with a range of modes of inheritance (Table 2). The following sections in this chapter detail some of the diseases seen frequently in GSD's and their impact on the animal.

Table 2 – Heritable diseases commonly seen in German Shepherd Dogs and their mode of inheritance. Adapted from Wahl et al. (2008).

Inherited disease found in GSD	Mode of Inheritance
Acral lick dermatitis	Undetermined
Aortic stenosis	Autosomal dominant
Atopic dermatitis	Autosomal dominant
Autoimmune lymphocytic thyroiditis	Autosomal recessive
Bladder stones	Autosomal recessive
Calcinosis circumscripta	Undetermined
Cataracts	Undetermined
Cerebellar abiotrophy	Autosomal recessive
Cervical vertebral instability	Undetermined
Cleft lip/palate	Undetermined
Corneal dystrophy	Undetermined
Cutaneous asthenia	Undetermined
Cutaneous lupus erythematosus	Undetermined
Deafness	Undetermined
Degenerative myelopathy	Undetermined
Demodecosis	Undetermined
Dermatomyositis	Undetermined
Dermoids	Undetermined
Elbow dysplasia	Polygenic
Epilepsy	Undetermined
Exocrine pancreatic insufficiency	Autosomal recessive
Familial vasculopathy	Autosomal recessive
Footpad disorder	Undetermined
Gastric dilatation-volvulus	Undetermined
German shepherd pyoderma	Autosomal recessive
Glycogen storage disease type III	Autosomal recessive
Hemivertebra	Autosomal recessive
Haemophilia	X-linked recessive
Hip dysplasia	Polygenic
Hyperadrenocorticism	Undetermined
Hypertrophic osteodystrophy	Undetermined
Lymphedema	Autosomal dominant
Masticatory myositis	Undetermined
Megaoesophagus	Undetermined
Mitral valve disease	Undetermined
Myasthenia gravis	Autosomal recessive
Nodular dermatofibrosis	Autosomal dominant
Optic nerve hypoplasia	Undetermined
Pancreatic acinar atrophy	Polygenic

Table 2 – Cont. – Heritable diseases commonly seen in German Shepherd Dogs and their mode of inheritance. Adapted from Wahl et al. (2008).

Pannus	Undetermined
Panosteitis	Undetermined
Patent ductus arteriosus	Polygenic
Pemphigus erythematosus	Undetermined
Perianal fistula	Undetermined
Persistent right aortic arch	Polygenic
Pituitary dwarfism	Autosomal recessive
Progressive retinal atrophy	Autosomal recessive
Pulmonic stenosis	Undetermined
Retinal dysplasia	Autosomal recessive
Sebaceous adenitis	Undetermined
Seborrhoea	Autosomal dominant (variable expression)
Selective IgA deficiency	Undetermined
Small intestinal bacterial overgrowth	Undetermined
Tricuspid dysplasia	Undetermined
Vertebral stenosis	Polygenic
Vitiligo	Undetermined
von Willebrand disease Type I	Autosomal (incomplete dominance)
von Willebrand disease Type II and III	Autosomal recessive

2.4.1 – Degenerative Myelopathy

Degenerative myelopathy is an adult-onset neurodegenerative disease characterised by the progressive degradation of the white matter within the spinal cord (Capucchio et al., 2014; Holder et al., 2014). This leads to loss of motor function and progressive ataxia paresis in the pelvic limbs, eventually worsening to affect the thoracic limbs (Capucchio et al., 2014; Coates & Wininger, 2010). This degeneration is caused by a missense mutation, with incomplete penetrance, within the superoxide dismutase 1 gene (*SOD1*) leading to oxidative damage from the accumulation of superoxide free radicals (Awano et al., 2009; Holder et al., 2014; Wininger et al., 2011).

Diagnosis is difficult as it mostly relies on the exclusion of other diseases, with definitive confirmation only possible with a postmortem histological assessment of the spinal cord tissue (Capucchio et al., 2014). A genetic test is now available to confirm diagnosis (Holder et al., 2014). The difficulty of obtaining a diagnosis is furthered by the rapid deterioration of the animal, with many dogs losing the ability to walk within six to twelve months of the onset of

symptoms (Capucchio et al., 2014), worsening to complete tetraplegia and impacts on brain stem function with further time (Coates & Wininger, 2010). There are no options or treatments for the disease (Coates & Wininger, 2010), typically leaving owners with supportive care until quality of life decreases to the point where euthanasia may be considered.

2.4.2 – Haemophilia A

Haemophilia A is a recessive X-linked trait in both humans and dogs which results in the disease being prevalent within males and uncommon in females, who are often carriers of the gene (Aslanian et al., 2014; Parry et al., 1988). German shepherd dogs are one of the dog breeds most affected by haemophilia (Aslanian et al., 2014). The disease was first reported in the breed as early as the 1950's–60's (Mustard & Packham, 1968), with increasing frequency since that time as popular stud dogs unintentionally spread the allele (Wootton, 1988). Parry et al. (1988) reported that 15 out of 17 dogs with haemophilia A could be traced back to a single prolific sire: Canto. Wootton (1988) reported that this dog sired 100 litters with many of his sons performing well at dog shows, thereby resulting in them also becoming prolific sires leading to the rapid spread of his genes.

Haemophilia A is a bleeding disorder caused by the disfunction of or deficiency in coagulation factor VIII (Aslanian et al., 2014; Parry et al., 1988; Stokol et al., 1994). Without sufficient factor VIII, the conversion of fibrinogen to fibrin is slow and ineffective, causing the platelet plug to be unstable and bleeding to continue from the damaged site (Stokol et al., 1994). Due to the high risk of bleeding, affected dogs are usually diagnosed very early, at less than a year of age (Aslanian et al., 2014) as symptoms can be seen following minor injuries, vaccinations or surgery.

The level of deficiency is measured by the amount of factor VIII coagulant within the blood serum and classified to one of three levels of severity (Aslanian et al., 2014; Parry et al., 1988). Mild, with factor VIII coagulant levels of 6–20%, moderate, with factor VIII coagulant levels of 2–5%, and severe, with factor VIII coagulant levels of <2% (Aslanian et al., 2014). Dogs with mild cases typically only experience haemorrhage after a major trauma or surgery, whereas moderate cases may also experience hemarthrosis and spontaneous haemorrhage (Aslanian et al., 2014; Stokol et al., 1994). Severe cases can experience prolonged and severe haemorrhage after injury and surgery, spontaneous hemarthrosis and intramuscular haemorrhage, and are at risk of life-threatening internal haemorrhage (Stokol et al., 1994).

Prognosis of haemophilia A in canines is fair nowadays as methods have been developed to provide factor VIII replacement therapy with blood components such as cryoprecipitate, fresh-frozen plasma, packed red blood cells or fresh whole blood (Aslanian et al., 2014). However, effectiveness of treatment can depend on the level of severity, the resources available to the owner and any concurrent diseases or comorbidities (Aslanian et al., 2014; Stokol et al., 1994).

2.4.3 – Hereditary Multifocal Renal Cystadenocarcinoma and Nodular Dermatofibrosis

Canine hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis is a heritable disorder first described in GSDs in 1985 (Lingaas et al., 2003). It is characterised by bilateral multifocal tumours on the kidneys, often accompanied by dense collagen fibre nodules in the skin and uterine leiomyomas in female dogs (Conrado et al., 2020; Lingaas et al., 2003). The condition is due to an autosomal dominant mutation in the Birt–Hogg–Dubé locus and is thought to have a homozygous lethal effect (Lingaas et al., 2003).

Diagnosis can typically be attained through palpation or radiography of the enlarged abnormal kidneys, alongside symptoms including increased drinking, blood in the urine, depression, fever and loss of appetite (Moe & Lium, 1997). Conformation of diagnosis is through histological examination of renal tissue. As it is typically a later onset disease with a rapid progression, treatment is mostly limited to supportive care and removing tumours (Moe & Lium, 1997).

2.4.4 – Megaoesophagus

Megaesophagus is a complex condition which can either be congenital or acquired in response to a dog contracting another disease that impacts neuromuscular function (Bell et al., 2022; Johnson et al., 2009). Both forms of the condition disproportionately impact GSDs compared to other breeds. Congenital idiopathic megaoesophagus is caused by faulty acetylcholine receptors which impact the nervous function of the musculature of the oesophagus, leading to oesophageal dilation and decreased peristalsis (Bell et al., 2022; Johnson et al., 2009; Mace et al., 2012).

Puppies with congenital idiopathic megaoesophagus are often diagnosed soon after weaning due to frequent coughing, regurgitation and weight loss, with some animals also presenting with generalised weakness and/or aspiration pneumonia (Mace et al., 2012). Diagnosis is typically achieved through thoracic radiographs, using barium contrast in some cases to observe oesophageal dilation (Johnson et al., 2009). Prognosis is mixed and highly dependent on the level of severity, with affected animals requiring highly specialised care and constant

monitoring, including a high calorie liquid diet fed in small portions frequently while in an upright position to use gravity to assist swallowing (Bell et al., 2022; Mace et al., 2012). However, some cases may spontaneously resolve by one year of age, possibly due to delayed nerve development (Bell et al., 2022).

In some breeds congenital idiopathic megaoesophagus is inherited in an autosomal recessive fashion, however, Bell et al. (2022) proposed that GSDs showed a more complex method of inheritance, with a variable number tandem repeat in the Melanin Concentrating Hormone Receptor 2 gene (*MCHR2*) on chromosome 12. Additionally, it was also found that oestrogen was a protective factor for female dogs due to its role in relaxing the oesophageal sphincter and facilitating food passage, leading to a sex bias where males are twice as likely to be affected by congenital idiopathic megaoesophagus, independent of their body size (Bell et al., 2022).

2.4.5 – Musculoskeletal Disorders

Due to the intense physical requirements placed on police and military working dogs, musculoskeletal disorders are, unsurprisingly, one of the leading causes of retirement, death or euthanasia of GSDs (Evans et al., 2007). This includes degenerative joint diseases such as arthritis (Evans et al., 2007; Moore et al., 2001; Parr & Otto, 2013; Worth et al., 2013) particularly in the lumbosacral and coxo-femoral joints (Worth et al., 2013), spinal disease (Evans et al., 2007; Moore et al., 2001; Parr & Otto, 2013; Worth et al., 2013), and hip/elbow dysplasia (Wahl et al., 2008; Worth et al., 2013). While these issues are known to be heritable, GSDs have a particular predisposition compared to other breeds (Evans et al., 2007; Moore et al., 2001), and little is known about their mode of inheritance.

Many studies have assessed the possible heritability of some of these diseases (Table 3) with greatly varying results. Many factors and variables can influence these results such study size, the measure being assessed, the rate of inbreeding in the study population, as well as inconsistent assessment of subjective measures.

Table 3 – Comparison of heritability scores of some different complex musculoskeletal diseases from various studies on populations of German Shepherd Dogs, arthritis and spinal disease have been excluded from this table due to low to no papers presenting genetic parameters.

Source	Disease	Heritability	Study size
Oberbauer et al. (2017)	Hip dysplasia	0.58 ± 0.03	107,048
Soo et al. (2014)	Hip dysplasia	0.33 ± 0.08	804
Hamann et al. (2003)	Hip dysplasia	0.26 ± 0.03	38,604
Karsada et al. (2007)	Hip dysplasia	0.56 ± 0.01	21,828
Sturaro et al. (2006)	Hip dysplasia	0.15	18,696
Stock et al. (2011)	Hip dysplasia Elbow dysplasia	0.25 ± 0.01 0.31 ± 0.02	48,367
Oberbauer et al. (2017)	Elbow dysplasia	0.26 ± 0.08	37,233
Gluding et al. (2021)	Lumbosacral transitional vertebrae	0.27	27,597
Berg et al. (2025)	Lumbosacral transitional vertebrae	0.06 ± 0.01	2,843

2.4.6 – Pancreatic Acinar Atrophy

Pancreatic acinar atrophy is characterised by the atrophy of the acinar cells responsible for producing digestive enzymes in the digestive tract. The deficiency in the organ's secretory capacity leads to a functional diagnosis of exocrine pancreatic insufficiency (Moeller et al., 2002; Tsai et al., 2011, 2013; Wiberg, 2004). Onset of the disease is typically seen through excess appetite, weight loss and voluminous faeces (Wahl et al., 2008; Wiberg, 2004). Onset of symptoms is generally at five years of age but can be as early as 13 months old (Moeller et al., 2002). Clinical disease, however, may not be evident until 90% of pancreatic function is already lost (Tsai et al., 2011, 2013).

Exocrine pancreatic insufficiency can be measured through a serum canine trypsin-like immunoreactivity radioimmunoassay (Moeller et al., 2002; Tsai et al., 2011, 2013; Wiberg, 2004). A low reading is likely indicator that pancreatic acinar atrophy is the cause of the exocrine pancreatic insufficiency (Moeller et al., 2002). A definitive diagnosis requires histological examination of the pancreatic tissue (Tsai et al., 2011, 2013). The only available treatment is ongoing replacement of digestive enzymes which gives a mixed prognosis depending on the animal's response to the exogenous enzymes (Wahl et al., 2008).

Pancreatic acinar atrophy is a progressive autoimmune disease (Moeller et al., 2002; Wiberg, 2004) that had been thought to be inherited in either an autosomal recessive patten, or dominant with incomplete penetrance (Wiberg, 2004). However, Tsai et al. (2013) has since proposed a

more complex mode of inheritance involving the major histocompatibility complex in GSD populations.

2.4.7 – Pituitary Dwarfism

Pituitary dwarfism has been seen in GSD's since as early as the 1940's (Voorbij & Kooistra, 2009). It is a rare endocrinopathy caused by an autosomal recessive mutation in the LIM Homeobox 3 (*LHX3*) gene (Kitzmann et al., 2021; Nicholas, 1978). This results in the poor formation of the adenohypophysis, or anterior pituitary (Kitzmann et al., 2021; Voorbij & Kooistra, 2009). Due to the anterior pituitary's role in hormone production, affected dogs typically experience a deficiency in growth hormone, thyroid stimulating hormone and prolactin, however, no difference is seen in levels of adrenocorticotrophic hormone (Kitzmann et al., 2021).

Most affected dogs are diagnosed by three to four months of age through phenotypic symptoms (Voorbij & Kooistra, 2009). These include extremely stunted growth, retention of the puppy coat, symmetrical bilateral alopecia and skin hyperpigmentation (Kitzmann et al., 2021; Voorbij & Kooistra, 2009). Diagnosis can be confirmed by administering growth hormone releasing hormone, ghrelin or α-adrenergic drugs. In an affected dog, growth hormone levels remain low after stimulation, whereas an unaffected dog will display a growth hormone spike in response (Kitzmann et al., 2021; Voorbij & Kooistra, 2009). A genetic test is also available to identify carriers, who are asymptomatic, as well as to diagnose affected animals (Kitzmann et al., 2021).

There are several treatment options for pituitary dwarfism, including levothyroxine, porcine growth hormone and progestogens, all of which typically provide a guarded prognosis (Kitzmann et al., 2021). These treatment options can prolong an animal's lifespan compared to untreated dogs, with side effects being uncommon. Additionally, dogs treated with growth hormone and/or progestogens are significantly taller and heavier than other, untreated, affected dogs (Kitzmann et al., 2021). Without treatment, affected dogs usually either die or are euthanised prior to five years of age (Kitzmann et al., 2021; Voorbij & Kooistra, 2009), whereas with treatment they can reach ten years of age or more (Kitzmann et al., 2021).

2.4.8 – von Willebrand Disease

Von Willebrand disease (VWD) is one of the most common bleeding disorders in dogs and is caused by various defects within the von Willebrand factor (VWF) protein (Aslan et al., 2016; Brooks & Catalfamo, 2022; Stokol et al., 1994). The VWF proteins are an essential

glycoprotein involved in the adhesion of platelets at bleeding sites (Brooks & Catalfamo, 2022; Stokol et al., 1994). The various defects within VWF protein led to an array of functional and quantitative deficiencies, which results in the disease having different types and subtypes (Aslan et al., 2016; Brooks & Catalfamo, 2022; Stokol et al., 1994).

Type one VWD is a quantitative deficiency of <50% of expected VWF, where severity is variable depending on the level of deficiency (Brooks & Catalfamo, 2022; Stokol et al., 1994). Type two is a functional defect within VWF where four further subtypes express different dysfunctional interactions with platelets and factor VIII. Type three is a complete lack of any VWF within the plasma. As each type is caused by different errors in the VWF protein, inheritance is also dependent on the type of disease. Types two and three are typically autosomal recessive while type one is complex and thought to be dominant with incomplete penetrance, hence causing the variable severity of disease. Type one is the most commonly seen version of VWD in GSDs (Brooks & Catalfamo, 2022; Stokol et al., 1994).

Regardless of disease type, severe and prolonged bleeding following trauma is the main symptom and can be diagnosed through a species specific VWF plasma assay (Brooks & Catalfamo, 2022). Management of VWD primarily involves preventing and minimising haemorrhage where possible such as having transfusions and coagulation promoting drugs ready in situations where surgery may be necessary (Brooks & Catalfamo, 2022; Stokol et al., 1994).

2.4.9 – Previously Undescribed Genetic Mutations in German Shepherd Dogs

2.4.9.1 – BOP1

Block of Proliferation or Ribosomal Biogenesis Factor; *BOP1,* is a well described gene within literature due to its involvement with the PeBoW complex alongside genes Pescadillo (*PES*) and WD Repeat Domain 12 (*WDR12*) (Ahn et al., 2016; Borah et al., 2019). Together these genes modulate and process ribosomal RNA. Alongside regulation of ribosome biogenesis, *BOP1* has also been identified to contribute to regulation of gene expression, DNA synthesis and metabolic regulation, hence disruption of the gene can cause severe deleterious effects (Borah et al., 2019). Additionally, several sources have identified an upregulation of *BOP1* within cancerous tissue, making it a potentially valuable biomarker for targeted therapy (Ahn et al., 2016; Chung et al., 2011; Wu et al., 2023).

2.4.9.2 – IQANK1

IQ Motif and Ankyrin Repeat Containing (*IQANK1*) gene, previously known as *FAM83H-AS1*, is less well known. Dronova et al. (2024) identified the gene's role in the production and organisation of cytoplasmic membranes and the cytoskeleton, however, the products and their functions are not described. Mutations within *IQANK1* have been linked with diseases including amelogenesis imperfecta, affecting tooth enamel (Tachie-Menson, 2020; Wang et al., 2020) and congenital keratoconjunctivitis sicca and ichthyosiform dermatosis, a congenital coat and eye condition (Forman et al., 2012). Neither of these diseases align with the clinical and phenotypic presentation of the affected animals.

2.4.9.3 – THEMIS2

Thymocyte Selection Associated Family Member 2 (*THEMIS2)* is a part of a small family of 'Themis' proteins with shared characterisations (Cheng et al., 2016). Its function however is not well known, though some studies have reported interactions with the immune system (Cheng et al., 2016; Letko et al., 2023; Nabekura et al., 2023). Nabekura et al. (2023) reported that the gene is a regulator of natural killer cell memory formation and may also contribute to pathways causing the downregulation of Ca^{2+} influx. However, they did not report whether a deficiency in *THEMIS2* caused a negative effect on natural killer cells, leaving their relationship unclear. Cheng et al. (2016) identified a similar function of *THEMIS2* in the positive selection of B cells by lowering the activation threshold. Similarly, deficiency in the gene did not impact B cell development, also leaving their interaction unclear (Cheng et al., 2016). Given this information, *THEMIS2* appears to interact with the immune system, but the extent of which and the mechanisms involved are unclear. Furthermore, clinical results are insufficient to rule out an issue with the immune system, but currently there is also little evidence to suggest that the primary cause of the phenotype seen in the affected animals is due to a dysfunctional immune system.

2.4.9.4 – ZNF517

In dogs, the C2H2 zinc finger (*C2H2-ZNF*) gene superfamily consists of 184 genes across 57 clusters (Tadepally et al., 2008), with a large subfamily of these genes also containing a Krüppel-associated box (*KRAB*), one of which is *ZNF517* (Lorenz et al., 2010). Classical *C2H2-ZNF* are one of the largest gene families involved in transcriptional gene regulation and related to DNA binding, however, few genes within the superfamily have been explored in detail and therefore their specific interactions remain unknown (Lorenz et al., 2010; Tadepally

et al., 2008). The *KRAB* effector domain is always located at the N-terminal of the zinc finger gene and is a transcriptional repressor (Tadepally et al., 2008).

2.5 – Gene Characterisation and Discovery

Given the rapid development of genetic mapping technology seen in recent years, it is quicker, cheaper and easier than ever to genetically test animals (Baker et al., 2019; Ng & Kirkness, 2010). This can be done for several reasons such parentage confirmation, health screening, which is required in some cases for breed registration, breed origin analysis, particularly popular with mixed bred shelter dogs, and for selective breeding to focus on very specific traits such as producing dogs of a certain colour (Baker et al., 2019).

Depending on how much of the genome is being sequenced, there are several methods that can be employed, including de novo or reference based whole genome sequencing (WGS) and partial sequencing of specific SNPs (Baker et al., 2019; Ng & Kirkness, 2010). Reference based WGS can only be utilised where a closely related reference genome is already available and works by mapping individual reads to matching SNPs in the reference, producing a consensus sequence similar to that of the reference genome (Ng & Kirkness, 2010). Where no appropriate reference genome exists, de novo WGS is applied where short length reads, though longer reads tend to be more accurate, are compared and overlapped at matching locations to produce a longer sequence. However, there are now many different reference genomes publicly available for an array of different breeds (Benson et al., 2012). While de novo WGS is more resource heavy, it is also able to capture and include novel sequences not present, and therefore not mapped, in a reference based WGS (Ng & Kirkness, 2010).

Partial genome sequencing can be a useful screening tool but can only be applied when the location is known for a well-documented SNP that is related to or causative of a certain phenotype (Ng & Kirkness, 2010). These tests can often also be restricted to specific breeds due to linkage disequilibrium, the non-random assortment of alleles in different loci, or haplotypes, causing segments of DNA to be inherited together due to selective breeding (Baker et al., 2019). Partial sequencing also runs the risk of missing larger-scale issues like structural variants that are impacting protein production and function (Ng & Kirkness, 2010).

With some or all an animal's genome sequenced, comparisons to the expected "normal" functioning sequence can be made and variants such as missense SNPs, or frame shift indels

can be isolated (Ng & Kirkness, 2010). The impact of these variants on the related protein's structure and function can then be predicted using tools such as SnpEff (Cingolani, Platts, et al., 2012) or Ensembl's Variant Effect Predictor (VEP) (McLaren et al., 2016) and related to phenotypic impacts such as disease traits.

This thesis will undertake a genomic investigation of the probable causes of an unknown and previously undescribed exercise induced disease presentating in a family of police dogs.

Chapter 3 – Methods

3.1 – Animal Background

The 'W' litter was born in December of 2022 and includes three male, and four female puppies (Figure 5). Apart from the dam having a mild allergic skin and ear condition, both parents are healthy, pregnancy and whelping were normal.

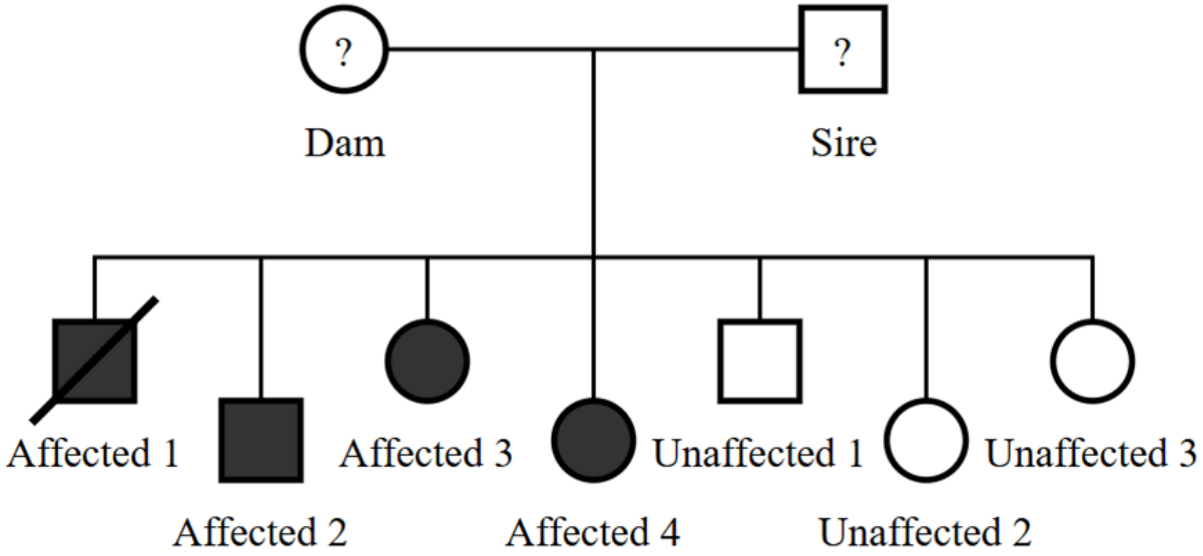

Figure 5 – 'W' litter to parents of unknown genotype. A square indicates a male animal and a circle a female. Filled in shapes indicate affected dogs. Strikethrough animals are deceased.

Healthwise, affected 2 was underweight at 12 weeks of age and a week later was vomiting from an unknown cause that was resolved. Affected 4 had poor conformation of the back legs, with cow hocks and knock-knees. These traits comprised an angular deformity of the hind legs due to incorrect growth of the tibia and fibula (McKee, 2010), causing her to have a shuffling gait and mild hip pain, although radiographs showed no visible issues. At eight to nine months of age, four out of the seven puppies (affected 1, 2, 3 and 4) had begun to show extreme intolerance to exercise. To varying extents, each of the affected puppies would stop during exercise (such as chasing a ball), refuse to move, and display clear signs of pain. Affected 4 was also observed to pant excessively. Following a five-to-ten-minute break, the puppies were able to move normally again.

Affected 1 was presented to Massey University veterinary hospital for assessment at nine months of age in September of 2023. The physical exam was normal apart from reluctance to move following exercise without an apparent localisation of pain. An orthopaedic exam and radiographs of the pelvis, spine, forelimbs and hindlimbs also revealed no abnormalities.

Following this, a blood panel and blood gas levels were assessed after exercise which revealed extremely high levels of creatine kinase (CK), aspartate transferase (AST) and alanine aminotransferase (ALT). These three enzymes are all common biomarkers used as indicators of both skeletal and cardiac muscle injury (Hunt, 2018); hence, high levels were strong indication of muscle damage. Elevated ALT can also be caused by damage to the liver, however, given the combination of elevated enzymes; and no other indicators of liver damage, it is more likely to be due to damaged muscle tissue (Hunt, 2018).

During the following months, all of the dogs in a 'W' litter participated in a post exercise blood test to assess if they also had elevated levels of CK, AST and ALT. A muscle biopsy was also taken from affected 1 to assess the physical muscle structure.

In early 2025, unaffected 1 became the only dog from the 'W' litter to complete training and become a patrol dog. The remaining puppies were rehomed, with unaffected 2 and 3 living a normal life as pets. Affected 2 and 4 are also living relatively normal lives, only showing pain and reluctance to move after excessive exercise, which can be easily managed. Both dogs have also had repeated ear infections, likely an allergic ear disease. Affected 3's condition worsened since 2023, with a low threshold of exercise tolerance, particularly in summer due to overheating. Affected 1's condition was relatively stable until September 2024 when he became very reluctant to move, and experienced difficulty breathing and subsequently died on the way to the vet. We were not aware of his death at the time and therefore further investigation on the cause of death was not able to be performed.

The dam had no further litters and has now been retired from the breeding programme. The sire had produced four other litters to three different bitches before the 'W' litter began showing any symptoms of possible heritable disease, all born between January and November of 2023 (Figure 6). Within these litters there were several different health issues; however, none were similar to that of the 'W' litter. This included one, now deceased, puppy in the 'A' litter that suffered from a patent right aortic arch, a congenital heart defect. Two puppies in the 'B' litter: one with a probable allergic ear disease and another with ear lesions of an unknown cause. At least one puppy in the 'I' litter had atopic dermatitis (an inflammatory skin disease). In the 'U' litter, two puppies had hypertrophic osteodystrophy (an auto-inflammatory disease of the bones), and another puppy had teeth malocclusion.

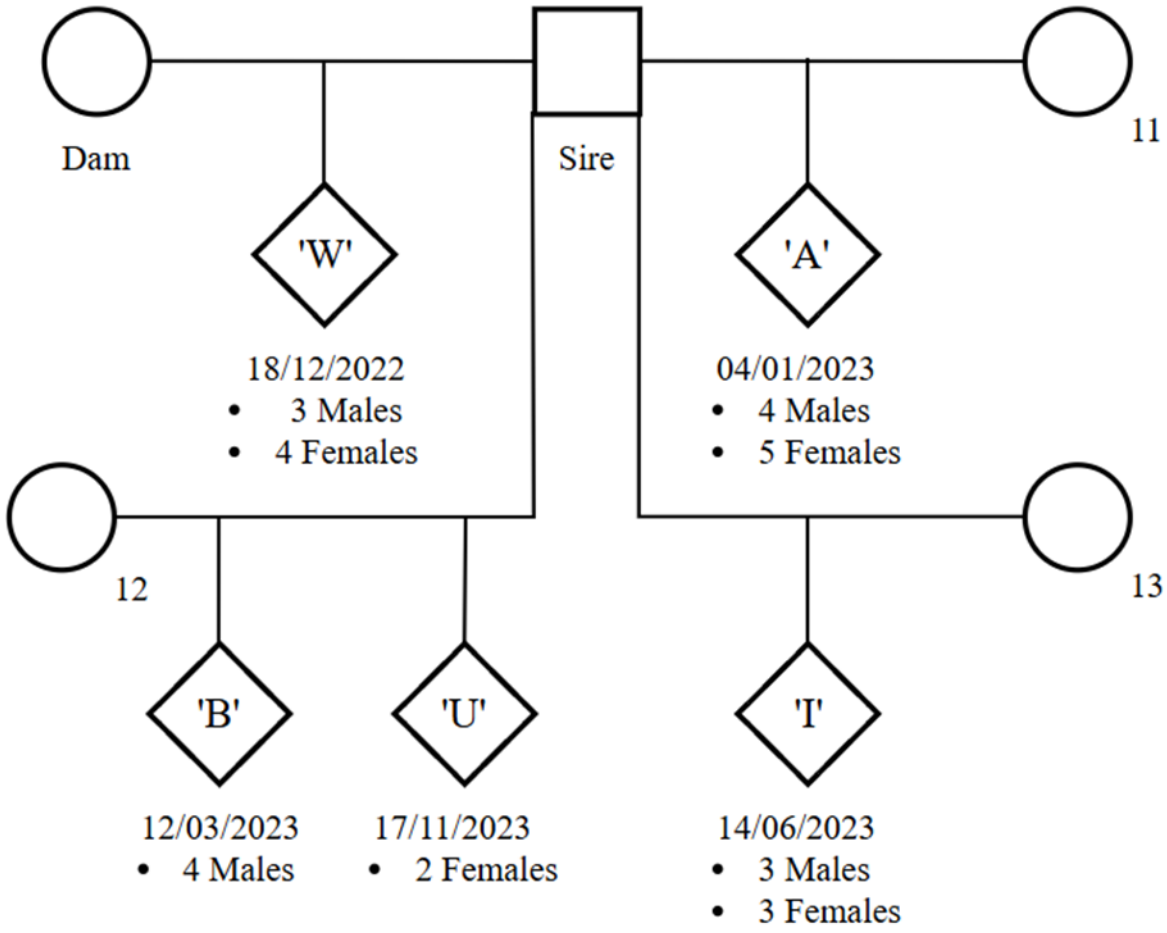

Figure 6 – The sire's offspring to different bitches with the litters date of birth and numbers of each sex born. The letter within the diamond indicates the name applied to the litter in the naming scheme.

3.1.1 – Inbreeding Coefficients

Using Wrights equation (Wright, 1922) and the tabular method, inbreeding coefficients for and between individuals can be calculated to inspect the relatedness of animals within a pedigree. Equation one (below) is first used to find the diagonal elements, the animals own inbreeding coefficient, F_X.

$$F_X = \sum_{CA=1}^{k} \left[\left(1/2\right)^{n1+n2} (1 + F_{CA}) \right]$$

Where, *CA* is a common ancestor of the dam and sire of individual X with *k* being the number of common ancestors in X's pedigree. *n1* and *n2* is the number of generations between the common ancestor and the sire of X, and the dam of X, respectively. F_{CA} is the inbreeding coefficient of the common ancestor. If an animal is not known to be inbred, the diagonal element for the individual is 1. An animal that is inbred will have a value greater than 1.

Wright's coefficient of relationships between individuals *X* and *Y*, R_{XY}, is calculated using equation two:

$$R_{XY} = \frac{\sum_{CA=1}^{k}\left[\left(1/2\right)^{n1+n2}\left(1 + F_{CA}\right)\right]}{\sqrt{1 + F_X}\sqrt{1 + F_Y}}$$

Where, *CA* is a common ancestor of individual X and Y with *k* being the number of common ancestors between the two animals. *n1* and *n2* is the number of generations between the common ancestor and individual X and Y, respectively. Finally, F_{CA}, is the inbreeding coefficient of the common ancestor, and F_X and F_Y are the inbreeding coefficients of individuals X and Y. If animals X and Y have no common ancestors, the off-diagonal element for the two is zero.

3.1.2 – Blood Tests

Blood was collected from dogs following exercise with two times 1.8mL in a BD Vacutainer® Serum Tube (Becton, Dickinson and Company, New Jersey, USA). These were submitted to a commercial diagnostic laboratory (IDEXX NZ Ltd., Palmerston North, NZ) for analysis. Blood samples were then stored in a -80°C freezer for later DNA analysis.

Analysis of serum CK, AST and ALT were run to confirm which animals were affected within the 'W' litter. Affected 1 had two samples taken, one prior to and one post exercise, with both results significantly greater than the reference range compared to a normal metabolic response to exercise. Affected 2 also had a second sample taken following heavier exercise, approximately two weeks after the first, as the results of the first sample were not consistent with the symptoms he was experiencing. The dam was also sampled in 2025 to rule out the possibility of subclinical disease. Her results indicated a slight elevation in AST but not to the point of clinical concern. Creatine kinase and ALT were both at normal levels.

3.1.3 – Muscle Biopsy

Four muscle biopsies were taken from affected 1 under anaesthesia, ventilated on the 'Zoll' ventilator to avoid using circuits or machines containing inhalation agents to prevent potential malignant hypothermia, a hereditary condition where certain anaesthetic agents can trigger a cascade of excessive enzyme production in the muscles. Alfaxalone was used to maintain anaesthesia to help decrease the myoclonus effects, the involuntary contraction of muscle fibers, which result from the administration of propofol, so that the muscle biopsies were accurate.

One biopsy sample was taken from each of the biceps femoris, epaxial, vastus lateralis (quadriceps) and a proximal forelimb (trapezius) muscle. At each of the four sites, an incision was made through the skin and subcutaneous tissue overlying the muscle, and an approximately 5mm x 20mm longitudinal section of muscle was biopsied.

A 1-2mm square piece of each biopsy was separated and fixed in glutaraldehyde for electron microscopy. Samples in glutaraldehyde were post-fixed in osmium tetroxide and embedded into epoxy resin. Thin sections from the resin blocks were stained using lead citrate/uranyl acetate and viewed using a FEI Tecnai G2 Biotwin Transmission Electron Microscope at the Manawatu Microscopy and Imaging Centre (Massey University, Palmerston North, NZ). The remainder of each biopsy was placed in 10% formalin and routinely processed for histology, with 5 micrometre thick sections cut and stained with haematoxylin and eosin.

3.2 – Genetic Testing

3.2.1 – DNA Sequencing

The DNA was extracted from whole blood using the Qiagen MagAttract HMW DNA Kit (Cat67563 Qiagen NZ). Sequence libraries were prepared with the Illumina DNA Prep Kit (Illumina, NZ) and sequenced using the Illumina NovaSeq 6000 using an SP flowcell (2×150bp reads; conducted at GeneMark, Hamilton, NZ). Four raw FastQ files were produced for each animal from the forward and reverse reads of two flow cell lanes. Poor quality reads and artifacts of sequencing were removed using Trimmomatic v0.39 (Bolger et al., 2014).

SAMtools v1.13 (Danecek et al., 2021) was used to map the FastQ files onto the CanFam4 reference genome (UU_Cfam_GSD_1.0) with bwa-mem2 v2.2.1 (Vasimuddin et al., 2019). One SAM file per animal was produced and were subsequently transformed into BAM files with BAMtools v2.5.1 (Barnett et al., 2011). Reads were mapped and aligned to the CanFam4 genome (Wang et al., 2021) by bwa-mem2 v2.2.1 (Vasimuddin et al., 2019) and variant calling by GATK Haplotypecaller v4.3.0.0 (Poplin et al., 2017).

3.2.2 – DNA Processing

Following sequencing, filtering methods were applied to the raw VCF using SnpEff v5.2 (Cingolani, Platts, et al., 2012). Filters were applied using BCFtools v1.10.2 (Danecek et al., 2021) for i) a recessive trait from both parents, where only the affected dogs were homozygous, ii) a dominant traits from a single affected parent, where non unaffected dogs were homozygous

recessive, and iii) a de novo mutations where the affected dogs displayed a different haplotype to either parent.

These initial candidate variants were further filtered down to those of high or moderate impact using SnpSift v5.2 (Cingolani, Patel, et al., 2012). Manual examination of the genome and candidate genes, including comparison to the RNA sequence was done in the integrative genomics viewer (IGV) (Robinson et al., 2011). The canine reference RNA, produced by the Dog Biomedical Variant Database Consortium (Jagannathan et al., 2019), was downloaded in the form of raw FastQ files using SRA-Toolkit v3.0.2 (National Institutes of Health, 2022). This was then mapped to CanFam4 with STAR software v2.7.9a (Dobin et al., 2012) and compiled with SAMtools v1.6 (Danecek et al., 2021). Candidate variants remaining after applying the filters were then compared to the RNA sequence data of 17 different tissue types to ensure that they were located within an exon of a gene, as expected for moderate or high impact genes. Final candidate genes were also filtered to remove any moderately frequent variants present in a catalogue of publicly available multibreed whole genome sequences (Plassais et al., 2019).

Chapter 4 – Results

4.1 – Clinical Results

4.1.1 – Blood Tests

All four affected dogs that presented with exercise intolerance, showed a markedly elevated level of serum CK, AST and ALT (Table 4) compared to the upper reference range limits of 609 IU/L (international units per litre), 79 IU/L and 75 IU/L, respectively (IDEXX NZ Ltd., Palmerston North, NZ). The three unaffected dogs not exhibiting any affected traits were all within normal limits for all tests.

Table 4 – Blood serum levels (IU/L) of creatine kinase (CK), aspartate transferase (AST) and alanine aminotransferase (ALT) of the affected and unaffected dogs before or after exercise.

Dog ID	Pre/Post exercise	Date taken	Blood serum concentrations		
			CK (IU/L)	AST (IU/L)	ALT (IU/L)
Affected 1	Pre exercise	16/10/2023	*20813*	*704*	*619*
	Post exercise	16/10/2023	*34160*	*928*	*673*
Affected 2	Post exercise	21/11/2023	374	50	*177*
	Post exercise	06/12/2023	*190150*	*3603*	*837*
Affected 4	Post exercise	21/11/2023	*12170*	*401*	*282*
Affected 3	Post exercise	21/11/2023	*3455*	*259*	*775*
Unaffected 1	Post exercise	20/11/2023	293	41	54
Unaffected 3	Post exercise	21/11/2023	184	35	42
Unaffected 2	Post exercise	27/11/2023	127	39	53
Dam	Post exercise	04/07/2025	405	*108*	54
Normal upper limit			**609**	**79**	**75**

4.1.2 – Histology and Electron Microscopy

The muscle biopsies taken from affected 1 were histologically examined under standard light microscopy (Figure 7, A – D) and electron microscopy (Figure 8, A – E). Figure 7A and 7B shows a longitudinal section of the epaxial muscle at magnifications of 40x, and 200x, showing a degenerating muscle fibre, and a mix of different inflammatory cells. In comparison, Figure 7C shows a transverse section of the same muscle at a magnification of 400x and shows a large range in size, shape and staining of the muscle fibres, with the presence of small, angular fibres (red arrows), which is an indication of atrophy. There was also a degenerating myofiber with inflammatory cells present (green arrow). A transverse section of the biceps femoris at 400x magnification, showed the acute degeneration of a muscle fibre (red arrow), accompanied by

other myofibers at different stages of degeneration (Figure 7D). However, this degeneration was not paired with an inflammatory response.

The samples from affected dog 1 examined under electron microscope were also compared to a sample of the triceps muscle from a normal, unaffected dog. The unaffected dog sample (Figure 8A and 8B) presented normally shaped and spaced mitochondria between the muscle fibres (red arrows). The biceps femoris of the affected dog (Figure 8C and 8D) showed a large number of excessively large and poorly shaped mitochondria between muscle fibres and beneath the cell membrane in the sarcolemmal space (green arrows on Figure 8C and 8D). The epaxial muscle of the affected dog (Figure 8E and 8F) presented large mitochondria that fill the space between normal myofibrils. Some of these also have an abnormal shape, such as the one indicated by a blue arrow that resembles a golf club. There are large numbers of mitochondria under the cell membrane (yellow arrow), located near the nucleus (N). There was no evidence of significant damage to the myofibers in Figure 8 C – F.

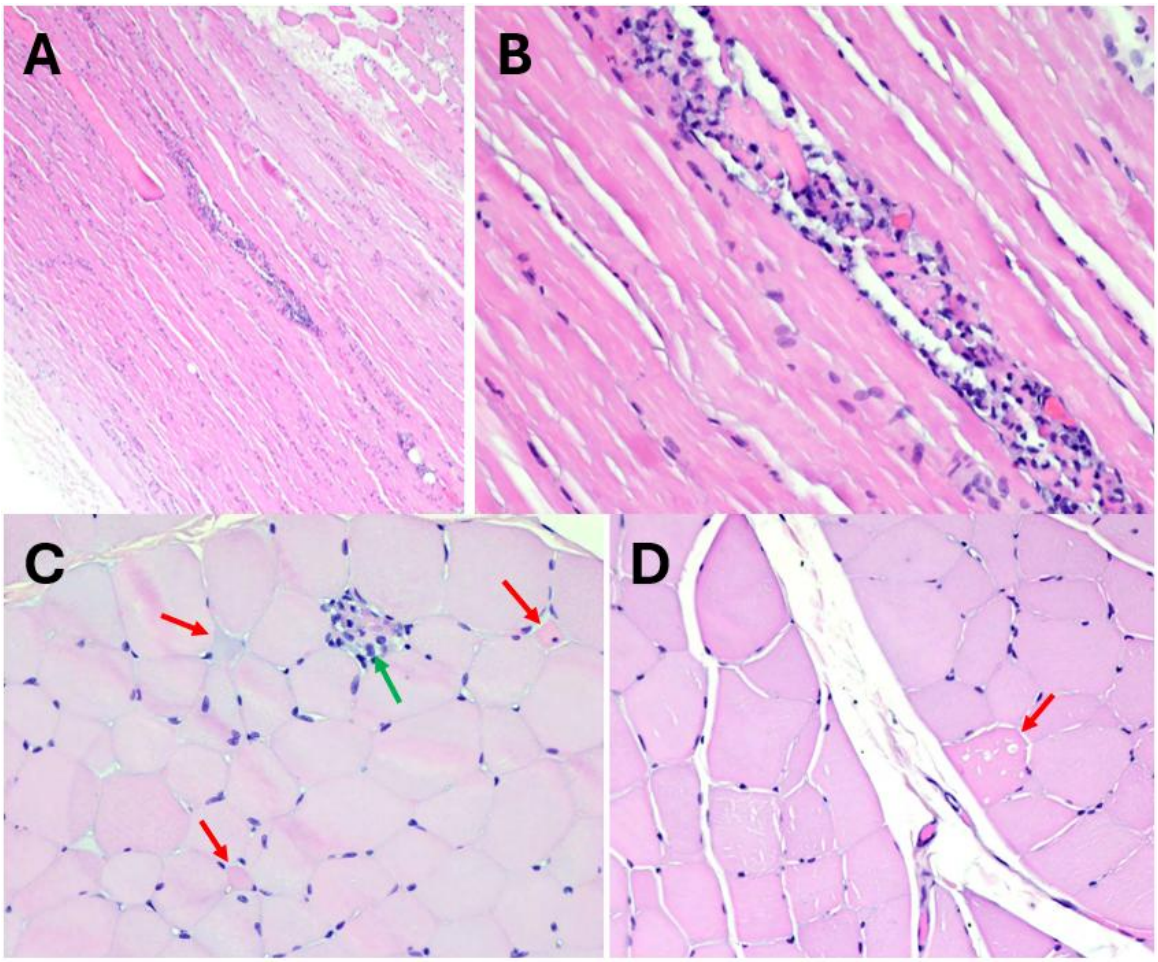

Figure 7 – Photomicrographs of skeletal muscle from affected 1, stained with heamatoxylin and eosin A 40x magnification, B 200x magnification, C 400x magnification and D 400x magnification of acute degeneration.

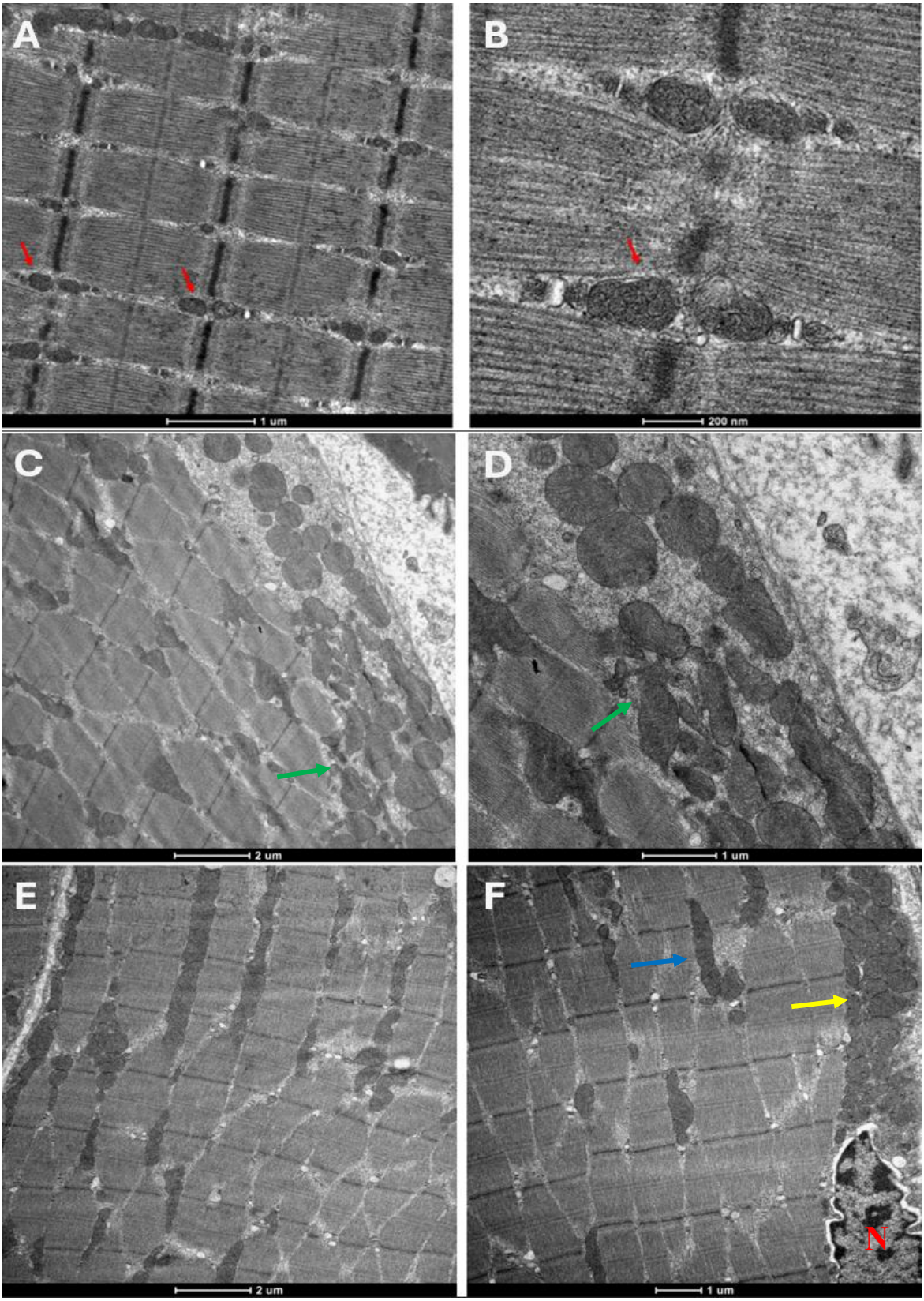

Figure 8 – Electron microscopy of skeletal muscle from a normal dog (A and B) and affected 1 (C – F).

4.2 – Genetic Findings

4.2.1 – Pedigree Analysis

Analysis of the dam and sire's individual pedigrees revealed two common ancestors, animals two and three (Figure 9). This pedigree was used to calculate inbreeding for and between individuals using the tabular method (Table 5).

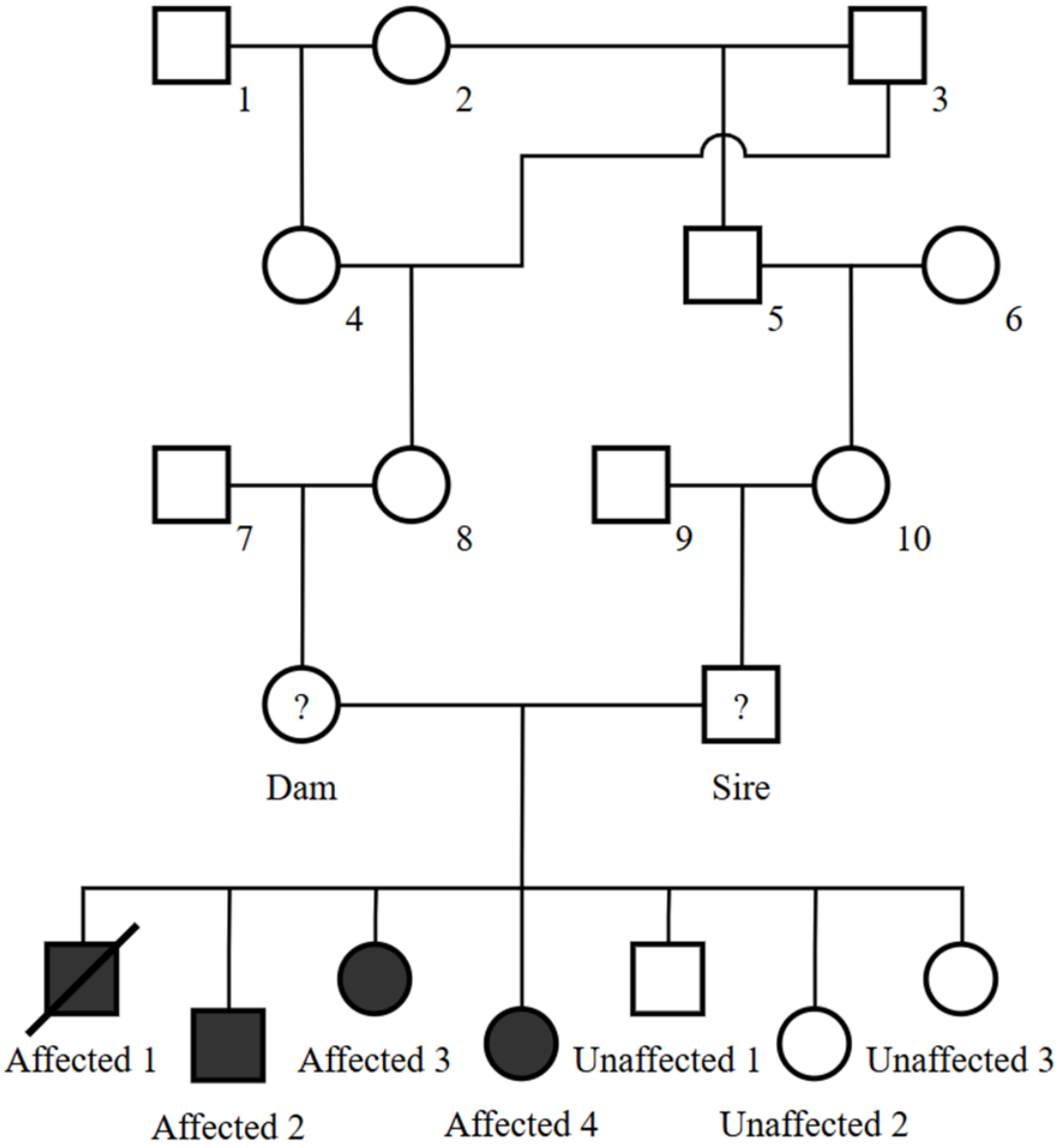

Figure 9 – Dam and sire's family tree with common ancestors 2 and 3. A square indicates a male animal and a circle a female. Filled in shapes indicate affected genotype dogs. Strikethrough animals are deceased.

Table 5 – Tabulated inbreeding coefficient of animals from the dam and sires' pedigree in Figure 9. Values on the diagonal represent the inbreeding coefficient of the individual(s) plus one, meaning that numbers greater than one indicate a level of inbreeding. Off-diagonal values describe the degree of relatedness between individual animals.

Animal ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Dam**	**Sire**	**'W' litter**
1	1												
2	0	1											
3	0	0	1										
4	0.5	0.5	0	1									
5	0	0.5	0.5	0.25	1								
6	0	0	0	0	0	1							
7	0	0	0	0	0	0	1						
8	0.25	0.25	0.5	0.5	0.375	0	0	1					
9	0	0	0	0	0	0	0	0	1				
10	0	0.25	0.25	0.125	0.5	0.5	0	0.1875	0	1			
Dam	0.125	0.125	0.25	0.25	0.1875	0	0.5	0.5	0	0.09375	1		
Sire	0	0.125	0.125	0.0625	0.25	0.25	0	0.09375	0.5	0.5	0.046875	1	
'W' litter	0.0625	0.125	0.1875	0.15625	0.21875	0.125	0.25	0.296875	0.25	0.296875	0.523438	0.523438	1.023438

The calculated inbreeding coefficients showed that only the 'W' litter were inbred within this pedigree where the individuals in the litter were 2.34% identical by decent. While the other animals were not inbred, there was a high frequency of related individuals between the sire and the dam's ancestry. It should be noted that this method is limited to available information. As only a four-generation pedigree for the dam and sire was available, it is assumed that the other animals are not inbred and there were no other common ancestors for the 'W' litter. Additionally, this does not account for the baseline level of inbreeding that exists within GSDs.

In this instance, the inbreeding likely occurred because only three generations of the parent's pedigree were compared, while the common ancestors were in the fourth generation.

4.2.2 – DNA Quality

Several parameters were used to assess the quality of the DNA samples taken from the dogs in the study including average mapping quality scores, error rate and average coverage depth (or average read depth). Average quality and error rates are linked factors, where average quality increases on a logarithmic scale with a decreasing error rate, however, they can be used simultaneously to describe the quality of a genetic sample. Conversely, in the literature average coverage depth is incredibly variable, with little agreement on what constitutes a 'good' average coverage depth. This factor can also be significantly influenced by how much money is being spent as bigger flowcells and more samples per flowcell will impact the subsequent reading. Given that all the animals sampled returned an error rate of ≤1% and the average quality was consistent between animals with a range of 35.0 to 35.6, the quality of the samples is high (Table 6). With the addition of manual examination of the DNA to investigate SNPs of interest, possible false positive flags are likely to be found and removed. Furthermore, the transition/transversion (Ti/Tv) ratio, a quality measurement that varies based on genome region and functionality, was recorded at 2.81 following initial filtering. Generally, a higher Ti/Tv ratio, not greater than four for non-exome exclusive regions indicates a better-quality sample (Wang et al., 2014).

Table 6 – Average quality, coverage depth and error rate of the individual BAM files produced for each animal.

	Average quality	Error rate	Average coverage depth
Sire	35.0	0.0071	12.17
Dam	35.5	0.0054	28.32
Affected 1	35.6	0.0053	14.68
Affected 2	35.5	0.0055	27.10
Affected 3	35.2	0.0063	14.00
Affected 4	35.4	0.0055	20.76
Unaffected 1	35.4	0.0060	23.01
Unaffected 2	35.3	0.0061	17.32
Unaffected 3	35.2	0.0063	16.67
Average	35.3	0.0059	19.34

4.2.3 – Variant Calling

Initial annotation to the CanFam4 genome by GATK produced a total of 5,134,827 SNPs and 2,360,313 indels across the nine dogs. To reduce the number of mutations to be screened different filtering methods were applied using SnpEff. Three models were used to initially filter variants, based on different possible methods of inheritance: a recessive mutation where one allele copy was inherited from both parents in the affected dogs, a dominant mutation where one parent (either sire or dam) passed one copy of the affected allele to affected offspring, and a de novo mutation where the affected allele is only seen in the affected dogs. Additionally, the non-nuclear DNA and allosomes were manually examined for evidence of mutations. Both were ruled out from further investigation when no evidence was found.

This gave an output of four variant calling files alongside the original, unfiltered version, with varying levels of variants, including SNPs, insertions and deletions across the autosomes (Table 7). De novo mutations, where affected dogs would have a different haplotype to the nonaffected dogs, yielded no variants matching the filtering criteria of a haplotype present exclusively and consistently across the affected animals, regardless of the unaffected animals haplotype. A dominant trait mutation from either the dam or sire produced minimal results, with dam filtering showing 163 variants, none of which were within a gene and therefore are highly unlikely to impact protein production and function. Sire filtering, however, produced 89 variants, five of which were high impact, meaning that they may likely be located within a stop/start codon or is an insertion/deletion (INDEL) causing a frameshift.

Table 7 – Number of variants (n) found for each filtering method by the type of mutation and subsequent impact.

Filtering method	Variants (n)	SNPs (n)	INDEL (n)	High impact (n)	Moderate impact (n)
Recessive	9,181	7,407	1,774	2	92
De novo	0	0	0	0	0
Sire	89	36	53	5	0
Dam	163	64	99	0	0

Note: Recessive filtering keeps only alleles where affected animals are homozygous and both parents carry a copy. De novo filters out haplotypes unique to only the affected animals. Sire/dam filtering finds dominant heterozygous haplotypes in the affected dogs as well as a parent based on filter rules (only sire or only dam).

Recessive trait filtering yielded the most results with 9,181 variants, the majority of which were SNPs, spread across the autosomes but primarily clustered between position 60,000,000 and 82,000,000 on chromosome 2 (Figure 10) and between position 34,000,000 and 42,000,000 on chromosome 13 (Figure 11).

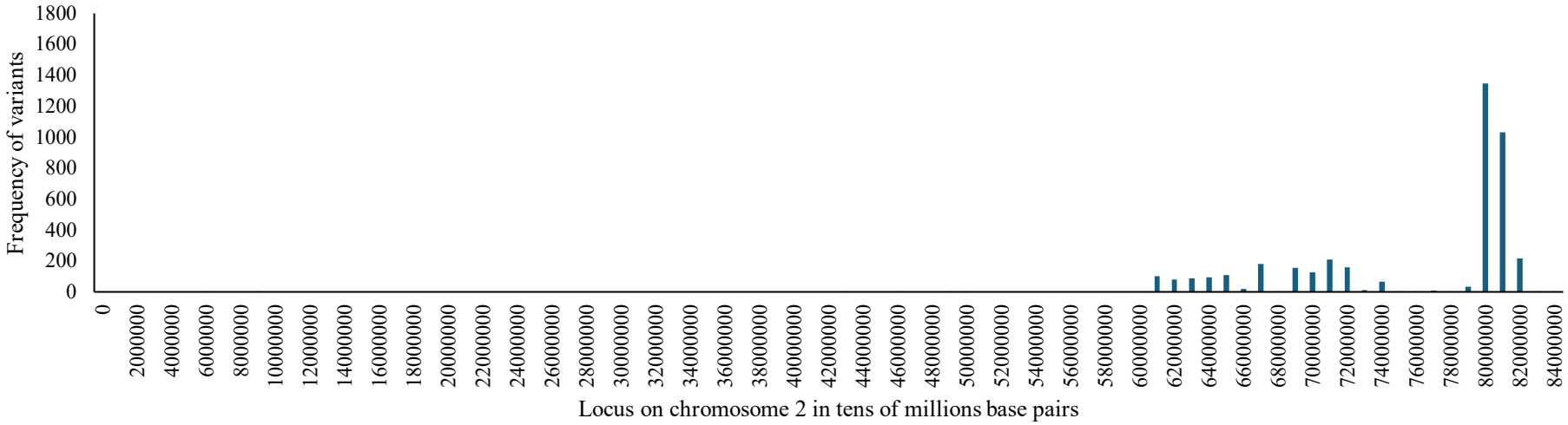

Figure 10 – Distribution of recessive variants across chromosome 2.

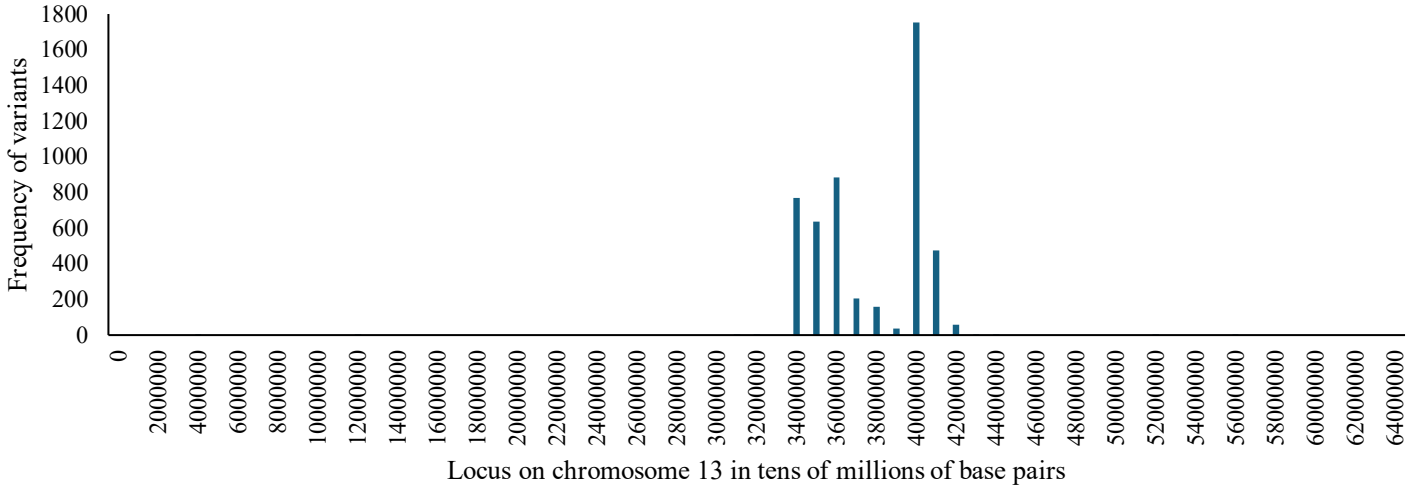

Figure 11 – Distribution of recessive variants across chromosome 13.

Of the variants identified, two were high impact and 92 were moderate impact, which are often due to missense mutations in an exon. The remaining variants were either low impact or within modifier regions which would be unlikely to cause any significant effect. The distribution of the types of variant effect as well as the regions in which the variants occur is described in Tables 8 and 9.

Table 8 – Number and proportion of variant effects produced through recessive variant filtering by variant type. Includes modifier variants.

Type of variant	Count (n)	Percent (%)
3 prime UTR variant	388	1.19
5 prime UTR premature start codon gain variant	5	0.02
5 prime UTR variant	91	0.28
Downstream gene variant	2,929	9.01
Intergenic region	4,003	12.31
Intragenic variant	5	0.02
Intron variant	21,475	66.06
Missense variant	92	0.28
Non-coding transcript exon variant	410	1.26
Splice acceptor variant	2	0.01
Splice donor variant	2	0.01
Splice region variant	45	0.14
Synonymous variant	198	0.61
Upstream gene variant	2,866	8.82

Table 9 – Number and proportion of variant effects produced through recessive variant filtering by region variant occurred in. Includes modifier variants.

Variant type by region	Count (n)	Percent (%)
Downstream	2,929	9.02
Exon	700	2.16
Intergenic	4,003	12.33
Intron	21,430	66.01
Splice site acceptor	2	0.01
Splice site region	45	0.14
Transcript	5	0.02
Upstream	2,866	8.83
UTR 3 prime	388	1.20
UTR 5 prime	96	0.30

Annotated variant call files were then manually examined in IGV (Robinson et al., 2011) to ensure that the mutations were within an exon of a gene and that any errors from the filtering and annotation process were removed. During this manual examination, the high-impact variants produced from the filter for dominant alleles from the sire were removed from further investigation as they were considered to be the result of poor annotation and not true nonsense variants. This left the recessive filtering as the only model remaining with variants of high impact for further investigation.

In addition to these filters, the X chromosome and non-nuclear DNA from the mitochondria were examined to ensure that there were no structural variants or inconsistencies that may not have been flagged. Manual examination in IGV revealed no areas of interest on either chromosome, hence they were disqualified as a potential source of the issue.

4.2.4 – Chromosomal Abnormalities

Variants found through recessive trait filtering are displayed by chromosome as well as the rate at which they appear (Table 10). Of the 38 canine autosomes, 26 contained variants, ranging from one to 4,988. Furthermore, chromosomes 2 and 13 had a large number of variants, with 4,054 and 4,988 respectively. Excluding these chromosomes, those remaining with variants averaged 5.5 per chromosome. Chromosomes where variants only appear at a very low frequency (all but 2 and 13), are likely a result of annotation errors during the filtering process or are located within introns and therefore would have no impact as they are not known to encode for gene products.

The number of variants across the genome were displayed on a logarithmic scale, due to the disproportionate distribution of variants (Figure 12). This clearly showed the disparity in frequency between chromosome 2 and 13 compared to the other autosomes.

Table 10 – Frequency and rate at which variants occur in each autosome and non-nuclear DNA.

Chromosome	Length (bp)	Variants (n)	Variants rate (bp/n)
1	123,556,469	4	30,889,117
2	84,979,418	4,054	20,961
3	92,479,059	18	5,137,725
4	89,535,178	2	44,767,589
5	89,562,946	7	12,794,706
6	78,113,029	3	26,037,676
7	81,081,596	1	81,081,596
8	76,405,709	1	76,405,709
10	70,643,054	2	35,321,527
11	74,805,798	12	6,233,816
13	64,299,765	4,988	12,890
14	61,112,200	4	15,278,050
15	64,676,183	1	64,676,183
16	60,362,399	1	60,362,399
17	65,088,165	2	32,544,082
19	55,516,201	1	55,516,201
21	51,742,555	17	3,043,679
22	61,573,679	10	6,157,367
23	53,134,997	3	17,711,665
25	51,730,745	6	8,621,790
27	46,662,488	24	1,944,270
30	40,643,782	1	40,643,782
31	39,901,454	1	39,901,454
34	42,397,973	2	21,198,986
35	28,051,305	7	4,007,329
38	24,803,098	3	8,267,699
Non-nuclear DNA	1,559,206	6	259,868
Total	**1,674,418,451**	**9,181**	**182,378**

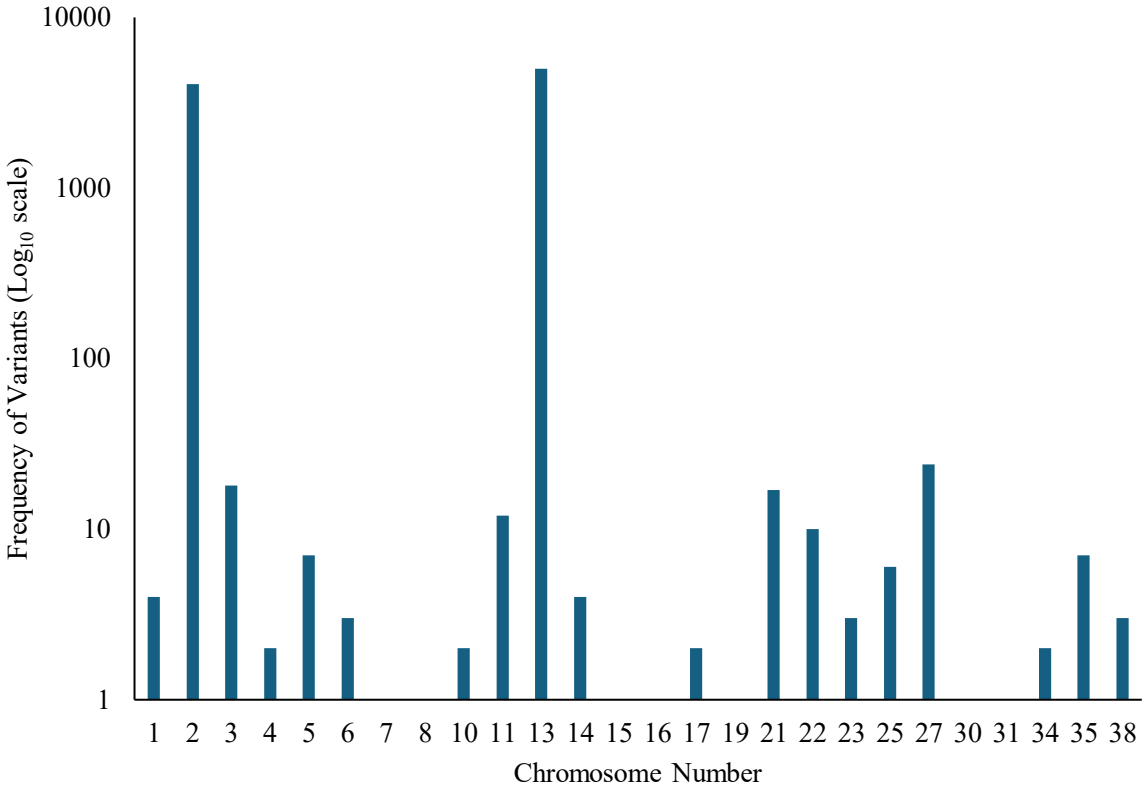

Figure 12 – Number of variants (log_{10}) produced through filtering for recessive traits where affected animals are homozygous for loci that parents are both heterozygous.

4.2.5 – Variants of Interest

As all other filtering methods did not produce any variants of interest, further investigation only proceeded with the variants found under recessive condition filtering in the affected animals. These variants were processed using SnpSift v5.2 (Cingolani, Patel, et al., 2012), to locate and isolate the variants of moderate and high impact, reducing the initial 9,181 recessive mutations to 94, spread across 30 different genes (see Appendix 1). One gene of particular interest was *COL22A1*, due to its related pathways being involved with the stabilisation of myotendinous junctions and strengthening skeletal muscle attachment during contractile activity. However, the variants within this gene were later excluded due to annotation errors and high frequency within a larger sample population, as described in section 4.2.4. Furthermore, given the variants presented, all of the potential diseases discussed in chapter 2.4 were excluded from further investigation since none of the relevant genes were present within the affected chromosomal regions.

Moderate impact variants were typically missense variants, where a change in the nucleotide sequence causes the related amino acid to be changed. High impact variants were usually an

insertion or deletion, causing a frameshift and therefore usually causing a loss of functional expression, with both reduced or absent protein function and potential reduced expression due to nonsense-mediated decay. Given the number of variants remaining after filtering, each variant was located and manually examined in IGV, to identify and exclude any poorly annotated variants. Synonymous and missense variants located in introns were removed since they have no impact on the related protein and will not cause a frameshift of the final gene product as they are removed during RNA splicing to make mRNA, leaving 56 variants located within exons. From these variants, the reference and alternative codon at each location was found alongside the subsequent amino acid. This allowed synonymous codons to be removed, as they usually cause no change in the encoded protein. While there was a chance that synonymous variants could cause significant impacts, they can be difficult to predict and therefore beyond the scope of this study.

4.2.6 – Population Filtering and RNA Coverage

Of the remaining 56 variants, two final filters were applied: an RNA coverage map of 17 tissue types (including: adipose tissue, adrenal gland, bone marrow, cerebellum, the cortex and medulla of the kidney, liver, lung, lymph node, occipital cortex, pancreas, right ventricle of the heart, skeletal muscle, skin, small intestine, spleen and stomach), and a publicly available catalogue of moderately and higher frequency variants from the whole genome sequence 722 canids of various breeds, retrieved from Plassais et al. (2019). Adequate RNA coverage was used to ensure that each variant was located in a translated area and therefore would impact the

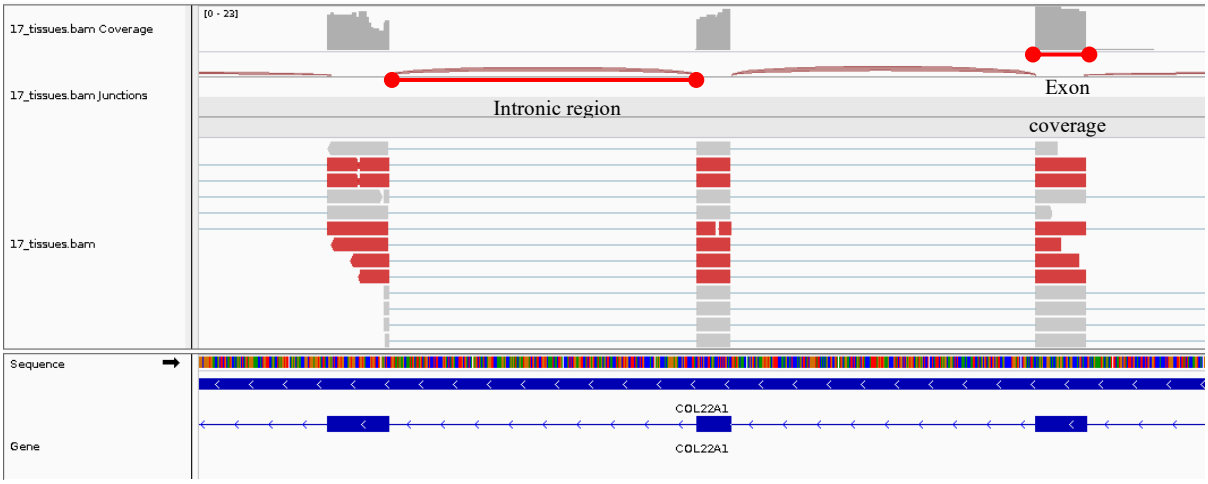

Figure 13 – Integrative genomics viewer (IGV) screenshot of RNA coverage aligned to the CanFam4 reference genome. Includes areas of coverage within the exon (red line, right) and areas without RNA coverage (intron, red line, left). Areas with no RNA coverage are not expressed within the tissue.

resulting protein sequence. Examples of good and no coverage are displayed in Figure 13. Comparison with the breed catalogue allowed variants with a mean allele frequency of 10% or greater within the wider population to be removed. These variants were less likely to cause disease and did not match the unique presentation of the 'W' litter. Following these final filters, only four variants remain (Table 11).

These variants were spread across genes with different functions, including: *BOP1*, a regulator gene of ribosome biogenesis and cell proliferation; *IQANK1*, which organises a cell's cytoplasmic membranes and cytoskeleton; *THEMIS2*, an immune cell regulator; and *ZNF517*, a zinc finger protein within a gene family involved in transcriptional gene regulation.

Table 11 – Genes remaining with variants present within an exon after all filtering steps, including their location and impact on the gene and change to amino acids.

Gene ID	Gene Name	Chr	Position	Variant Type	Impact Type	Reference	Alternative	Reference Amino Acid	Alternative Amino Acid
BOP1	BOP1 Ribosomal Biogenesis Factor	13	38,452,284	Missense	Moderate	C	T	Arg	Gln
IQANK1	IQ Motif and Ankyrin Repeat Containing 1	13	38,099,655	Missense	Moderate	A	G	Arg	Gly
THEMIS2	Thymocyte Selection Associated Family Member 2	2	71,859,696	Missense	Moderate	G	A	Thr	Met
ZNF517	Zinc Finger Protein 517	13	38,839,994	Missense	Moderate	C	T	Arg	Gln

Chapter 5 – Discussion

5.1 – Clinical Findings

The blood tests undertaken in this study provided important evidence to determine which animals were affected compared to those that were not, while the histological samples helped to describe the structural and functional impact on the muscle tissue of the affected dogs. However, these results, alongside the clinical presentation of temporary exercise induced intolerance did not match any previously described disease. As a result, the dogs DNA was explored.

5.1.1 – Mitochondrial Myopathies

The main disease of concern based on the clinical presentation of the affected animals was a mitochondrial myopathy or other mitochondrial disease from a veterinary perspective. Primary mitochondrial myopathies were postulated as it mainly affects the skeletal muscle through defects in the oxidative phosphorylation process (Arena et al., 2022). They are accompanied by exercise intolerance and muscle weakness, as well as progressive external ophthalmoplegia (weakness or paralysis of the eye muscles) and eyelid ptosis (droopy eyelids) (Arena et al., 2022), neither of which were seen in any of these animals. Diagnostic criteria from histopathological evidence also includes decreased cytochrome oxidase and succinate dehydrogenase staining, and the presence of ragged red fibres in the muscle biopsy (Paciello et al., 2003). In the affected dogs, elevated CK in blood serum and misshapen and excessive numbers of mitochondria under electron microscopy align with a possible primary mitochondrial myopathy diagnosis (Paciello et al., 2003). Based on the clinical symptoms alone, veterinarians were unable to diagnose the affected animals with a mitochondrial myopathy, especially considering that often, primary mitochondrial myopathies can only be confirmed through post-mortem examination or invasive testing techniques (Gomes, 2021).

From the following genetic investigation, no evidence was found that there were any variants in the mitochondrial DNA sequence. However, there was a possibility that a mitochondrial issue was caused by defects within the nuclear DNA. These defects are difficult to determine, and little literature exists on the topic, especially within dogs. Given the unclear but possible structural issues seen on chromosome 2 and 13, there is insufficient evidence to confirm or

exclude the possibility of mitochondrial disease, however, it is unlikely that, if present, that it would be exclusively maternally inherited.

5.2 – Chromosomal Issues

The non-nuclear DNA from the sampled dogs revealed no variants or areas of interest. As the majority of mitochondrial DNA is maternally inherited (Gomes, 2021; Paciello et al., 2003), an exclusively maternally inherited disease can be ruled out as a cause of the symptoms seen within the 'W' litter. Consequently, it is probable that this disease presentation has contributions from both the dam and the sire, and both of the animals should be investigated further before being bred from again. Alternatively, if the disease was due to variants within the mitochondrial DNA, the sire would be cleared to be bred from again since he would have minimal contribution to the issue. Similarly, the X chromosome contained no areas of interest, so the likelihood of a sex-linked disease was excluded.

While no clear evidence of any distinct structural issues was revealed through genome filtering and manual examination in IGV (Figures 11 and 12), there is some indication that there is some structural inconsistencies in chromosomes 2 and 13. The significant peaks of recessive variants following filtering is distributed in distinct bands, which is a potential marker of a deleterious structural anomaly (Abel & Duncavage, 2013). Further investigation of this variant distribution to determine whether or not it is a causal variant is beyond the scope of this study, and no determination could be made.

5.3 – Candidate Variants

Several different filtering methodologies were applied to the whole genome sequences of the nine dogs to assess the possible mode of inheritance through process of elimination. This revealed four candidate genes (Table 11) that met the filtering criteria of consistency in mode of inheritance across the affected animals, RNA coverage and a ≤10% allele frequency within a large population filter. Each of these candidate genes are described further within the following sections.

5.3.1 – BOP1

Located on chromosome 13, genome analysis showed a missense mutation at locus 38,452,284 where the nucleotide cytosine was replaced by thymine. Consequently, this altered the related amino acid from arginine to glutamine, thereby removing the charged side chain present on

arginine and disrupting the proteins' structure (Sotomayor-Vivas et al., 2022). This change to the protein structure of the protein may disrupt at least one of the functions that this gene is associated with. In the case of metabolic regulation, it is possible that this change could impact on exercise tolerance in these animals, however, metabolomic work would be required to determine the impact on the function of these genes.

5.3.2 – IQANK1

A missense mutation was found within *IQANK1* at locus 38,099,655 on chromosome 13 where a nucleotide was changed from adenine to guanine. The resulting codon led to change in the amino acid from arginine, a charged and basic molecule, to glycine, which is polar and uncharged, therefore potentially creating a structural change in the protein (Sotomayor-Vivas et al., 2022). It should also be noted that while the gene had some RNA coverage, it was inconsistent and less reliable compared to coverage seen over other genes.

5.3.3 – THEMIS2

The missense mutation within *THEMIS2* was located at 71,859,696 on chromosome 2, where a guanine was changed to adenine, causing the affected amino acid to also change from the polar threonine to the non-polar methionine, thereby altering the proteins resulting structure (Sotomayor-Vivas et al., 2022).

5.3.4 – ZNF517

The final candidate variant identified within this study is a missense mutation on chromosome 13 at the locus 38,839,994. These variant changes the cytosine to a to thymine with the resulting amino acid altered from the charged and basic arginine to the uncharged, polar glutamine, potentially leading to a structural change in the protein (Sotomayor-Vivas et al., 2022).

5.4 – Implications of Inbreeding for Animals

Inbreeding poses a serious risk to the health and longevity of an animal due to the accumulation of deleterious alleles. It is virtually impossible to avoid where 'purebred' animals are concerned since breeders are choosing from a limited gene pool to avoid outbreeding and is, further worsened by severe founder effect at the breed conception. The founder effect can be seen in the GSD, due to the original rigid breed standard created by von Stephanitz (Strickland, 1998; Tenner, 2017). There was also a significant impact of champion dogs on the subsequent generations of puppies. Because of the limited effective population size, many individuals can

be traced back to some of the founding animals of the breed (Strickland, 1998). There is an inevitable baseline level of inbreeding present in GSD's which has not been accounted for in this study (Goldbecker & Hart, 1984; Strickland, 1998; Wootton, 1988). The impact of this baseline of inbreeding can be seen with the number of hereditary conditions seen within the breed (Wahl et al., 2008) (see chapter 2.4).

Breeders therefore need to maintain a level of awareness for the level of inbreeding within their animals when choosing pairings. Further inbreeding can have significant impacts, such as can be seen within the 'W' litter. This is further worsened by the fact that they were intended as working animals and need to maintain a high level of health to ensure that the time and financial investment in these animals is worthwhile.

5.5 – Limitations on Study

While this study identified several candidate variants (chapter 5.3) currently none show a clear link seen with the phenotype of the affected animals. Additionally, there were several synonymous variants as well the stretches of inconsistency on chromosomes 2 and 13 which were not able to be investigated further within the scope of this study, either of which may provide stronger evidence for a causal mutation. Regardless of not having a clearly evidenced causal variant, from the phenotypic presentation within the dogs of the 'W' litter alongside the genetic data, it is likely that a recessive hereditary condition is the cause of the symptoms in these animals. Without further investigation to determine the source of the genetic defect, all nine animals should be removed from any further breeding programs, to remove the chance of having these issue occur in other animals. Sequencing other related animals that are still alive may also provide more context for the mode of inheritance.

The major limitation of this study was the lack of post-mortem on the deceased affected dog. Examination of this animal may have provided significant information such as a muscle biopsy to compare if there was an increase in muscle decay or if there was damage to other organs and body systems. Furthermore, some complex diseases, particularly involving the nervous system, such as myopathies require diagnosis post-mortem due to the highly invasive nature of the investigation.

Chapter 6 – Conclusion

The aim of this thesis was to investigate a potential genetic cause of a low exercise tolerance phenotype present in a litter of GSD puppies intended for police dog training. They showed extreme exercise intolerance at the beginning of training, and clinical investigation provided inconclusive results. All nine dogs were whole genome sequenced (seven pups, dam and sire).

Initial variant filtering was conducted based on haplotypes for: a dominant condition passed on from either parent; a recessive condition from both parents; a sex-linked condition; non-nuclear inheritance and a de novo mutation. Variants were screened to find those with an impact on the resulting protein. The second screening process eliminated all inheritance patterns except for a recessive condition, where both parents contributed to the phenotype seen. The recessive variants, were manually audited, providing a set of four final candidate genes, *BOP1*, *IQANK1*, *THEMIS2*, and *ZNF517*. None of the gene functions would be expected to correlate to the phenotype of the affected animals. Given the inconclusive genetic investigation, further research could be conducted investigating the potential structure issues in chromosome 2 and 13, as well as a potentially causal synonymous variant. Additionally, conclusive diagnosis can only typically found post-mortem, therefore if any of the other affected animals die, this could provide significant evidence to the presence of mitochondrial disease.

Given the inconclusive outcome of this investigation, all of the animals involved in the study should be removed from any further breeding due to the severe negative effects of the affected phenotype. This is on grounds of both animal welfare as well as regarding the time and financial investment put into these animals to become working police dogs, which only one of the seven puppies went on to become. Likewise, extended pedigrees should be consulted when choosing pairings to breed to prevent further inbreeding that may lead to similar deleterious effects.

Chapter 7 – References

Abel, H. J., & Duncavage, E. J. (2013). Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches. Cancer Genetics, 206(12), 432–440. https://doi.org/10.1016/j.cancergen.2013.11.002

Ahn, C. S., Cho, H. K., Lee, D., Sim, H., Kim, S., & Pai, H. (2016). Functional characterization of the ribosome biogenesis factors PES, BOP1, and WDR12 (PeBoW), and mechanisms of defective cell growth and proliferation caused by PeBoW deficiency in Arabidopsis. Journal of Experimental Botany, 67(17), 5217–5232. https://doi.org/10.1093/jxb/erw288

Arena, I. G., Pugliese, A., Volta, S., Toscano, A., & Musumeci, O. (2022). Molecular Genetics Overview of Primary Mitochondrial Myopathies. Journal of Clinical Medicine, 11(3), 632. https://doi.org/10.3390/jcm11030632

Aslan, Ö., Akyüz, B., Arslan, K., Keleş, İ., Ulusan, M., İlgar, E. G., & Akçay, A. (2016). The Relationship Between von Willebrand Factor Gene and von Willebrand Factor Antigen Levels in Dogs. Kafkas Universitesi Veteriner Fakultesi Dergisi, 22(4), 585–590. https://doi.org/10.9775/kvfd.2016.15106

Aslanian, M. E., Sharp, C. R., Rozanski, E. A., De Laforcade, A. M., Rishniw, M., & Brooks, M. B. (2014). Clinical outcome after diagnosis of hemophilia A in dogs. Journal of the American Veterinary Medical Association, 245(6), 677–683. https://doi.org/10.2460/javma.245.6.677

Awano, T., Johnson, G. S., Wade, C. M., Katz, M. L., Johnson, G. C., Taylor, J. F., Perloski, M., Biagi, T., Baranowska, I., Long, S., March, P. A., Olby, N. J., Shelton, G. D., Khan, S., O'Brien, D. P., Lindblad-Toh, K., & Coates, J. R. (2009). Genome-wide association analysis reveals a SOD1 mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. Proceedings of the National Academy of Sciences, 106(8), 2794–2799. https://doi.org/10.1073/pnas.0812297106

Baker, L., Muir, P., & Sample, S. J. (2019). Genome-wide association studies and genetic testing: understanding the science, success, and future of a rapidly developing field. Journal of the American Veterinary Medical Association, 255(10), 1126–1136. https://doi.org/10.2460/javma.255.10.1126

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., & Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. Bioinformatics, 27(12), 1691–1692. https://doi.org/10.1093/bioinformatics/btr174

Bell, S. M., Evans, J. M., Evans, K. M., Tsai, K. L., Noorai, R. E., Famula, T. R., Holle, D. M., & Clark, L. A. (2022). Congenital idiopathic megaesophagus in the German shepherd dog is a sex-differentiated trait and is associated with an intronic variable number tandem repeat in Melanin-Concentrating Hormone Receptor 2. PLoS Genetics, 18(3), e1010044. https://doi.org/10.1371/journal.pgen.1010044

Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. Nucleic Acids Research, 41(D1), D36–D42. https://doi.org/10.1093/nar/gks1195

Berg, J. A., Sævik, B. K., Trangerud, C., Madsen, P., & Lingaas, F. (2025). Genetic analyses of lumbosacral transitional vertebra and hip dysplasia in nine dog breeds in Norway. Acta Veterinaria Scandinavica, 67(1). https://doi.org/10.1186/s13028-025-00810-z

Blancher, A. (2013). Evolution of the ABO supergene family. ISBT Science Series, 8(1), 201–206. https://doi.org/10.1111/voxs.12044

Bleackley, J. May 12, 2025. Personal communication, regarding police dog training schedule.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borah, N., Nukala, K. M., Vadlapudi, V., Gubbala, S. P., & Amanchy, R. (2019). BOP1– a key player of ribosomal biogenesis. Current Science, 117(3), 422. https://doi.org/10.18520/cs/v117/i3/422-433

Breen, M., Switonski, M., & Binns, M. M. (2001). Cytogenetics and Physical Chromosome Maps. In The Genetics of the Dog (pp. 299–328). CAB International.

Brooks, M. B., & Catalfamo, J. L. (2022). Chapter 83: von Willebrand Disease. In Schalm's Veterinary Hematology (7th ed., pp. 731–738). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119500537.ch83

Brooks, M., & Sargan, D. R. (2001). Genetic Aspects of Disease in Dogs. In The Genetics of the Dog (pp. 191–266). CAB International.

Capucchio, M. T., Spalenza, V., Biasibetti, E., Bottero, M. T., Rasero, R., Dalmasso, A., & Sacchi, P. (2014). Degenerative myelopathy in German Shepherd Dog: comparison of two molecular assays for the identification of the SOD1:c.118G>A mutation. Molecular Biology Reports, 41(2), 665–670. https://doi.org/10.1007/s11033-013-2904-9

Chavanas, S., Méchin, M., Takahara, H., Kawada, A., Nachat, R., Serre, G., & Simon, M. (2004). Comparative analysis of the mouse and human peptidylarginine deiminase gene clusters reveals highly conserved non-coding segments and a new human gene, PADI6. Gene, 330, 19–27. https://doi.org/10.1016/j.gene.2003.12.038

Cheng, D., Deobagkar-Lele, M., Zvezdova, E., Choi, S., Uehara, S., Baup, D., Bennett, S. C., Bull, K. R., Crockford, T. L., Ferry, H., Warzecha, C., Marcellin, M., De Peredo, A. G., Lesourne, R., Anzilotti, C., Love, P. E., & Cornall, R. J. (2016). Themis2 lowers the threshold for B cell activation during positive selection. Nature Immunology, 18(2), 205–213. https://doi.org/10.1038/ni.3642

Chung, K., Cheng, I. K., Ching, A. K., Chu, J., Lai, P. B., & Wong, N. (2011). Block of proliferation 1 (BOP1) plays an oncogenic role in hepatocellular carcinoma by promoting epithelial-to-mesenchymal transition. Hepatology, 54(1), 307–318. https://doi.org/10.1002/hep.24372

Cingolani, P., Patel, V. M., Coon, M., Nguyen, T., Land, S. J., Ruden, D. M., & Lu, X. (2012). Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Frontiers in Genetics, 3. https://doi.org/10.3389/fgene.2012.00035

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly, 6(2), 80–92. https://doi.org/10.4161/fly.19695

Coates, J. R., & Wininger, F. A. (2010). Canine degenerative myelopathy. Veterinary Clinics of North America Small Animal Practice, 40(5), 929–950. https://doi.org/10.1016/j.cvsm.2010.05.001

Conrado, A. L. V., Iunes, R. S., Balduíno, A. L. L., Santanna, M. C. F. B., & Da Silva, J. R. M. C. (2020). Serum symmetric dimethylarginine levels in a half-breed German shepherd dog with renal cystadenocarcinoma and nodular dermatofibrosis. Comparative Clinical Pathology, 29(4), 905–909. https://doi.org/10.1007/s00580-020-03135-7

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., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. Bioinformatics, 27(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience, 10(2). https://doi.org/10.1093/gigascience/giab008

Deng, Z., Cangkrama, M., Butt, T., Jane, S. M., & Carpinelli, M. R. (2021). Grainyhead-like transcription factors: guardians of the skin barrier. Veterinary Dermatology, 32(6), 553. https://doi.org/10.1111/vde.12956

Dillard, K. J., Ochs, M., Niskanen, J. E., Arumilli, M., Donner, J., Kyöstilä, K., Hytönen, M. K., Anttila, M., & Lohi, H. (2020). Recessive missense LAMP3 variant associated with defect in lamellar body biogenesis and fatal neonatal interstitial lung disease in dogs. PLoS Genetics, 16(3), e1008651. https://doi.org/10.1371/journal.pgen.1008651

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Dronova, T. A., Babyshkina, N. N., Kostromitsky, D. N., Eremin, D. A., & Cherdyntseva, N. V. (2024). Transcriptome of metastatic colorectal cancer. Cardiometry, 33, 33–35. https://doi.org/10.18137/cardiometry.2024.33.conf.12

Edmands, S. (2006). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Molecular Ecology, 16(3), 463–475. https://doi.org/10.1111/j.1365-294x.2006.03148.x

Erceg, V. February 25, 2025. Personal communication, regarding police dog breeding programme.

Evans, R. I., Herbold, J. R., Bradshaw, B. S., & Moore, G. E. (2007). Causes for discharge of military working dogs from service: 268 cases (2000–2004). Journal of the American Veterinary Medical Association, 231(8), 1215–1220. https://doi.org/10.2460/javma.231.8.1215

Forman, O. P., Penderis, J., Hartley, C., Hayward, L. J., Ricketts, S. L., & Mellersh, C. S. (2012). Parallel Mapping and Simultaneous Sequencing Reveals Deletions in BCAN and FAM83H Associated with Discrete Inherited Disorders in a Domestic Dog Breed. PLoS Genetics, 8(1), e1002462. https://doi.org/10.1371/journal.pgen.1002462

Genome assembly (UU_Cfam_GSD_1.0): Canis Lupus Familiaris. (2010, March). National Center for Biotechnology Information. Retrieved November 3, 2024, from https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_011100685.1/

Gherman, L., Isachesku, E., Zanoaga, O., Braicu, C., & Berindan-Neagoe, I. (2024). Molecular markers in canine mammary tumors. Acta Veterinaria, 74(2), 159–182. https://doi.org/10.2478/acve-2024-0012

Gluding, D., Stock, K. F., Tellhelm, B., Kramer, M., & Eley, N. (2021). Genetic background of lumbosacral transitional vertebrae in German shepherd dogs. Journal of Small Animal Practice, 62(11), 967–972. https://doi.org/10.1111/jsap.13380

Goldbecker, W., & Hart, E. H. (1984). The German shepherd dog. TFH Publications.

Gomes, S. (2021). A review of mitochondrial disease in dogs. Companion Animal, 26(11), 257–264.

Gray, R. S., Abitua, P. B., Wlodarczyk, B. J., Szabo-Rogers, H. L., Blanchard, O., Lee, I., Weiss, G. S., Liu, K. J., Marcotte, E. M., Wallingford, J. B., & Finnell, R. H. (2009). The planar cell polarity effector Fuz is essential for targeted membrane trafficking, ciliogenesis and mouse embryonic development. Nature Cell Biology, 11(10), 1225–1232. https://doi.org/10.1038/ncb1966

Groot, K. R., Sevilla, L. M., Nishi, K., DiColandrea, T., & Watt, F. M. (2004). Kazrin, a novel periplakin-interacting protein associated with desmosomes and the keratinocyte plasma membrane. The Journal of Cell Biology, 166(5), 653–659. https://doi.org/10.1083/jcb.200312123

Gunster, M. J., Satijn, D. P., Hamer, K. M., Blaauwen, J. L. D., De Bruijn, D., Alkema, M. J., Van Lohuizen, M., Van Driel, R., & Otte, A. P. (1997). Identification and Characterization of Interactions between the Vertebrate Polycomb-Group Protein BMI1 and Human Homologs of Polyhomeotic. Molecular and Cellular Biology, 17(4), 2326–2335. https://doi.org/10.1128/mcb.17.4.2326

Hamann, H., Kirchhoff, T., & Distl, O. (2003). Bayesian analysis of heritability of canine hip dysplasia in German Shepherd Dogs. Journal of Animal Breeding and Genetics, 120(4), 258–268. https://doi.org/10.1046/j.1439-0388.2003.00395.x

History of the police dog section. (n.d.). New Zealand Police. Retrieved September 12, 2024, from https://www.police.govt.nz/about-us/structure/teams-units/dog-section/history

Holder, A. L., Price, J. A., Adams, J. P., Volk, H. A., & Catchpole, B. (2014). A retrospective study of the prevalence of the canine degenerative myelopathy associated superoxide dismutase 1 mutation (SOD1:c.118G > A) in a referral population of German Shepherd dogs from the UK. Canine Genetics and Epidemiology, 1(1), 10. https://doi.org/10.1186/2052-6687-1-10

Huang, X., Miyata, H., Wang, H., Mori, G., Iida-Norita, R., Ikawa, M., Percudani, R., & Chung, J. (2023). A CUG-initiated CATSPERθ functions in the CatSper channel assembly and serves as a checkpoint for flagellar trafficking. Proceedings of the National Academy of Sciences, 120(39). https://doi.org/10.1073/pnas.2304409120

Hunt, H. (2018). Epidemiological, pathological and metabolomic characterisation of an acquired myopathy of dogs in New Zealand : a thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Veterinary Science at Massey University, Manawatū, New Zealand [PhD Dissertation, Massey University]. https://mro.massey.ac.nz/handle/10179/14413

Iida, A., & Nakamura, Y. (2004). Identification of 45 novel SNPs in the 83-kb region containing peptidylarginine deiminase types 1 and 3 loci on chromosomal band 1p36.13. Journal of Human Genetics, 49(7), 387–390. https://doi.org/10.1007/s10038-004-0156-1

Jagannathan, V., Drögemüller, C., & Leeb, T. (2019). A comprehensive biomedical variant catalogue based on whole genome sequences of 582 dogs and eight wolves. Animal Genetics, 50(6), 695–704. https://doi.org/10.1111/age.12834

Jing, W., Yabas, M., Bröer, A., Coupland, L., Gardiner, E. E., Enders, A., & Bröer, S. (2019). Calpain cleaves phospholipid flippase ATP8A1 during apoptosis in platelets. Blood Advances, 3(3), 219–229. https://doi.org/10.1182/bloodadvances.2018023473

Johnson, B. M., Denovo, R. C., & Mears, E. A. (2009). Canine Megaesophagus. In Kirk's Current Veterinary Therapy (14th ed., pp. 486–492). Saunders Elsevier.

https://www.currentveterinarytherapy.com/content/webchapters/Bonagura_Web%20Chapter%2047_main.pdf

Karsada, R., Mészáros, G., Kadleík, O., & Blueca, JJ. (2007). Impact of inbreeding and heritability of Canine Hip Dysplasia in German Shepherd population. Folia Veterinaria, 51, 142–145. https://www.researchgate.net/profile/Radovan-Kasarda/publication/259214454_Evaluation_of_dysplasia_at_German_shepherd_population/links/00b7d52a7122539139000000/Evaluation-of-dysplasia-at-German-shepherd-population.pdf

Kassymbekova, S., Bimenova, Z., Iskhan, K., Sobiech, P., Jastrzebski, J., Brym, P., Babis, W., Kalykova, A., Otebayev, Z., Kabylbekova, D., Baneh, H., & Romanov, M. (2025). Uncovering genetic diversity and adaptive candidate genes in the Mugalzhar horse breed using Whole-Genome sequencing data. Animals, 15(18), 2667. https://doi.org/10.3390/ani15182667

Kemp, M. G., Mason, A. C., Carreira, A., Reardon, J. T., Haring, S. J., Borgstahl, G. E. O., Kowalczykowski, S. C., Sancar, A., & Wold, M. S. (2009). An alternative form of replication protein A expressed in normal human tissues supports DNA repair. Journal of Biological Chemistry, 285(7), 4788–4797. https://doi.org/10.1074/jbc.m109.079418

Kim, Y., Bang, H., & Kim, D. (2000). TASK-3, a new member of the tandem Pore K+ Channel family. Journal of Biological Chemistry, 275(13), 9340–9347. https://doi.org/10.1074/jbc.275.13.9340

Kitzmann, S., Hartmann, K., Zablotski, Y., Rieger, A., Mueller, R., & Wehner, A. (2021). Wellbeing, quality of life, presence of concurrent diseases, and survival times in untreated and treated German Shepherd dogs with dwarfism. PLoS ONE, 16(8), e0255678. https://doi.org/10.1371/journal.pone.0255678

Koch, M., Schulze, J., Hansen, U., Ashwodt, T., Keene, D. R., Brunken, W. J., Burgeson, R. E., Bruckner, P., & Bruckner-Tuderman, L. (2004). A Novel Marker of Tissue Junctions, Collagen XXII. Journal of Biological Chemistry, 279(21), 22514–22521. https://doi.org/10.1074/jbc.m400536200

Korneenko, T. V., Pestov, N. B., Okkelman, I. A., Modyanov, N. N., & Shakhparonov, M. I. (2015). P4-ATPase Atp8b1/FIC1: Structural features and physiological functions in

health and disease. Russian Journal of Bioorganic Chemistry, 41(1), 1–9. https://doi.org/10.1134/s1068162015010070

Kristensen, T. N., & Sørensen, A. C. (2005). Inbreeding – lessons from animal breeding, evolutionary biology and conservation genetics. Animal Science, 80(2), 121–133. https://doi.org/10.1079/asc41960121

Kyono, M. (2002). The New Zealand Police Dogs: A thesis presented in partial fulfilment of the requirements for the degree of Master of Veterinary Studies in Animal Science at Massey University [MA Thesis, Massey University]. https://mro.massey.ac.nz/bitstream/10179/10245/1/01_front.pdf

Langousis, G., Cavadini, S., Boegholm, N., Lorentzen, E., Kempf, G., & Matthias, P. (2022). Structure of the ciliogenesis-associated CPLANE complex. Science Advances, 8(15). https://doi.org/10.1126/sciadv.abn0832

Leeb, T., Bannasch, D., & Schoenebeck, J. J. (2022). Identification of Genetic Risk Factors for Monogenic and Complex Canine Diseases. Annual Review of Animal Biosciences, 11, 183–205. https://doi.org/10.1146/annurev-animal-050622-055534

Letko, A., Hédan, B., Snell, A., Harris, A. C., Jagannathan, V., Andersson, G., Holst, B. S., Ostrander, E. A., Quignon, P., André, C., & Leeb, T. (2023). Genomic diversity and runs of homozygosity in Bernese mountain dogs. Genes, 14(3), 650. https://doi.org/10.3390/genes14030650

Levin, M. G., Tsao, N. L., Singhal, P., Liu, C., Vy, H. M. T., Paranjpe, I., Backman, J. D., Bellomo, T. R., Bone, W. P., Biddinger, K. J., Hui, Q., Dikilitas, O., Satterfield, B. A., Yang, Y., Morley, M. P., Bradford, Y., Burke, M., Reza, N., Charest, B., . . . Damrauer, S. M. (2022). Genome-wide association and multi-trait analyses characterize the common genetic architecture of heart failure. Nature Communications, 13(1). https://doi.org/10.1038/s41467-022-34216-6

Li, L., Liu, X., Yang, S., Li, M., Wu, Y., Hu, S., Wang, W., Jiang, A., Zhang, Q., Zhang, J., Ma, X., Hu, J., Zhao, Q., Liu, Y., Li, D., Hu, J., Yang, C., Feng, W., & Wang, X. (2024). The HEAT repeat protein HPO-27 is a lysosome fission factor. Nature, 628(8008), 630–638. https://doi.org/10.1038/s41586-024-07249-8

Li, Y., Wang, X., Qi, S., Gao, L., Huang, G., Ren, Z., Li, K., Peng, Y., Yi, G., Guo, J., Yang, R., Wang, H., Zhang, X., & Liu, Y. (2021). Spliceosome-regulated RSRP1-dependent NF-κB activation promotes the glioblastoma mesenchymal phenotype. Neuro-Oncology, 23(10), 1693–1708. https://doi.org/10.1093/neuonc/noab126

Lingaas, F., Comstock, K. E., Kirkness, E. F., Sørensen, A., Aarskaug, T., Hitte, C., Nickerson, M. L., Moe, L., Schmidt, L. S., Thomas, R., Breen, M., Galibert, F., Zbar, B., & Ostrander, E. A. (2003). A mutation in the canine BHD gene is associated with hereditary multifocal renal cystadenocarcinoma and nodular dermatofibrosis in the German Shepherd dog. Human Molecular Genetics, 12(23), 3043–3053. https://doi.org/10.1093/hmg/ddg336

Liu, C., Chen, Z., Zhang, Z., Wang, Z., Guo, X., Pan, Y., & Wang, Q. (2024). Unveiling the Genetic Mechanism of Meat Color in Pigs through GWAS, Multi-Tissue, and Single-Cell Transcriptome Signatures Exploration. International Journal of Molecular Sciences, 25(7), 3682. https://doi.org/10.3390/ijms25073682

Lorenz, P., Dietmann, S., Wilhelm, T., Koczan, D., Autran, S., Gad, S., Wen, G., Ding, G., Li, Y., Rousseau-Merck, M., & Thiesen, H. (2010). The ancient mammalian KRAB zinc finger gene cluster on human chromosome 8q24.3 illustrates principles of C2H2 zinc finger evolution associated with unique expression profiles in human tissues. BMC Genomics, 11(1). https://doi.org/10.1186/1471-2164-11-206

Lynch, M., & Walsh, B. (1998). Genetics and analysis of quantitative traits. Sinauer Associates Incorporated.

Mace, S., Shelton, G., & Eddlestone, S. (Eds.). (2012). Megaesophagus. vetlearn.com.

Mahdieh, N., & Rabbani, B. (2013). An overview of mutation detection methods in genetic disorders. Iranian Journal of Pediatrics, 23(4), 375–388. https://pubmed.ncbi.nlm.nih.gov/24427490

Mair, B., Tomic, J., Masud, S. N., Tonge, P., Weiss, A., Usaj, M., Tong, A. H. Y., Kwan, J. J., Brown, K. R., Titus, E., Atkins, M., Chan, K. S., Munsie, L., Habsid, A., Han, H., Kennedy, M., Cohen, B., Keller, G., & Moffat, J. (2019). Essential gene profiles for human pluripotent stem cells identify uncharacterized genes and substrate dependencies. Cell Reports, 27(2), 599-615.e12. https://doi.org/10.1016/j.celrep.2019.02.041

Malbouyres, M., Guiraud, A., Lefrançois, C., Salamito, M., Nauroy, P., Bernard, L., Sohm, F., Allard, B., & Ruggiero, F. (2022). Lack of the myotendinous junction marker col22a1 results in posture and locomotion disabilities in zebrafish. Matrix Biology, 109, 1–18. https://doi.org/10.1016/j.matbio.2022.03.002

Marangi, G., Leuzzi, V., Manti, F., Lattante, S., Orteschi, D., Pecile, V., Neri, G., & Zollino, M. (2012). TRAPPC9-related autosomal recessive intellectual disability: report of a new mutation and clinical phenotype. European Journal of Human Genetics, 21(2), 229–232. https://doi.org/10.1038/ejhg.2012.79

McKee, M. (2010). Growth deformities of the long bones in dogs. In Practice, 32(7), 282–291. https://doi.org/10.1136/inp.c3914

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl variant effect Predictor. Genome Biology, 17(1). https://doi.org/10.1186/s13059-016-0974-4

Micale, L., Cialfi, S., Fusco, C., Cinque, L., Castellana, S., Biagini, T., Talora, C., Notarangelo, A., Bisceglia, L., Taruscio, D., Salvatore, M., & Castori, M. (2020). Novel TONSL variants cause SPONASTRIME dysplasia and associate with spontaneous chromosome breaks, defective cell proliferation and apoptosis. Human Molecular Genetics, 29(18), 3122–3131. https://doi.org/10.1093/hmg/ddaa195

Moe, L., & Lium, B. (1997). Hereditary multifocal renal cystadeno-carcinomas and nodular dermatofibrosis in 51 German shepherd dogs. Journal of Small Animal Practice, 38(11), 498–505. https://doi.org/10.1111/j.1748-5827.1997.tb03306.x

Moeller, E. M., Steiner, J. M., Clark, L. A., Murphy, K. E., Famula, T. R., Williams, D. A., Stankovics, M. E., & Vose, A. S. (2002). Inheritance of pancreatic acinar atrophy in German Shepherd Dogs. American Journal of Veterinary Research, 63(10), 1429–1434. https://doi.org/10.2460/ajvr.2002.63.1429

Momen, M., & Muir, P. (2025). Polygenic risk score prediction of complex diseases in companion animals: prospects, opportunities, and challenges. American Journal of Veterinary Research, 86(5), 1–8. https://doi.org/10.2460/ajvr.25.01.0018

Moore, G. E., Burkman, K. D., Carter, M. N., & Peterson, M. R. (2001). Causes of death or reasons for euthanasia in military working dogs: 927 cases (1993–1996). Journal of the

American Veterinary Medical Association, 219(2), 209–214. https://doi.org/10.2460/javma.2001.219.209

Moore, T., Hecquet, S., McLellann, A., Ville, D., Grid, D., Picard, F., Moulard, B., Asherson, P., Makoff, A. J., McCormick, D., Nashef, L., Froguel, P., Arzimanoglou, A., LeGuern, E., & Bailleul, B. (2001). Polymorphism analysis of JRK/JH8, the human homologue of mouse jerky, and description of a rare mutation in a case of CAE evolving to JME. Epilepsy Research, 46(2), 157–167. https://doi.org/10.1016/s0920-1211(01)00275-3

Mustard, J. F., & Packham, M. A. (1968). The unrealized potenital of animal diseases in the study of human diseases. PubMed, 98(19), 887–890. https://pubmed.ncbi.nlm.nih.gov/5689807

Nabekura, T., Deborah, E. A., Tahara, S., Arai, Y., Love, P. E., Kako, K., Fukamizu, A., Muratani, M., & Shibuya, A. (2023). Themis2 regulates natural killer cell memory function and formation. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-42578-8

Nachat, R., Cipolat, S., Sevilla, L. M., Chhatriwala, M., Groot, K. R., & Watt, F. M. (2009). KazrinE is a desmosome-associated liprin that colocalises with acetylated microtubules. Journal of Cell Science, 122(22), 4035–4041. https://doi.org/10.1242/jcs.047266

National Institutes of Health. (2022). SRA-Toolkit (v3.0.2) [Software]. National Center for Biotechnology Information. https://hpc.nih.gov/apps/sratoolkit

NCBI Canis lupus familiaris annotation. (2021). In National Center for Biotechnology Information (Release 106). https://www.ncbi.nlm.nih.gov/refseq/annotation_euk/Canis_lupus_familiaris/106/

New Zealand Police. (2022). Police Manual - Police Dogs. https://www.police.govt.nz/about-us/publication/police-dogs-police-manual-chapter

Ng, P. C., & Kirkness, E. F. (2010). Whole genome sequencing. Methods in Molecular Biology, 215–226. https://doi.org/10.1007/978-1-60327-367-1_12

Nganvongpanit, K., Euppayo, T., Siengdee, P., Buddhachat, K., Chomdej, S., & Ongchai, S. (2020). Post-treatment of hyaluronan to decrease the apoptotic effects of carprofen in

canine articular chondrocyte culture. PeerJ, 8, e8355. https://doi.org/10.7717/peerj.8355

Nicholas, F. (1978). Pituitary dwarfism in German Shepherd Dogs: a genetic analysis of some Australian data. Journal of Small Animal Practice, 19, 167–174. https://onlinelibrary.wiley.com/doi/10.1111/j.1748-5827.1978.tb05471.x

Niu, L., Zhou, Y., Zhang, W., Yan, Y., & Ren, Y. (2021). ARHGEF19 promotes the growth of breast cancer in vitro and in vivo by the MAPK pathway. Physiology International, 108(4), 399–411. https://doi.org/10.1556/2060.2021.00187

Oakley, G. G., Tillison, K., Opiyo, S. A., Glanzer, J. G., Horn, J. M., & Patrick, S. M. (2009). Physical Interaction between Replication Protein A (RPA) and MRN: Involvement of RPA2 Phosphorylation and the N-Terminus of RPA1. Biochemistry, 48(31), 7473–7481. https://doi.org/10.1021/bi900694p

Oberbauer, A. M., Keller, G. G., & Famula, T. R. (2017). Long-term genetic selection reduced prevalence of hip and elbow dysplasia in 60 dog breeds. PLoS ONE, 12(2), e0172918. https://doi.org/10.1371/journal.pone.0172918

Oberbauer, A. M., & Sampson, J. (1999). Pedigree Analysis, Genotype Testing and Genetic Counselling. In The Genetics of the Dog (pp. 461–485). CAB International.

Paciello, O., Maiolino, P., Fatone, G., & Papparella, S. (2003). Mitochondrial Myopathy in a German Shepherd Dog. Veterinary Pathology, 40(5), 507–511. https://doi.org/10.1354/vp.40-5-507

Padgett, G. A. (1998). Control of canine genetic diseases. Howell Book House.

Parr, J. R., & Otto, C. M. (2013). Emergency visits and occupational hazards in German Shepherd police dogs (2008–2010). Journal of Veterinary Emergency and Critical Care, 23(6), 591–597. https://doi.org/10.1111/vec.12098

Parry, B. W., Howard, M. A., Mansell, P. D., & Holloway, S. A. (1988). Hemophilia A in German Shepherd Dogs. Australian Veterinary Journal, 65(9), 276–279.

Plassais, J., Kim, J., Davis, B. W., Karyadi, D. M., Hogan, A. N., Harris, A. C., Decker, B., Parker, H. G., & Ostrander, E. A. (2019). Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-09373-w

Police dog section. (n.d.). New Zealand Police. Retrieved September 12, 2024, from https://www.police.govt.nz/about-us/structure/teams-units/dog-section

Police Dog Trust | New Zealand. (n.d.). Retrieved September 12, 2024, from https://www.policedogtrust.co.nz/

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van Der Auwera, G. A., Kling, D. E., Gauthier, L. D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M. J., Neale, B., MacArthur, D. G., & Banks, E. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv (Cold Spring Harbor Laboratory). https://doi.org/10.1101/201178

Redwood, M. M. (1980). A dog's life: Working Dogs in New Zealand. A.H. & A.W. Reed.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. Nature Biotechnology, 29(1), 24–26. https://doi.org/10.1038/nbt.1754

Roessler, E., Ouspenskaia, M. V., Karkera, J. D., Vélez, J. I., Kantipong, A., Lacbawan, F., Bowers, P., Belmont, J. W., Towbin, J. A., Goldmuntz, E., Feldman, B., & Muenke, M. (2008). Reduced NODAL signaling strength via mutation of several pathway members including FOXH1 is linked to human heart defects and holoprosencephaly. The American Journal of Human Genetics, 83(1), 18–29. https://doi.org/10.1016/j.ajhg.2008.05.012

Salt, V., & Salt, C. (1972). Born to Obey. Collins.

Saredi, G., Huang, H., Hammond, C. M., Alabert, C., Bekker-Jensen, S., Forne, I., Reverón-Gómez, N., Foster, B. M., Mlejnkova, L., Bartke, T., Cejka, P., Mailand, N., Imhof, A., Patel, D. J., & Groth, A. (2016). H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. Nature, 534(7609), 714–718. https://doi.org/10.1038/nature18312

Sargan, D. R., Sampson, J., & Binns, M. M. (2001). Molecular Genetics of the Dog. In The Genetics of the Dog (pp. 139–159). CAB International.

Sevilla, L. M., Nachat, R., Groot, K. R., & Watt, F. M. (2008). Kazrin regulates keratinocyte cytoskeletal networks, intercellular junctions and differentiation. Journal of Cell Science, 121(21), 3561–3569. https://doi.org/10.1242/jcs.029538

Soo, M., Sneddon, N., Lopez-Villalobos, N., & Worth, A. (2014). Genetic evaluation of the total hip score of four populous breeds of dog, as recorded by the New Zealand Veterinary Association Hip Dysplasia Scheme (1991–2011). New Zealand Veterinary Journal, 63(2), 79–85. https://doi.org/10.1080/00480169.2014.961581

Sotomayor-Vivas, C., Hernández-Lemus, E., & Dorantes-Gilardi, R. (2022). Linking protein structural and functional change to mutation using amino acid networks. PLoS ONE, 17(1), e0261829. https://doi.org/10.1371/journal.pone.0261829

Sponenberg, D. P., & Rothschild, M. F. (2001). Genetics of Coat Colour and Hair Texture. In The Genetics of the Dog (pp. 61–87). CAB International.

Stock, K., Klein, S., Tellhelm, B., & Distl, O. (2011). Genetic analyses of elbow and hip dysplasia in the German shepherd dog. Journal of Animal Breeding and Genetics, 128(3), 219–229. https://doi.org/10.1111/j.1439-0388.2010.00901.x

Stokol, T., Parry, B., Mansell, P., & Richardson, J. (1994). Hematorrhachis associated with hemophilia A in three German shepherd dogs. Journal of the American Animal Hospital Association, 30(3), 239–243. https://www.researchgate.net/publication/279718296_Hematorrhachis_associated_with_hemophilia_A_in_three_German_shepherd_dogs

Stölting, G., Fischer, M., & Fahlke, C. (2014). CLC channel function and dysfunction in health and disease. Frontiers in Physiology, 5. https://doi.org/10.3389/fphys.2014.00378

Strickland, W. G. (1998). The German Shepherd today. Turner Publishing Company.

Sturaro, E., Menegazzo, L., Piccinini, P., Bittante, G., Carnier, P., & Gallo, L. (2006). Prevalence and genetic parameters for hip dysplasia in Italian population of purebred dogs. Italian Journal of Animal Science, 5(2), 107–116. https://doi.org/10.4081/ijas.2006.107

Tachie-Menson, T. (2020). Investigating the role of FAM83H, a protein mutated in Amelogenesis Imperfecta [PhD dissertation, University of Dundee].

https://discovery.dundee.ac.uk/ws/portalfiles/portal/42014406/Theresa_Tachie_Menson_Thesis_2_.pdf

Tadepally, H. D., Burger, G., & Aubry, M. (2008). Evolution of C2H2-zinc finger genes and subfamilies in mammals: Species-specific duplication and loss of clusters, genes and effector domains. BMC Evolutionary Biology, 8(1). https://doi.org/10.1186/1471-2148-8-176

Tanaka, K., Kato, A., Angelocci, C., Watanabe, M., & Kato, Y. (2014). A potential molecular pathogenesis of cardiac/laterality defects in Oculo-Facio-Cardio-Dental syndrome. Developmental Biology, 387(1), 28–36. https://doi.org/10.1016/j.ydbio.2014.01.003

Tarnopolsky, M. A., & Raha, S. (2005). Mitochondrial myopathies: diagnosis, exercise intolerance, and treatment options. Medicine & Science in Sports & Exercise, 37(12), 2086–2093. https://doi.org/10.1249/01.mss.0000177341.89478.06

Tenner, E. (2017). Constructing the German Shephard Dog. Raritan, 36 (3), 90–115. https://www.researchgate.net/publication/316522740_Constructing_the_German_Shepherd_Dog

Tkaczyk-Wlizło, A., Kowal, K., & Ślaska, B. (2022). Mitochondrial DNA alterations in the domestic dog (Canis lupus familiaris) and their association with development of diseases: A review. Mitochondrion, 63, 72–84. https://doi.org/10.1016/j.mito.2022.02.001

Ton, Q. V., Leino, D., Mowery, S. A., Bredemeier, N. O., Lafontant, P. J., Lubert, A., Gurung, S., Farlow, J. L., Foroud, T. M., Broderick, J., & Sumanas, S. (2018). Collagen COL22A1 maintains vascular stability and mutations in COL22A1 are potentially associated with intracranial aneurysms. Disease Models & Mechanisms, 11(12). https://doi.org/10.1242/dmm.033654

Tsai, K. L., Noorai, R. E., Starr-Moss, A. N., Quignon, P., Rinz, C. J., Ostrander, E. A., Steiner, J. M., Murphy, K. E., & Clark, L. A. (2011). Genome-wide association studies for multiple diseases of the German Shepherd Dog. Mammalian Genome, 23(1–2), 203–211. https://doi.org/10.1007/s00335-011-9376-9

Tsai, K. L., Starr-Moss, A. N., Venkataraman, G. M., Robinson, C., Kennedy, L. J., Steiner, J. M., & Clark, L. A. (2013). Alleles of the major histocompatibility complex play a role

in the pathogenesis of pancreatic acinar atrophy in dogs. Immunogenetics, 65(7), 501–509. https://doi.org/10.1007/s00251-013-0704-y

Vasimuddin, M., Misra, S., Li, H., & Aluru, S. (2019). Efficient Architecture-Aware acceleration of BWA-MEM for multicore systems. 2022 IEEE International Parallel and Distributed Processing Symposium (IPDPS). https://doi.org/10.1109/ipdps.2019.00041

Vila, C., Maldonado, J. E., & Wayne, R. K. (1999). Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. Journal of Heredity, 90(1), 71–77. https://doi.org/10.1093/jhered/90.1.71

Von Stephanitz, M., & Schwabacher, J. (1923). The German shepherd dog in word and picture (C. Charke, Trans.). Jena.

Voorbij, A., & Kooistra, H. (2009). Pituitary Dwarfism in German Shepherd Dogs. Journal of Veterinary Clinical Science, 2(1), 4–11. https://www.researchgate.net/publication/46714439

Wahl, J. M., Herbst, S. M., Clark, L. A., Tsai, K. L., & Murphy, K. E. (2008). A review of hereditary diseases of the German shepherd dog. Journal of Veterinary Behavior, 3(6), 255–265. https://doi.org/10.1016/j.jveb.2008.05.004

Waldron, R., De Los Angeles Becerra Rodriguez, M., Williams, J. M., Ning, Z., Ahmed, A., Lindsay, A., & Moore, T. (2024). JRK binds satellite III DNA and is necessary for the heat shock response. Cell Biology International, 48(8), 1212–1222. https://doi.org/10.1002/cbin.12216

Wang, C., Wallerman, O., Arendt, M., Sundström, E., Karlsson, Å., Nordin, J., Mäkeläinen, S., Pielberg, G. R., Hanson, J., Ohlsson, Å., Saellström, S., Rönnberg, H., Ljungvall, I., Häggström, J., Bergström, T. F., Hedhammar, Å., Meadows, J. R. S., & Lindblad-Toh, K. (2021). A novel canine reference genome resolves genomic architecture and uncovers transcript complexity. Communications Biology, 4(1). https://doi.org/10.1038/s42003-021-01698-x

Wang, J., Raskin, L., Samuels, D. C., Shyr, Y., & Guo, Y. (2014). Genome measures used for quality control are dependent on gene function and ancestry. Bioinformatics, 31(3), 318–323. https://doi.org/10.1093/bioinformatics/btu668

Wang, S., Zhang, H., Hu, C., Liu, J., Chadha, S., Kim, J., Simmer, J., & Hu, J. (2020). FAM83H and autosomal dominant hypocalcified amelogenesis imperfecta. Journal of Dental Research, 100(3), 293–301. https://doi.org/10.1177/0022034520962731

Wang, X., & Shen, K. (2024). New scissor for lysosome. The Innovation Life, 2(2), 100064. https://doi.org/10.59717/j.xinn-life.2024.100064

Weigel, K. (2001). Controlling inbreeding in modern breeding programs. Journal of Dairy Science, 84, E177–E184. https://doi.org/10.3168/jds.s0022-0302(01)70213-5

Weizmann Institute of Science & LifeMap Sciences. (2025). BEND4 Gene - BEN Domain Containing 4. Gene Cards - the Human Genome Database. Retrieved March 18, 2025, from https://www.genecards.org/cgi-bin/carddisp.pl?gene=BEND4&keywords=bend4#publications

Wiberg, M. (2004). Pancreatic acinar atrophy in German shepherd dogs and rough-coated Collies. Etiopathogenesis, diagnosis and treatment. A review. Veterinary Quarterly, 26(2), 61–75. https://doi.org/10.1080/01652176.2004.9695169

Williams, J. P., Mouilleron, S., Trapero, R. H., Bertran, M. T., Marsh, J. A., & Walport, L. J. (2024). Structural insight into the function of human peptidyl arginine deiminase 6. Computational and Structural Biotechnology Journal, 23, 3258–3269. https://doi.org/10.1016/j.csbj.2024.08.019

Willis, M. B. (1992). The German Shepherd Dog: a genetic history. H.F. & G. Witherby. https://archive.org/details/germanshepherddo0000will/

Wininger, F., Zeng, R., Johnson, G., Katz, M., Johnson, G., Bush, W., Jarboe, J., & Coates, J. (2011). Degenerative Myelopathy in a Bernese Mountain Dog with a Novel SOD1 Missense Mutation. Journal of Veterinary Internal Medicine, 25(5), 1166–1170. https://doi.org/10.1111/j.1939-1676.2011.0760.x

Wootton, B. H. (1988). The German shepherd dog. David & Charles Publishers.

Worth, A. (2015). A study of debilitating orthopaedic conditions of working New Zealand Police German Shepherd Dogs: A thesis presented in partial fulfilment of the requirements for a doctorate degree of Veterinary Science at Massey University, Manawatu, New Zealand [PhD Dissertation, Massey University]. https://mro.massey.ac.nz/handle/10179/7150

Worth, A., Sandford, M., Gibson, B., Stratton, R., Erceg, V., Bridges, J., & Jones, B. (2013). Causes of loss or retirement from active duty for New Zealand police German shepherd dogs. Animal Welfare, 22(2), 167–174. https://doi.org/10.7120/09627286.22.2.167

Wright, P. D., Veale, E. L., McCoull, D., Tickle, D. C., Large, J. M., Ococks, E., Gothard, G., Kettleborough, C., Mathie, A., & Jerman, J. (2017). Terbinafine is a novel and selective activator of the two-pore domain potassium channel TASK3. Biochemical and Biophysical Research Communications, 493(1), 444–450. https://doi.org/10.1016/j.bbrc.2017.09.002

Wright, S. (1922). Coefficients of inbreeding and relationship. The American Naturalist, 56(645), 330–338. https://doi.org/10.1086/279872

Wu, X., Jing, Z., Huang, T., & Jing, Y. (2023). BOP1 Promotes Prostate Cancer through the DUSP6/MAPK Pathway. Archivos Españoles De Urología, 76(6), 445. https://doi.org/10.56434/j.arch.esp.urol.20237606.54

Xu, L., Zhang, W., Zhang, H., Yang, X., Ceccobelli, S., Zhao, Y., & E, G. (2024). Identification of GOAt supernumerary Teat phenotype using Wide-Genomic Copy Number variants. Animals, 14(22), 3252. https://doi.org/10.3390/ani14223252

Yamamoto, F. (2017). Evolutionary divergence of the ABO and GBGT1 genes specifying the ABO and FORS blood group systems through chromosomal rearrangements. Scientific Reports, 7(1). https://doi.org/10.1038/s41598-017-09765-2

Yamamoto, F., Yamamoto, M., & Blancher, A. (2009). IMMUNOHEMATOLOGY: Generation of histo-blood group B transferase by replacing the N-acetyl-d-galactosamine recognition domain of human A transferase with the galactose-recognition domain of evolutionarily related murine α1,3-galactosyltransferase. Transfusion, 50(3), 622–630. https://doi.org/10.1111/j.1537-2995.2009.02463.x

Zhou, D. (2024). Genetic basis of alpha-1,3 galactose epitopes in animal reservoir of coronaviruses. Research Square (Research Square). https://doi.org/10.21203/rs.3.rs-3927077/v1

Zhou, J., Li, Q., Deng, X., Peng, L., Sun, J., Zhang, Y., & Du, Y. (2024). Comprehensive analysis identifies ubiquitin ligase FBXO42 as a tumor-promoting factor in neuroblastoma. Scientific Reports, 14(1). https://doi.org/10.1038/s41598-024-69760-2

Chapter 9 – Appendix

Appendix A: Candidate variants resulting from recessive trait filtering including the gene name and ID, the chromosome (Chr) the variant was found on, the variant type (Var) and impact (Imp). Also includes a broad summary of function for each gene.

Gene Name	Gene ID	Chr	Var	Imp	Function	References
Alpha 1,3-Galactosyltransferase 2	*A3GALT2*	2	SNP	Mod	Transfer protein involved in immune function and lipid glycosylation.	Blancher (2013) Yamamoto et al. (2009) Yamamoto (2017) Zhou (2024)
Rho Guanine Nucleotide Exchange Factor 19	*ARHGEF19*	2	SNP	Mod	Associated with wound healing through organisation actin filaments. Facilitates directional migration of keratinocytes.	Deng et al. (2021) Niu et al. (2021)
ATPase Phospholipid Transporting 8A1	*ATP8A1*	13	SNP	Mod	Highly expressed in murine and human platelets, involved in vesicle mediated trafficking.	Jing et al. (2019) Korneenko et al. (2015)
BEN Domain Containing 4	*BEND4*	13	SNP	Mod	Involved with DNA binding. Little information available.	Weizmann Institute of Science & LifeMap Sciences (2025)
BOP1 Ribosomal Biogenesis Factor	*BOP1*	13	SNP	Mod	Regulator of ribosome biogenesis and cell proliferation.	Ahn et al. (2016) Borah et al. (2019) Chung et al. (2011) Wu et al. (2023)
Caspase 9	*CASP9*	2	SNP	Mod	Involved with apoptosis of chondrocytes (cartilage formation and bone development). Marker for mammary tumour via dysregulated cell death.	Gherman et al. (2024) Nganvongpanit et al. (2020)
Chloride Voltage-Gated Channel Ka	*CLCNKA*	2	SNP	Mod	Renal chloride channel for salt reabsorption. Risk factor for heart failure and increased blood pressure.	Levin et al. (2022) Stölting et al. (2014)
Collagen Type XXII Alpha 1 Chain	*COL22A1*	13	SNP	Mod	Involved in the stabilisation of myotendinous junctions, collagen production and maintains vascular integrity.	Koch et al. (2004) Malbouyres et al. (2022) Ton et al. (2018)
Ciliogenesis And Planar Polarity Effector Complex Subunit 2	*CPLANE2*	2	SNP	Mod	Involved with cilium assembly and planar cell polarity for embryonic development. May be involved with lipid binding. Causes issues with CNS and skeletal development	Gray et al. (2009) Langousis et al. (2022)
Family With Sequence Similarity 131 Member C	*FAM131C*	2	SNP	Mod	Unclear function. Related to supernumerary teats in goats and meat colour in pigs.	Liu et al. (2024) Xu et al. (2024)
F-Box Protein 42	*FBXO42*	2	SNP	Mod	Post translational modification.	Zhou et al. (2024)

Appendix A – Cont.: Candidate variants resulting from recessive trait filtering including the gene name and ID, the chromosome (Chr) the variant was found on, the variant type (Var) and impact (Imp). Also includes a broad summary of function for each gene.

Gene Name	Gene ID	Chr	Var	Imp	Function	References
Forkhead Associated Phosphopeptide Binding Domain 1	*FHAD1*	2	SNP	Mod	Involved with motile cilia proteins, regulates sperm motility and spermatocyte meiosis.	Kassymbekova et al. (2025)
Forkhead Box H1	*FOXH1*	13	SNP	Mod	Transcription factor, binds to ASE region, related to left-right symmetry. Also involved with cell differentiation and regulation of developmental patterning.	Mair et al. (2019) Roessler et al. (2008) Tanaka et al. (2014)
IQ Motif and Ankyrin Repeat Containing 1	*IQANK1*	13	SNP	Mod	Organisation of cells cytoplasmic membranes and cytoskeleton.	Forman et al. (2012) Tachie-Menson (2020) Wang et al. (2020)
Jrk Helix-Turn-Helix Protein	*JRK*	13	SNP	Mod	DNA binding protein, regulates heat shock response.	Moore et al. (2001) Waldron et al. (2024)
Kazrin, Periplakin Interacting Protein	*KAZN*	2	SNP	Mod	Associated with desmosome activity and epidermal differentiation. Impacts cell shape through keratinocyte activity.	Groot et al. (2004) Nachat et al. (2009) Sevilla et al. (2008)
Potassium Two Pore Domain Channel Subfamily K Member 9	*KCNK9*	13	INDEL	High	Transmembrane potassium channel protein.	Kim et al. (2000) Wright et al. (2017)
Unclassified locus	*LOC102157316*	13	SNP	Mod	Unknown function.	
Unclassified locus	*LOC111093311*	2	SNP	Mod	Unknown function.	
Maestro Heat Like Repeat Family Member 1	*MROH1*	13	SNP	Mod	Involved in lysosomal function and tubulation.	Li et al. (2024) Wang and Shen (2024)
Peptidyl Arginine Deiminase 1	*PADI1*	2	SNP	Mod	Post-translational modification enzyme. Possibly regulates gene expression through histone modification.	Chavanas et al. (2004) Iida and Nakamura (2004)
Peptidyl Arginine Deiminase 6	*PADI6*	2	SNP	Mod	Post-translational modification enzyme. May be involved in chromatin decondensation. Involved with embryonic development.	Chavanas et al. (2004) Williams et al. (2024)

Appendix A – Cont.: Candidate variants resulting from recessive trait filtering including the gene name and ID, the chromosome (Chr) the variant was found on, the variant type (Var) and impact (Imp). Also includes a broad summary of function for each gene.

Gene Name	Gene ID	Chr	Var	Imp	Function	References
Polyhomeotic Homolog 2	*PHC2*	2	SNP	Mod	Related to chromatin-associated protein complex.	Gunster et al. (1997)
Replication Protein A2	*RPA2*	2	SNP	Mod	Involved in DNA replication and repair.	Kemp et al. (2009) Oakley et al. (2009)
Arginine And Serine Rich Protein 1	*RSRP1*	2	SNP	Mod	Associated with spliceosome activity and assembly.	Li et al. (2021)
Thymocyte Selection Associated Family Member 2	*THEMIS2*	2	SNP	Mod	Immune cell regulator.	Cheng et al. (2016) Letko et al. (2023) Nabekura et al. (2023)
Cation Channel Sperm-Associated Auxiliary Subunit	*TMEM249 or CATSPERQ*	13	SNP	Mod	Involved in functions of channel assembly and sperm flagellum formation.	Huang et al. (2023)
Tonsoku Like, DNA Repair Protein	*TONSL*	13	SNP	Mod	DNA repair protein, maintains genomic integrity through histone binding.	Micale et al. (2020) Saredi et al. (2016)
Trafficking Protein Particle Complex Subunit 9	*TRAPPC9*	13	INDEL	High	Involved in nerve growth and differentiation.	Marangi et al. (2012)
Zinc Finger Protein 517	*ZNF517*	13	SNP	Mod	Transcriptional gene regulator.	Lorenz et al. (2010) Tadepally et al. (2008)