Institutionen för husdjursgenetik

Evaluation of genomic DNA from paraffin-embedded tissue and *desmin* as candidate gene for dilated cardiomyopathy in Newfoundland dogs

by

Katarina Davidsson

Handledare:
Göran Andersson
Izabella Baranowska
Nicolette Salmon Hillbertz

Examensarbete 290
2007

Examensarbete ingår som en obligatorisk del i utbildningen och syftar till att under handledning ge de studerande träning i att självständigt och på ett vetenskapligt sätt lösa en uppgift. Föreliggande uppsats är således ett elevarbete och dess innehåll, resultat och slutsatser bör bedömas mot denna bakgrund. Examensarbete på D-nivå i ämnet husdjursgenetik, 20 p (30 ECTS).

Institutionen för husdjursgenetik

Evaluation of genomic DNA from paraffin-embedded tissue and *desmin* as candidate gene for dilated cardiomyopathy in Newfoundland dogs

by

Katarina Davidsson

Agrovoc: Paraffin-embedded, dilated cardiomyopathy, DCM, desmin, Newfoundland

Handledare:
Göran Andersson
Izabella Baranowska
Nicolette Salmon Hillbertz

Examensarbete 290
2007

Examensarbete ingår som en obligatorisk del i utbildningen och syftar till att under handledning ge de studerande träning i att självständigt och på ett vetenskapligt sätt lösa en uppgift. Föreliggande uppsats är således ett elevarbete och dess innehåll, resultat och slutsatser bör bedömas mot denna bakgrund. Examensarbete på D-nivå i ämnet husdjursgenetik, 20 p (30 ECTS).

Abstract

Dilated cardiomyopathy, DCM, affects both dogs and humans, and environmental factors, individual status and genetic causes are known or suspected. DCM is a heart disease where the cardiac muscle increases in size (hypertrophy), as the heart gets dilated and systolic (contractile) pressure decrease. The ineffective contraction is the fundamental defect in DCM. Inherited DCM in humans is a heterogeneous disease with many known causative genes. These known genes can be used as candidate genes for DCM in other species, for example the dog. The inherited variant of DCM in dog appears to be homogeneous within a breed and together with high linkage disequilibrium (LD), which means that large genomic regions are inherited together, dogs are preferable to use in genetic studies. Markers in the genome are compared between individuals to see which regions are shared between individuals.

The overall aim with the DCM project is to determine the genetic association with DCM in Newfoundland dogs. The study reported here is a part of the DCM project and includes an evaluation of genomic DNA from paraffin-embedded tissue for use in genetic studies and a candidate gene approach where association between *desmin* and DCM was evaluated.

Genomic DNA extracted from paraffin-embedded tissue was tested in PCR with and without additional purification steps. All tests indicated that the quality varied between different paraffin blocks and that most samples contained highly degraded genomic DNA. The genomic DNA could still be used to amplify fragments including microsatellites around *desmin* from several individuals. Only with two of the microsatellites, fragments were amplified in sufficient number of individuals and the *desmin* haplotype therefore only consists of two markers separated by 15.4 kb. One of these markers was in addition uninformative (95 % of the tested alleles had the same length). This means that the results should be considered with caution, but they indicated no association between a *desmin* haplotype and the disease.

Index

1. Abbreviations

bp	base pair/s
DCM	dilated cardiomyopathy
dNTP	deoxyribonucleoside triphosphate
gDNA	genomic DNA (deoxyribonucleic acid)
kb	kilobase/s
LD	linkage disequilibrium
PCR	Polymerase chain reaction
SNP	single nucleotide polymorphism
SLU	Sveriges lantbruksuniversitet (Swedish University of Agricultural Sciences)

2. Introduction

The overall purpose of the DCM project is to find genetic association with dilated cardiomyopathy, DCM, in Newfoundland dogs. This knowledge has importance to breed healthier dogs by using dogs without mutated genes in breeding. It can also have importance when the disease in human is studied. Although much is known about DCM in human, this knowledge can have a comparative value.

The Newfoundland breed was chosen because many blocks of paraffin-embedded cardiac tissue are present, which might be used as DNA source for genetic studies. Newfoundland dogs are also an over-represented breed to develop DCM compared to many other breeds (Tidholm & Jönsson 1997).

This study was a part of the overall DCM project and in this study the purpose was to evaluate if the gene encoding the muscle specific intermediate filament desmin contains a causative mutation for DCM in Newfoundland dogs. For this study genomic DNA, gDNA, from paraffin-embedded tissue was used and also evaluated as possible gDNA source for other studies.

2.1 Design of the project

First a method to extract gDNA from paraffin-embedded tissue was chosen by evaluating two methods for paraffin purification and two methods for DNA extraction. The quality of the extracted gDNA was evaluated and other purification steps were tested in effort to improve the quality. After evaluation of the extracted gDNA it was used as samples in the candidate gene approach to evaluate if the *desmin* gene is associated with DCM in Newfoundland dogs. Microsatellites around the *desmin* gene were used as markers and their lengths were gathered to haplotypes. Their possible association with DCM was then evaluated, by comparing haplotypes between individuals (cases and controls).

2.2 DCM

Literally the term cardiomyopathy means heart muscle disease and it is used to describe heart diseases resulting from a primary abnormality in the myocardium. Cardiomyopathies are traditionally divided in three major clinicopathologic groups; dilated, hypertrophic and restrictive (Kumar *et al.* 2003). Hypertrophic and restrictive will not be explained further in this study. Dilated cardiomyopathy occurs in a variety of species, notably dogs, cats and humans (Tidholm & Jönsson 1996) and it is characterized by progressive cardiac hypertrophy, which mean increased size of the heart (without cell division), dilation and contractile (systolic) dysfunction. The ineffective contraction is the fundamental defect in DCM.

The causes of DCM are not always known but can be a result of viral infections or

toxic components or be inherited (Kumar *et al.* 2003). DCM in dogs is suspected to be inherited because of its prevalence in certain breeds and in specific families. Also several other causes have been suggested in dog, including nutritional deficiencies, metabolic disorders, immunologic abnormalities, infectious diseases and decreased contractile function of the left ventricle (myocardial hypokinesis) induced by drugs, toxins or abnormally rapid heart beats (tachycardia) (Tidholm & Jönsson 2005).

In a swedish study by Tidholm and Jönsson (1997) DCM were found in 38 dog breeds, in an age range of 3.5 months to 13 years. 92% of the animals weighed over 15 kg (and 15% over 50 kg), resulting in predominance for DCM in large- and medium-sized breeds weighing more than 15 kg (but less than 50 kg). The study included 189 dogs diagnosed with congestive heart failure caused by DCM. The breeds Airedale terrier, Boxer, Doberman, Newfoundland, Standard poodles, St. Bernard and English cocker spaniel were significantly over-represented (Tidholm & Jönsson 1997). The clinical signs in dogs with DCM are cough, depression, exercise intolerance, inappetence, breathing difficulties (dyspnea), weight loss, abdominal distension, panting, syncope, excessive thirst (polydipsia), weak pulse, systolic murmur and excess fluid in the peritoneal cavity (ascites) (Tidholm & Jönsson 1995 & 1997).

Dogs that are clinically diagnosed with DCM reveal two distinct histological forms of DCM. Cardiomyopathy of Boxers and Doberman Pinschers are called “fatty infiltration-degenerative” type and in many giant, large- and medium-sized breeds DCM can be classified as “attenuated wavy fiber” type. Attenuated wavy fibers are myofibers that are thinner than normal and have wavy appearance. The fibers are separated by a clear space with oedematous fluid and in some cases there is also a diffuse infiltration of subendocardial fibrosis (Tidholm & Jönsson 2005). In a study, 64 of 65 (98%) dogs with confirmed DCM were positive for attenuated wavy fibers. In 147 dogs with other heart disease than DCM only one had trace of attenuated fibers, although not in sufficient numbers to be classified as a positive finding. With these results Tidholm and others (1998) suggested that a histological examination for attenuated wavy fibers might be a useful postmortem test for DCM in dogs and it is said to have a high specificity and sensitivity for DCM. Histopathological changes have also been reported as being characteristic for DCM in humans (Davies 1984).

Inherited DCM in humans is a genetically heterogeneous disease, meaning that many genes are linked to the disorder, possibly giving rise to different phenotypes (Fatkin & Graham 2002). When comparing individuals affected by DCM within dog breeds it appears to be relatively homogenous in separate breeds. This makes the dogs a good model to study the genetic background of spontaneously occurring DCM and possibly associate it to DCM in humans (Stabej *et al.* 2006). Lindbladh-Toh and others (2005) studied human and dog othologous sequences and suggested that the there is a common set of functional elements across the species. Dogs within each breed also have extensive linkage disequilibrium, LD, which means that large fragments in the genome are inherited as blocks. They also have relatively low diversity of haplotypes, where alleles in 80 % of a chromosome only have two to four alleles (Sutter *et al.* 2004). This means that fewer markers are needed when performing a whole-genome association study compared to humans (Lindbladh-Toh *et al.* 2005). In addition only population-based samples are required in difference to linkage mapping, where many generations of families are needed (Sutter *et al.* 2004).

2.3 Newfoundland breed

The Newfoundland breed has an increased incidence for DCM (Tidholm & Jönsson 1997) and shows attenuated wavy fibers (Tidholm *et al.* 1998). Even seven of 15 examined Newfoundland dogs without any abnormalities in echocardiographical examination or clinical signs for heart disease have been reported to be positive for attenuated wavy fibers. None of 32 other breeds without myocardial abnormalities showed attenuated wavy fibers (Tidholm *et al.* 2000). Since attenuated fibers have been shown to have high specificity and sensitivity for DCM (Tidholm *et al.* 1998) these results are, by Tidholm and others (2000), indicating that development of attenuated wavy fibers may represent an early stage of DCM. They also mean that development of attenuated wavy fibers is most probably not a response to chamber dilation and stretching of the myocytes, as suggested by Scheinin and others (1992). A biopsy from such individuals should make it possible to detect future development of DCM in a clinical healthy dog, but unfortunately biopsies for this kind of examination can not be taken in a living dog (Tidholm *et al.* 2000). Instead, it is of great interest to find a tool for predicting future cases using other methods; whole-genome association mapping is one suggestion.

The mode of inheritance indicates how many samples are required in a whole-genome association scan. The mode of inheritance of DCM seems to vary between breeds and has not been clearly defined in most breeds (Alroy *et al.* 2000) but an autosomal dominant mode of inheritance has been suggested in the Newfoundland breed (Göran Andersson, unpublished observation). If this is correct about 50 cases and 50 controls are needed to obtain statistically significant results from an association mapping, according to Karlsson and others (2007, in preparation). Whole genome association scan is expensive and sample collection is time consuming; however, it is possible to evaluate individual genes, prior performing a whole-genome association scan, utilizing a so-called candidate gene approach. In the current study the *desmin* gene has been chosen as a candidate.

2.4 Candidate gene

To find a candidate gene for DCM, three main molecular pathways involved in DCM listed by Stabej and others (2006) can be studied. The first is disturbed integrity of the cytoskeleton, second is disturbed Calcium kinetics and sensitivity and the third is impaired intracellular signalling mechanism. One way, as in this study, is to choose one that is known to cause DCM in humans. The one chosen for this study has also been excluded in Doberman (Stabej *et al.* 2004). A common feature of DCM in humans is disruption of cytoskeletal integrity (Franz *et al.* 2001) and one example of a cause of DCM is a missense mutation in the gene encoding the intermediate filament desmin (Li *et al.* 1999). Newfoundland dogs show attenuated wavy fibers (Tidholm *et al.* 1998), which might be due to a structural mutation. Therefore the gene encoding the structural protein desmin was chosen as candidate for DCM in Newfoundland dogs.

2.4.1 Desmin

Desmin expression is restricted mostly to muscle tissue and is concentrated at Z discs in smooth, skeletal and cardiac muscle cells. Desmin is an intermediate filament that link individual myofibrils, which build up the muscle cell, (Lazarides 1982) and form a three-dimensional scaffold throughout the extrasarcomeric cytoskeleton. The extrasarcomeric cytoskeleton is a complex network of proteins linking the sarcomere (the contractile unit that builds up the myofibrils) with the sarcolemma (which

surrounds the cell) and the extracellular matrix (in the space between cells). This network provides structural support for subcellular structures and transmits mechanical and chemical signals within and between cells (Franz *et al.* 2001). During assembly of myofibrils in biogenesis, desmin may play an important role in the generation of the striated appearance of a muscle (Lazarides 1982). A desmin defect might then be a possible reason for the wavy fiber appearance observed in diseased dogs.

The effect on myocardial mass, myocyte shape and cardiac systolic and diastolic function in absence of desmin has been studied in knock-out mice lacking desmin. This study, reported by Milner and others (1999), was performed in order to find the involvement of desmin in cardiomyopathies and in normal cardiac function. The mice lacking desmin showed development of cardiac hypertrophy with ventricular dilation, compromised systolic function and heart failure later in life. Their hearts were frequently enlarged when compared to hearts of normal mice and both right and left ventricular chambers were dilated. Histological and electron microscopic analysis in both heart and skeletal muscle tissues revealed severe disruption of muscle architecture and degeneration (Milner *et al.* 1996). Several reported cardiomyopathies have showed granular and filamentous aggregates of desmin (Goebel 1995), but there are no cases where desmin is completely absent (Milner *et al.* 1999). This might still not exclude that non-functional desmin can cause the disease. Milner and others (1999) report a similar theory while they argue that the phenotype of desmin null mice support the notion that the observed abnormalities in desmin distribution might have a causal effect.

Desmin has been evaluated as a causing gene in Doberman, and were concluded to not play a role in Doberman DCM (Stabej *et al.* 2004). This does not exclude *desmin* as a candidate for being the causative gene in Newfoundland since the two breeds do not even show the same type of DCM according to histopathological findings, where Doberman do not show wavy fibers (Tidholm & Jönsson 2005).

2.5 Tools in molecular genetics

2.5.1 Genetic polymorphism

All genomes consist of genetic polymorphism, which are sequences that are found in two or more variants (alleles) within a population. If the variants persist in the population with allele frequencies above 1 % for the rarest allele it is called a genetic polymorphism. Single nucleotide polymorphisms (SNPs) and microsatellites are examples of such genetic polymorphism. These can be used as genetic markers to create a haplotype, which can be used to determine association between a haplotype and a phenotype. A haplotype covers a genomic region that is not separated by recombination, which means that markers segregate together, in so-called linkage disequilibrium (LD). LD can for example arise from non-random mating (Gibson & Muse 2004). Since dogs have a history of inbreeding and restricted amount of founders for one breed, they have a high degree of LD. This result in longer haplotypes compared to humans, because the markers are linked and inherited together in a higher frequency than expected (Sutter *et al.* 2004).

2.5.2 Microsatellites

Microsatellites are tandemly repeated DNA and consists of a 1-, 2-, 3-, 4- or 5- base pair repeated unit (Brown 2002, Page & Holmes 1998). Microsatellites can, as mentioned, be polymorphic, where the number of repeated units differs between alleles in individuals of a species. The different lengths can create an individual

genetic profile, which can be used in forensic science or to establish relationship. Microsatellites are popular as markers because of this polymorphism and because they are randomly spread over the genome. They are also relatively short which make them easy (fast and accurate) to amplify in PCR (Brown 2002). Microsatellites are thought to be produced by mutation, unequal crossing-over and DNA slippage. DNA slippage occurs when DNA strands mispair during replication and recombination. One DNA strand creates a loop and the DNA repair mechanism will either remove the loop (decrease in size) or put in new nucleotides in the created gap (increase in size) (Page & Holmes 1998). This can also happen in PCR and should be considered when analysing the measured lengths of microsatellites amplified in PCR.

In this study microsatellites were used to define haplotypes found in the study population. In a microsatellite analysis microsatellites around a gene are chosen and the alleles of each selected microsatellite are defined in each individual. This data are used to calculate the observed and possible haplotypes in the individuals. The haplotypes are compared between cases and controls to see if a haplotype is associated with the phenotype. If there is no association between any of the observed haplotypes and the phenotype the number of cases are expected to be the same as the number of controls. The observed number is compared to the expected number and a χ^2 value can be calculated.

3. Material & method

3.1 Samples

Cardiac tissue, from deceased Newfoundland dogs with or without DCM, has been collected by veterinarian Anna Tidholm (Albano Animal hospital, Stockholm). The tissues were prepared with 4% buffered formalin and washed two times with 70% ethanol, one time with 80% ethanol, two times with 95% ethanol and one time with 99,5% ethanol followed by two times preparation with 100% xylene and finally embedding in paraffin for histology (Professor Lennart Jönsson, Department of Pathology, SLU, Uppsala, personal notification).

After histological findings the blocks were given to the Department of Animal Breeding and Genetics to be examined for use in genetic studies. 72 paraffin blocks were used (43 cases and 29 controls) and all these samples are listed in appendix 1. From a few individuals there are blood samples too, but they were not used in this study. They can be used to confirm these results and/or be saved for future studies, for example whole genome association mapping.

3.2 Evaluation of DNA-preparation method and gDNA from paraffin-embedded tissue

3.2.1 DNA extraction

Two methods to purify the tissue samples from paraffin were tested. In the first protocol; octane/methanol-protocol, (appendix 2a), the paraffin was detached from the tissue with n-octane and the tissue was protected with a small volume of 100% methanol when the n-octane layer with paraffin was removed. The methanol was then removed from the tissue before DNA extraction. In the other protocol; xylene-protocol, (appendix 2b), xylene was used to remove paraffin. After vortex and centrifugation xylene that had bound the paraffin was removed. The tissue was washed with absolute ethanol before DNA extraction. These paraffin purification protocols were tested before two DNA extraction methods; extraction with a standard

(salt) protocol (appendix 3) and E.Z.N.A. tissue DNA kit. All combinations were evaluated.

After evaluation octane/methanol protocol and standard (salt) protocol were used to purify the samples from paraffin and extract gDNA from the paraffin blocks listed in appendix 1. One additional step was made from the tested protocol. After precipitation the pellets were washed with 70% ethanol to remove any trace of NaCl, which was bound by the 30% water in the solution. All samples were tested to amplify a 600 bp fragment in PCR to see which have potential to be used in future studies. The samples were also used in a candidate gene approach. Both tests are explained later in this report.

3.2.2 Extraction evaluation

To evaluate which extraction method to use, samples from the same paraffin block (individual NF115) were used to make the comparison more accurate. These samples are called the evaluation samples. The block from NF115 was chosen only because more blocks were available for that individual. Both the evaluation samples and some other samples, included in this evaluation, were weighed before preparation to calculate the yield of ng DNA per mg tissue. The DNA concentration and purity (OD 260 nm/280 nm ratio) were measured in NanoDrop ND1000 Spectrophotometer. According to the NanoDrop manual the OD 260/280 ratio for gDNA are acceptable between 1.8 and 2.0.

3.2.2.1 Agarose gel electrophoresis

The prepared samples were run on a 1% agarose gel (SeaKem®GTG®Agarose ≥ 1kb) to see if there was any gDNA in the samples. This method was also used to see if tested PCRs were able to amplify a fragment of expected size.

3.2.2.2 PCR

Polymerase chain reaction, PCR, was used to evaluate the quality of the gDNA after preparation. In PCR the DNA strands are separated, specific primers anneal and new nucleotides build up a new strand (a copy) at different temperatures in several cycles. The nucleotide binding is supported by a polymerase and a PCR buffer and $MgCl_2$ are present to create optimal conditions. The primers anneal to specific locations in the DNA and the product will therefore have different size according to which primer pair is used.

Different PCR protocols were performed and primer pair for 250 bp fragments, 600 bp fragments and 1.500 bp (1.5 kb) were tested, (see appendix 4 for 250 bp- and 600 bp- PCR).

PCR products of 250 bp were gel purified (after separation with agarose gel electrophoresis) using E.Z.N.A. Gel Extraction Kit. After gel purification the PCR products were sequenced in MegaBACE 1000 to evaluate the quality of the gDNA. The analysis was made in the computer software CodonCode Alignment (v.1.6.2).

3.2.3 Evaluation of gDNA from paraffin-embedded tissue

To evaluate the quality of gDNA from paraffin-embedded tissue both evaluation samples and other samples were used. This evaluation was performed in collaboration with research assistant Katarina Stenshamn (Small Animal Clinical Sciences, SLU, Uppsala) who extracted gDNA from paraffin-embedded renal tissue from boxers. The same preparations were used for her samples and all samples were therefore considered to be comparable.

The evaluation consisted of testing different PCR with and without additional purification steps, which are described later.

3.2.3.1 PCR and primer design

The same PCR protocol for a 600 bp-fragments as tested with the evaluation sample was tested with some additional samples (see appendix 4 for PCR protocol). To see if a slightly longer fragment could be amplified a fragment of around 800 bp was chosen. The longer fragments that can be amplified in PCR, the higher is the possibility to use the sample in a genome scan.

Primers to amplify a fragment of around 800 bp including regulatory sequence upstream the *desmin* gene were designed (see table 1). The fragment should start approximately 50 bp downstream the Cap-site (transcription start) and then follow upstream to get a fragment size of around 800 bp. A sequence was retrieved from Ensembl Dog (release 40, Aug 2006) containing exon 1 and 1500 bp upstream ATG. The Cap-site was found by comparing the dog and human sequences with known position of the TATA-box and Cap-site (Zhenlin Li *et al.* 1989). The retrieved dog sequence with the target (50 bp downstream Cap-site and 700 bp upstream) marked with brackets was pasted in the program Primer3 (Primer3's homepage), which suggested some primer pairs. One primer pair (one forward and one reverse primer) were chosen and ordered from TAG Copenhagen A/S online (Tag Copenhagen's homepage). The regulatory sequence was chosen because of possible additional use in future regulatory studies.

Table 1. Primer design for an 850 bp fragment of regulatory sequence of *desmin*.

Name:	Sequence:	Tm (primer3 / TAGC)	Length:
regDesF	TGTCCCAGGACTGCTCTCTG	61.59 / 61.4	20 bases
regDesR	TAGGCCTGGCTCATGCTG	61.10 / 58.2	18 bases

Although both primers only had complete match to the wanted positions in the dog genome, according to the BLAT function at UCSC Genome Browser (Dog assembly May 2005) and Ensembl Dog BLAST (release 41, Oct 2006), this primer pair gave two PCR products of unexpected size instead of one 850 bp fragment in control samples of boxer gDNA and gDNA from a cross breed. The PCR program was modified to increase the stringency and betaine was added in the PCR-mix to make the template more available but the two PCR products remained unchanged.

To be sure to get at least one functional primer pair, two new primer pairs were designed and ordered. The sequence of the *desmin* gene with flanking regions was retrieved from Ensembl Dog (release 41, Oct 2006). It was not possible to design primers for a new fragment of a correct size in the promoter region close to *desmin*. If the target was moved more inside the gene the forward primer was placed in either a region of around 350 bp with many repetitions or in a region of around 350 bp that also could be found on another chromosome. If the target was moved more upstream the gene the primer design programme, Primer3 (Primer3's homepage), could not find any primers. Instead two fragments inside the gene were used to find possible primers. The first primer pair designed covered exon 4, 5 and 6 and should according to the Ensembl sequence give a fragment of 846 bp. The second primer pair should amplify a fragment of 826 bp covering exon 7. See table 2 for primer sequences and other information. The program Netprimer (Netprimer's homepage) was used to evaluate the structure of the primers and the primer sequences were also compared to the dog genome with the BLAT search in UCSC Genome Browser (Dog assembly, May

2005) to see that they only had perfect matches to the wanted positions. The chosen primers were ordered from TAG Copenhagen A/S online (Tag Copenhagen's homepage). See appendix 4 for PCR protocol.

Table 2. Primer design for fragments over 800 bp inside the *desmin* gene.

Name:	Sequence:	Tm (primer3 / TAGC)	Length:
Des_ex456F	TTACCCCTTTGACCCCTTGT	60.58 / 57.3	20 bases
Des_ex456R	TGAGAGCCAAGGTCATAGCA	59.55 / 57.3	20 bases
Des_ex7F	ACCTGGGTGTCCCTCTCCT	60.93 / 61.0	19 bases
Des_ex7R	CAAGATACATAACGTCTCCATCG	58.66 / 58.9	23 bases

3.2.3.2 Purification of the gDNA

DNA fragments of low molecular weight were suggested to disturb PCR and spin columns were used to remove different sizes of fragments before PCR. A spin column, "BD Chroma Spin™-400", that according to the BD CHROMA SPIN™ Columns User Manual removes fragments smaller than 170 bp were tested and some samples were also purified with phenol/chloroform (appendix 5) before, after or without spin column. The different combinations were tested in both 600 bp- and 1.5 kb- PCR.

Spin columns that removes smaller fragments were also tested. "BD Chroma Spin™-10", which removes most fragments below 4 bp (i.e. NTPs, dNTPs) and NaCl, and "BD Chroma Spin™-100", which remove most fragments below 50 bp (BD CHROMA SPIN™ Columns User Manual), were tested before PCR to amplify an 850 bp fragment. Samples run through a dry "BD Chroma Spin™-10", were tested in a 600 bp PCR. Before adding the sample to the spin column it should spin down. Unfortunately, this time the centrifuge was wrongly programmed and the spin columns went too dry. A large amount of the DNA from these samples remained in the column and the results from this test are therefore not completely reliable.

3.2.3.3 Optimizing PCR for gDNA from paraffin- embedded tissue

To improve the results, optimization of the PCR conditions were tried. To decrease the specificity and perhaps increase the yield the $MgCl_2$ and primer concentration were altered and tested in 600 bp PCR. $MgCl_2$ concentration of 2.0 mM, 2.5 mM and 3.0mM were tested in combination with a primer concentration of 0.2 µM and 0.4 µM. The best combination of primer- and $MgCl_2$- concentration was used in the 850 bp PCR with spin column (BD Chroma Spin™-10 and -100) samples.

After searching for ideas to optimize PCR with degraded and perhaps contaminated DNA, such as ancient DNA, a test with different dilutions of DNA template were performed in 600 bp PCR. The tested dilutions were 100 ng/reaction, 50 ng/reaction, 25 ng/reaction, 10 ng/reaction and 2 ng/reaction, in a 20 µl reaction. Diluting DNA template will also dilute disturbing contaminations and less contamination will be present in the PCR. Amplifying will anyhow be possible, because of the high sensitivity of PCR, which in theory only require a single DNA molecule as template.

3.3 Candidate gene approach

To evaluate whether *desmin* is a DCM causing gene in Newfoundlands, three genetic markers that cover the gene were used to define the haplotypes present in different individuals, cases and controls. gDNA from the paraffin blocks listed in appendix 1 were used.

3.3.1 Bioinformatics

The sequence of the gene and 10 kb upstream and downstream were first found in the first version of the dog genome (CanFam1), where the gene was searchable (in Ensembl release 40, Aug 2006). In the unfinished, latest version, of the dog genome sequence (CanFam2) in Ensembl (pre-dog) (Ensembl v.38) the sequence between the same positions was found, but did not match the first. The start of the first sequence (that was retrieved with the gene name in CanFam1) was found manually in the CanFam2-sequence. The start position and the size of the wanted fragment were calculated with the known positions from CanFam1. The new sequence (in CanFam2) started at 28.923.171 and ended at 28.949.704. The *desmin* gene is then found at chromosome 37 between 28.933.171 and 28.939.704 in the latest version of the Dog genome (CanFam2) with a size of 6.5 kb. These positions were needed to find the microsatellites found by UCSC Genome Browser (Dog May 2005 (CanFam2) assembly). In the upgraded Ensembl Dog (release 42, Dec 2006) the location of the *desmin* gene in CanFam2 is between 28.933.165 and 28.939.896 at chromosome 37.

3.3.1.1 Microsatellites around desmin

UCSC Genome Browser (Dog May 2005 (CanFam2) assembly) was used to find repeats and possible microsatellites in the sequence. Two from UCSC Genome Browser (the first and third) and one found manually (the second) were chosen as markers for the *desmin* gene. The first microsatellite is a CT-repeat (~20 repetitions) located around 5.1 kb upstream *desmin*, the second is a GT-repeat (~12 repetitions) located around 2.4 kb upstream *desmin* and the third is an ATTTT-repeat (~8 repetitions) located around 6.5 kb downstream *desmin*. In total, these microsatellites will define a haplotype with the size of 18.1 kb around the *desmin* gene. Table 3 present a microsatellite summary and figure 1 gives an overview of the region covered by the microsatellites and the region to be included in the *desmin* haplotype.

Table 3. Microsatellite summary

Name	Repeat	No of repeats	Distance in bp from *desmin*	PCR product (bp)	Forward primer (5'-3')	Reverse primer (5'-3')
Des1	CT	~20	5,091 (5')	295	TGCAAGACGCTGTACCACAT	GTTGGAAATGGGAAGGTTCA
Des2	GT	~12	2,385 (5')	227	GGCTCCAGTTTACGAATTGC	GGCAAGCTTTCTGTCCTGTC
Des3	ATTTT	~8	6,485 (3')	104	TTAGGGCATGGAACTGCTCT	ATACTGGGCATGGAACAACC

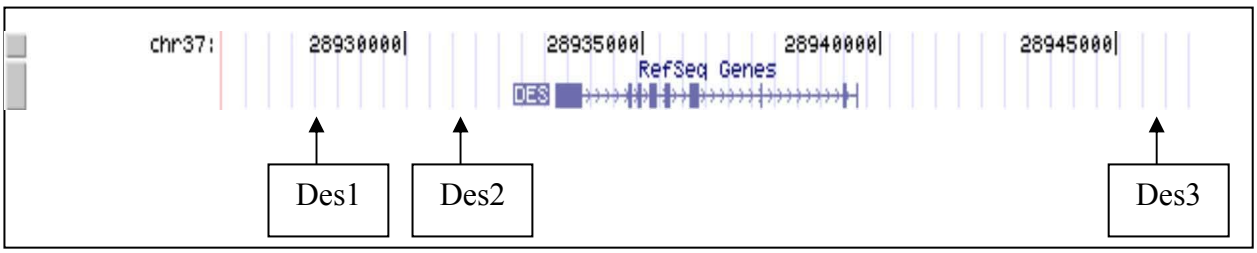


Figure 1. An overview of the *desmin* gene located at chromosome 37 between 28.933.171 and 28.939.704 (in CanFam2) with a size of 6.5 kb. The three microsatellites are marked and covers a region of 18.1 kb. *(UCSC Genome Browser - Dog May 2005 (CanFam2) assembly)*

3.3.2 PCR and primer design

Primer3 (Primer3's homepage) was used to design the primers to amplify fragments including the microsatellites, see table 3. The program Netprimer (Netprimer's homepage) was used to evaluate the structure of the primers and the primer sequences were also compared to the dog genome with the BLAT search in UCSC Genome Browser (Dog assembly, May 2005) to see that they only had perfectly matches to the wanted positions. A M13 (-21) tail, (5´-CACGACGTTGTAAAACGAC-3´), was added to the 5'-end of all forward primers and then the primers were ordered from TAG Copenhagen A/S online (Tag Copenhagen's homepage).

The M13 tail is needed for labelling the PCR products for detection by laser in MegaBACE when genotyping. Instead of labelling each primer with separate dyes a universal fluorescent-labelled M13 (-21) primer can be used. During the first cycles in PCR the forward primer with the M13 (-21) tail and the reverse primer anneal to the DNA template to amplify the wanted region. The M13 (-21) tail gets incorporated in the PCR product and becomes target for the labelled universal M13 (-21) primers in following cycles. In this way the labelling gets incorporated in the PCR product (Schuelke 2000).

Different PCR mixes and programs were tested and the best one was used to amplify the microsatellite fragments (see appendix 4). The fluorescent dyes in the labelled M13 (-21) primers used were TET (6-carboxy-fluorescine) for Des 1 and Des 3, and FAM (tetrachloro-6-carboxy-fluorescine) for Des 2.

All samples were tested on 1% agarose gel (Invitrogen AGAROSE ELECTROPHORESIS GRADE) to see if there were any PCR products. Even samples that showed weak bands were tested in MegaBACE for genotyping.

3.3.3 MegaBACE, Genotyping

A buffer plate containing 190 µl 1 x MegaBACE^{TM} LPA buffer in each well was prepared before each MegaBACE run. The PCR products were first diluted. Since Des 2 and Des 3 had different labelling they were run together for one individual in the same well (multiplexing). 1.5 µl of each PCR product and 3 µl milliQ water were mixed. Des 1 were run in separate wells and where diluted with 1.5 µl PCR product and 4.5 µl milliQ water. Some samples, which showed low intensity when analysing the MegaBACE result, were less diluted (3 µl PCR product in a total volume of 6µl) before a second try in MegaBACE. The diluted PCR products were put on a MegaBACE plate containing 5 µl diluted (1:20) size standard (ET 400-R) in each well. Before MegaBACE run the buffer plate and MegaBACE plate were centrifuged at 2.000 x g for 2 minutes. Six matrix tubes were also centrifuged at 3.000 x g for 3 minutes. When starting the MegaBACE run the instructions in the machine were followed.

The lengths of the PCR products (the fragments including the microsatellites) were detected by laser in the MegaBACE and by comparing the lengths with the size standard each microsatellite allele got a relative length. The raw data from the MegaBACE run were analysed with the MegaBACE^{TM} Genetic Profiler version 2.2 (Amersham Biosciences). In this analysis peaks with highest intensity are possibly the true alleles. Other peaks around, so-called stutter peaks, are probably due to DNA slippage in PCR. Assembling all alleles for one individual give the haplotype for this individual. The single alleles and the haplotypes can be compared to the ones of other individuals and by statistical calculations a certain allele or haplotype can be associated (linked) or not to a certain disease, in this case DCM. The calculations for

analysis of genetic linkage were performed with the program CONTIG version 2.71 (Utility programs for analysis of genetic linkage, copyright 1988, J. Ott).

4. Results

4.1 Evaluation of DNA-preparation method

The results for the evaluation samples (NF115) are listed in table 4.

Table 4. Results of evaluation samples.

New ID	DNA-prep.	Paraffin purifi-cation	DNA conc. (ng/uL)	DNA amount (ng)	260/ 280	Gel >1kb? (fig.3)	PCR 250bp?	PCR 600bp?	PCR 1.5kb?
S.oct.1	Salt prep.	Octane/ methanol	357,3	17865	1,78	-	yes	yes	no
S.oct.2	Salt prep.	Octane/ methanol	216,8	10840	1,72	yes	-	-	-
S.xyl.1	Salt prep.	Xylene	287,0	14350	1,81	yes	yes	yes	no
S.xyl.2	Salt prep.	Xylene	497,8	24890	1,83	-	-	-	-
EZ.oct.wat.1	EZNA:1 (water elution)	Octane/	94,5	9454	1,93	yes	yes	yes	no
EZ.oct.eb.1	EZNA:1 (2:nd elution with EB)	methanol	77,1	7710	1,99	yes	yes	yes	no
EZ.oct.wat.2	EZNA:2 (water elution)	Octane/	50,9	5086	1,95	-	-	-	-
EZ.oct.eb.2	EZNA:2 (2:nd elution with EB)	methanol	35,7	3567	2,05	-	-	-	-
EZ.xyl.wat.1	EZNA:3 (water elution)	Xylene	58,7	5873	1,86	-	yes	yes (weak)	no
EZ.xyl.eb.1	EZNA:3 (2:nd elution with EB)		42,4	4240	1,90	yes	yes	no	no
EZ.xyl.wat.2	EZNA:4 (water elution)	Xylene	91,1	9109	2,01	yes	-	-	-
EZ.xyl.eb.2	EZNA:4 (2:nd elution with EB)		29,7	2970	2,00	-	-	-	-

"EB" stands for Elution Buffer and "-" means that the sample has not been tested.

The 250 bp fragments amplified from the evaluation samples that were purified and sequenced in MegaBACE showed good quality. The evaluation samples were also run directly in agarose gel electrophoresis, which showed that the samples contained gDNA of both low- and high molecular weight, see figure 2.

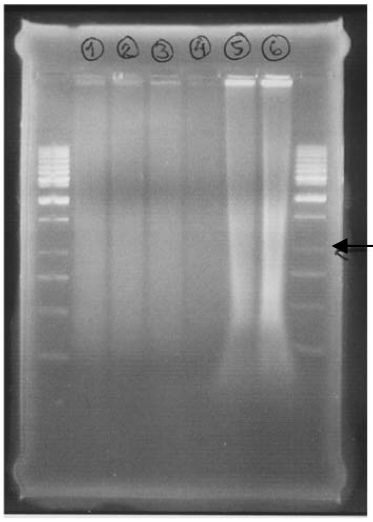


Figure 2. Gel picture of extracted gDNA from paraffin-embedded tissue.

Each well were loaded with 20µl sample of following concentrations; 1) 95 ng/µl
2) 90 ng/µl
3) 75 ng/µl
4) 40 ng/µl
5) 350 ng/µl
6) 300 ng/µl

Some samples, both evaluation samples and others, were weighed before purification and extraction. After concentration measurement with NanoDrop the yield of ng gDNA per mg tissue was calculated. Measurements for the yield calculations are listed in appendix 6. To visualise a possible correlation between the amount of tissue and the amount of extracted gDNA those values were plotted in a diagram, see figure 3. To correlated the two lines should follow each other, but this shows no correlation. A small tissue part sometimes gives more DNA than a bigger tissue part, for example "S.oct.1" compared to "EZ.xyl.1". This means that the yield is not due to how much tissue is prepared, but instead which tissue part is used. Therefore, the yield can not tell if one method is better.

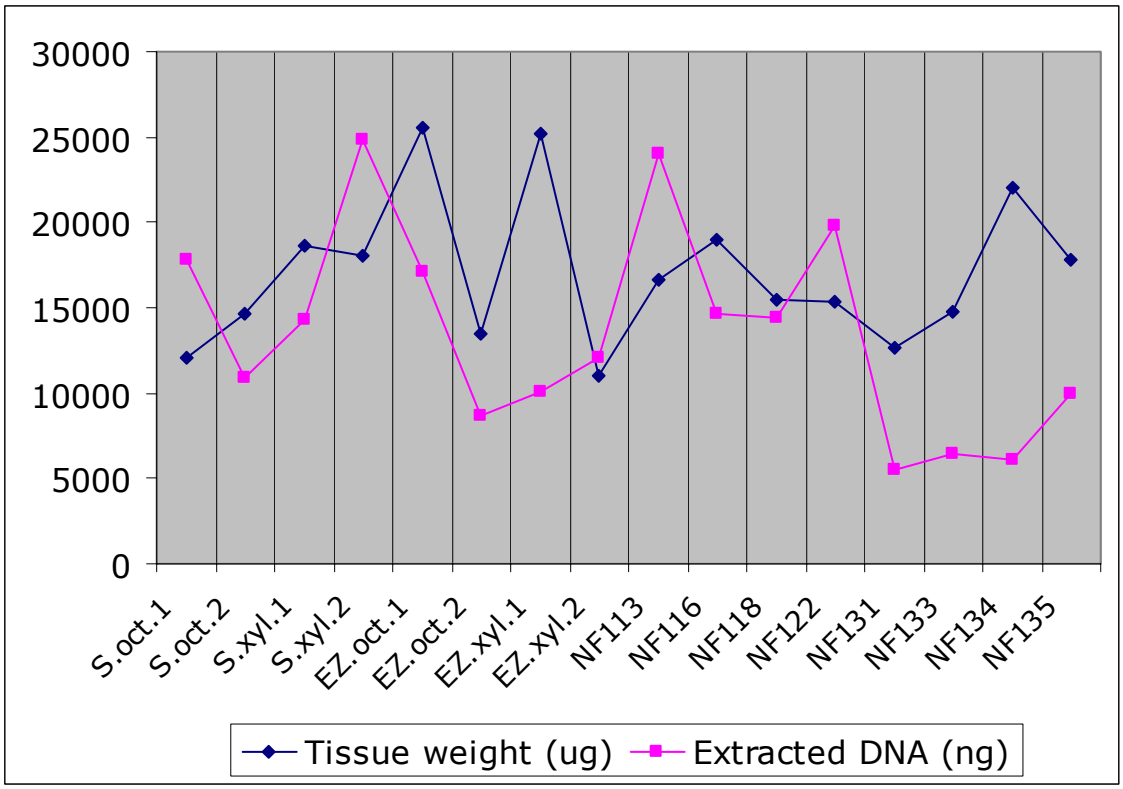


Figure 3. Tissue weight and amount of extracted DNA plotted to visualise a possible association.

In total, neither octane/methanol-protocol nor xylene-protocol gave better result than the other. The octane/methanol protocol was a bit easier and faster and was therefore used for paraffin purification of the original samples. There were no notable variation between the preparation protocols neither, therefore, by economical reasons; the standard (salt) protocol was chosen.

4.2 Evaluation of gDNA from paraffin-embedded tissue

Many combinations of purification and PCR were tested. Samples purified with a BD Chroma Spin™-400 could not amplify a 1.5 kb fragment. Not even when the samples were purified with phenol/chloroform before or after they could amplify a 1.5 kb fragment. Other results are presented in table 5.

Table 5. Result of different purification steps before PCR.

ID	Purification		850bp PCR	600bp PCR
NF113	Chroma Spin -400	Phenol/chloroform	-	no
NF113	Phenol/chloroform		-	no
NF113	Phenol/chloroform	Chroma Spin -400	-	no
105b (KS)	Chroma Spin -400	Phenol/chloroform	-	no
105b (KS)	Phenol/chloroform		-	no
105b (KS)	Phenol/chloroform	Chroma Spin -400	-	no
Control DNA			-	**yes**
Control DNA	Phenol/chloroform		-	**yes**
320b (KS)			-	no
320b (KS)	Phenol/chloroform		-	no
NF118			-	**yes**
NF118	Phenol/chloroform		-	**yes**
EZ.xyl.wat.1			-	**yes**
EZ.xyl.wat.1	Phenol/chloroform		-	**yes**
S.oct.1			-	**yes**
S.oct.1	Chroma Spin -10 (dry)		no	**yes**
320 (KS)			-	**yes**
320 (KS)	Chroma Spin -10 (dry)		no	no
7190 (KS)			-	no
7190 (KS)	Chroma Spin -10 (dry)		-	no
281 (KS)			-	no
281 (KS)	Chroma Spin -10 (dry)		no	**yes**
NF22			no	-
NF22	Chroma Spin -10		no	-
NF22	Chroma Spin -100		no	-
NF77			no	-
NF77	Chroma Spin -10		no	-
NF77	Chroma Spin -400		no	-
NF84			no	-
NF84	Chroma Spin -10		no	-
NF84	Chroma Spin -400		no	-
NF95			no	-
NF95	Chroma Spin -10		no	-
NF95	Chroma Spin -100		no	-
NF100			no	*
NF100	Chroma Spin -10		no	-
NF100	Chroma Spin -100		no	-
NF109			no	*
NF109	Chroma Spin -10		no	-
NF109	Chroma Spin -100		no	-

- = not tested * = samples from the same block have amplified a 600 bp fragment
KS = Katarina Stenshamn's samples

Purification with phenol/chloroform did not seem to improve the success of PCR. The samples that amplified a 600 bp fragment after phenol/chloroform purification (Control DNA, NF118 and EZ.xyl.wat.1) amplified it even without purification.

The darker grey are results from the dry spin column (described in "*3.2.3.2 Purification of the gDNA*") and are therefore unreliable. This is also indicated by the result of 320 (KS), which amplified a 600 bp fragment before but not after the spin column. Despite of this, the result of 281 (KS) might indicate an advantage of spin column use, as this sample amplified a 600 bp fragment after spin column, but not before.

The test with many spin columns would also be interesting to test in experiments of the 600 bp PCR. That might give more reliable results whether spin columns can improve PCR or not. While we have been unable to amplify an 850 bp fragment from

any of the paraffin-embedded samples it is currently impossible to see any change before and after the spin column.

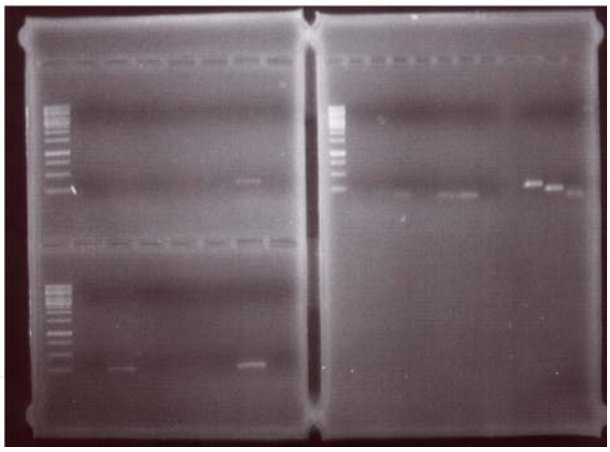
Using different dilutions of template DNA made no advantage and 50 ng in one reaction (of 20 µl) seemed to be an appropriate concentration. Altering $MgCl_2$- and primer concentration resulted in one slightly better combination. The same samples were tested for all combinations and one sample was amplified in all combinations. When $MgCl_2$ concentration was 2.5 mM and primer concentration was 0.4 µM one additional sample was amplified. This combination was then used in 850 bp PCR, but no sample was able to amplify the 850 bp fragment.

All extracted samples were tested in a 600 bp PCR. All samples are, as mentioned earlier, listed in appendix 1, where the ones that amplified a 600 bp fragment are marked.

The possible causes for the varying quality of the extracted gDNA will be discussed later and for this reason the samples with PCR result are sorted according to paraffin block age and listed in appendix 7.

4.3 Is *desmin* a DCM causing gene?

The PCR used to amplify the microsatellite fragments were first tested with few samples and run in agarose gel electrophoresis (see figure 4). Other PCR programs and PCR mixes were tested, but none gave as nice bands on gel as the chosen protocol. This one was chosen to amplify microsatellite fragments from all samples (appendix 1).



GEL 1:

1kb ladder	**Microsatellite 1** 18 22 77 84 95 100 blank
1kb ladder	**Microsatellite 2** 18 22 77 84 95 100 blank

GEL 2:

1kb ladder	**Microsatellite 3** 18 22 77 84 95 100 blank		Positive controls for microsatellite 1, 2 and 3

Figure 4. Gel picture after test amplification in the chosen microsatellite PCR.

After PCR, genotyping in MegaBACE and analysing raw data, different alleles and haplotypes were decided for the individuals, see appendix 8. Since only a few samples were able to amplify microsatellite 1, complete haplotypes could not be made for those individuals. Haplotypes of microsatellite 2 and 3, covering a region of 15.4 kb, which still include the *desmin* gene, were made and their association with DCM was calculated. No association was expected, and the haplotypes and alleles should then have an equal distribution between healthy and diseased. One haplotype that were found in one individual were not included in the calculations. The calculations for allele- and haplotype association are presented in table 6.

Table 6. Observed number of alleles/haplotypes compared to expected number if there is no association.

Observed number

Microsatellite 1

Alleles	**305**	**314**	**320**	
Cases	1	1	5	7
Controls	3	0	4	7
	4	1	9	**14**

Number expected under independence:

	305	**314**	**320**
Cases	2,00	0,50	4,50
Controls	2,00	0,50	4,50

Contributions to chi-square:

	305	**314**	**320**
Cases	0,50	0,50	0,06
Controls	0,50	0,50	0,06

Chi-square: 2.11 df=2
2-sided p-value: 0.347999

Microsatellite 2

Alleles	**248**	**252**	
Cases	2	21	23
Controls	0	16	16
	2	37	**39**

	248	**252**
Cases	1,18	21,82
Controls	0,82	15,18

	248	**252**
Cases	0,57	0,03
Controls	0,82	0,04

Chi-square: 1.47 df=1
2-sided p-value: 0.225877

Microsatellite 3

Alleles	**125**	**145**	**161**	**169**	
Cases	18	2	9	3	32
Controls	13	1	7	1	22
	31	3	16	4	**54**

	125	**145**	**161**	**169**
Cases	18,37	1,78	9,48	2,37
Controls	12,63	1,22	6,52	1,63

	125	**145**	**161**	**169**
Cases	0,00	0,03	0,02	0,17
Controls	0,01	0,04	0,04	0,24

Chi-square: 0.56 df=3
2-sided p-value: 0.906204

Haplotype (microsatellite 2 & 3)

	A	**B**	**C**	**D**	
Cases	12	2	6	2	22
Controls	11	1	5	1	18
	23	3	11	3	**40**

	A	**B**	**C**	**D**
Cases	12,65	1,65	6,05	1,65
Controls	10,35	1,35	4,95	1,35

	A	**B**	**C**	**D**
Cases	0,03	0,07	0,00	0,07
Controls	0,04	0,09	0,00	0,09

Chi-square: 0.41 df=3
2-sided p-value: 0.939205

The hypothesis that there is no association can not be excluded, since the chi-square values are lower than table chi-square values. These results thereby support the hypothesis that there is no association between the *desmin* gene and DCM.

5. Discussion

5.1 DNA-preparation

Both tested paraffin purification methods (octane/methanol- and xylene- protocol) and the two extraction methods (salt preparation and E.Z.N.A. tissue DNA kit) resulted in gDNA from the paraffin-embedded samples tested. From all combinations of methods tested we were able to PCR-amplify fragments up to 600 bp. No obvious differences were noticed between the methods and further analyses were performed with samples purified with octane/methanol protocol and extracted with salt preparation. All

methods should perhaps be tested with more samples in order to get stronger arguments to use a particular combination. The E.Z.N.A. kit might give better result if more samples are tested. VWR bioMarke (2006) has presented supporting results for E.Z.N.A. tissue DNA kit, as they showed that 200 bp fragments could be amplified and DNA of higher molecular weight were present in all tested samples from paraffin-embedded tissue.

Akalu and others (1999) reported that samples from paraffin-embedded tissue were able to amplify fragments up to 959 bp. In the reported study, gDNA was extracted with an extraction buffer (10 mM Tris–HCl, 1% Tween, 0.1 mg/ml proteinase K, 1 mM EDTA, pH 8.0) and purified with QIAquick kit or phenol/chloroform (with ethanol precipitation). The amount of high molecular weight DNA after purification with QIAquick was significantly higher when the total gDNA was analysed on agarose gel. The gDNA was tested in PCR and some samples purified with QIAquick kit amplified fragments of 959 bp. These tests indicated clear improvements when kits were used, supporting the notion that the E.Z.N.A. kit in this study should have been tested with more samples. A report by Isola and others (1994) should also be considered. They reported that prolonged digestion with proteinase K improves the yield of gDNA. New tests with E.Z.N.A. samples should be digested with proteinase K over night, as in the salt preparation protocol, to ensure optimal yield.

Other methods could also be tested. Tests of other paraffin purification and extraction methods were reported by Coombs and others (1999). A combination of digestion with proteinase K, paraffin purification with thermal cycler and gDNA extraction with Chelex-100 (media for DNA extraction) gave best results, as 61% of the analysed samples extracted PCR amplifiable DNA. They concluded that removal of paraffin and purification are main steps required to obtain good results. Techniques involving melting to remove paraffin were showed to be more effective compared to methods using organic solvents to dissolve the paraffin, and melting is also both safer and cheaper. This should be tested with the paraffin samples also in this study. Shi and others (2002) reported an extraction method involving heat-treatment and concluded that temperature and pH affect the outcome. High temperature, 120^{o}C, and pH 6-12 showed satisfactory results, considering yield and amplification in PCR.

Thus, it is apparent that multiple different methodologies are available for preparing gDNA from paraffin-embedded tissue. Different methods can be preferable in different tests and the optimal method in each given case has to be empirically determined.

5.2 gDNA from paraffin-embedded tissue

The amount of extracted gDNA was not correlated with the amount of tissue, (see figure 3). This is probably due to which paraffin block that was used for extraction. Perhaps which part of the paraffin block also matter, since no correlation between samples from the same paraffin block (S.oct.1 & 2, S.xyl. 1 & 2, EZ.oct. 1 & 2 and EZ.xyl. 1 & 2) could be seen. The quality of extracted gDNA also differs between different paraffin blocks. 61 % (44 of 72) of the tested samples were able to amplify a fragment of around 100 bp or 200 bp (microsatellite 2 and 3), but only 18 % (13 of 72) were able to amplify a fragment around 600 bp. No tested samples were able to amplify a fragment of around 850 bp. These results indicate that the gDNA is much degraded and this is a probable reason why additional purification steps do not improve the PCR.

Degraded gDNA might influence the yield, of obtained gDNA from a certain amount of tissue, since it is possible that the precipitation efficiency differs between

long and short DNA fragments when gDNA is concentrated with precipitation. If this is true, samples with high yield would contain more gDNA of high molecular weight and should be able to PCR amplify a long fragment. In this study, the yield of the samples, able to amplify a 600 bp fragment, was compared to the yield of an equal amount of randomly picked samples that have not been able to amplify a 600 bp fragment. The samples used for this comparison are marked in appendix 1 and a figure comparing the values is presented in appendix 9. This comparison showed no difference in yield between the compared samples and both samples with high and low yield were able to amplify a 600 bp fragment. This might indicate that the samples that have not been able to amplify a 600 bp fragment not are as degraded as suspected and instead it could have been contaminations that disturbed the PCR. Another possible explanation is that the samples are slightly degraded and therefore give high yield, but are still too degraded to be able to amplify a long PCR fragment. To evaluate these possibilities, the gDNA samples should be separated on an agarose gel to estimate the distribution of gDNA of different molecular weight.

The age of the paraffin block can be one possible cause of the poorer quality of the extracted gDNA. The list in appendix 7 was used to calculate proportions of how many samples from paraffin blocks of different age were are able to amplify PCR fragments of different sizes. The sample from the oldest paraffin block that could amplify a 600 bp fragment was embedded 1995, and this year was chosen as a border for the proportion calculation. 52 % (15/29) of the samples from paraffin blocks embedded from 1982 to 1994 were able to amplify PCR fragments. None could however amplify a fragment of 300 bp or longer. 72 % (31/43) of the samples from paraffin blocks embedded from 1995 to 2004 were able to amplify PCR fragments of 100 bp to 600 bp. 35 % of all samples from 1995 to 2004 were able to amplify fragments of 300 bp to 600 bp. Many new samples were determined to have better quality than older samples, but the age of the paraffin block cannot be the only cause, as some samples from new paraffin blocks did not amplify any fragments and some old were able to amplify at least shorter fragments. Other causes than age alone must influence the quality and here follow some suggestions of possible explanations. The age and/or other condition of the dog, when deceased, might influence the quality of the tissue before fixation in paraffin. Another important aspect is the time from death, tissue removal and paraffin fixation. Perhaps could also the storage of the paraffin blocks influence the quality.

The samples with enough quality, that is the samples which could amplify a 600 bp fragment, can perhaps be used in whole genome association studies with the Illumina array system that allows shorter fragments to be analyzed compared to the Affymetrix array. Other samples that are more degraded can be used in other, microsatellite-based, association studies. The gDNA might be good enough for other candidate gene approaches, and to obtain better results, shorter fragments including the microsatellites could be used. If the fragments are around 100 bp, more samples will probably give result and more complete haplotypes will give more significant association calculations.

No samples from paraffin-embedded tissue were able to amplify an 850 bp fragment, but the primers were functional for blood samples and can therefore be used in other studies..

5.3 Candidate gene approach

95 % (35/37) of the alleles of microsatellite 2 had the same length. This means that this is an uninformative genetic marker. This can be a reason why UCSC Genome

Browser did not recommend that repeat as a marker. The result based on this marker should therefore be considered with caution. A sequence containing GT-units repeated around 12 times, as microsatellite 2, would be expected to have more alleles and be polymorphic. The result in this study indicates loss of polymorphism in this region. One explaination to this can be that the repeated sequence is a part of a regulatory region and that the length has been selected for. The allele length of microsatellite 3 had large differences and that is probably due to the length of the repeated unit, which is 5 bases (ATTTT). Only one deletion or insertion will decrease or increase the size notably.

Even the MegaBACE result could be discussed, as only some samples are tested twice. To confirm the obtained results the samples should be tested at least once more, in a new PCR and a new MegaBACE run. Perhaps blood samples can be added as controls.

Although the results are not completely conclusive they indicate no association between *desmin* and DCM in Newfoundland dogs. *Desmin* can still have associations with the disease, since the results do not exclude a regulatory mutation that influences the expression of *desmin*. It can also be a totally different gene which causes DCM and there are many candidates, which, as mentioned earlier, are known disease genes for DCM in humans.

When studying the pedigree of the sampled Newfoundland dogs, made in Progeny (ver. 6) by student Katarzyna Koltowska, there is a great variation in who get the disease or not. In one litter almost all get DCM, but in another almost none. This is not in total agreement with classical autosomal dominant inheritance. This originates the question if something else might influence the appearance of DCM. In a Ph.D. thesis by Polona Stabej (2005) referred to in a collected writing about molecular genetics of DCM (Stabej *et al.* 2006), the *titin* gene (TTN) has been genotyped in affected and unaffected dogs. The affected group displayed a variety of haplotypes, whereas the unaffected group mostly shared one haplotype. Stabej (2005) suggested, by these results, that the *titin* allele, which is common in the unaffected dogs, might confer protection against DCM (Stabej *et al.* 2006).

5.4 Conclusions and future studies

The quality of gDNA prepared from paraffin-embedded tissue differs between blocks and gDNA from some blocks seems to be highly degraded. This study still shows that paraffin blocks can be used as source for gDNA in genetic studies. The paraffin purification and DNA extraction method used in this study was considered as a potential preparation method, as some samples were able to PCR-amplify a 600 bp fragment. Most samples were able to amplify samples around 100-200 bp, which make them useful for microsatellite analyses with fragments around that size.

To conclude the result of the candidate gene approach in this study, which indicated no association between *desmin* and DCM, at least one more informative marker should be involved. My suggestion is to use the known microsatellite, Des1, and design new primers for a smaller fragment that can be amplified in a sufficient number of samples to get complete haplotypes. All samples should also be tested at least twice, to confirm the result.

The samples that have not been able to amplify a 600 bp fragment should not be excluded from this project. Other purification steps or individual optimization of PCR might improve their success in PCR. Other extraction methods, as mentioned earlier, could also be tested. Samples, which can amplify a 600 bp fragment, can possibly be used in whole genome association mapping with the Illumina methodology or similar

approaches. The other samples, of poorer quality, can then be used for fine mapping if a suspected associated region is found.

6. Acknowledgements

First I would like to thank my supervisors Göran Andersson, Izabella Baranowska and Nicolette Salmon Hillbertz for all help and support, and Katarina Stenshamn, who has been a very good collaborator and has given me a lot of support. I would also like to thank Anna Tidholm, who has collected all paraffin blocks used in this study and has read the manuscript for this report.

Further I want to thank the whole department of Animal Breeding and Genetics at SLU, in particular Ulla Gustafson for sequencing and advice.

7. Svensk sammanfattning

Syftet med DCM-projektet är att hitta en genetisk koppling till hjärtsjukdomen dilaterad cardiomyopati (DCM) hos hundar av rasen Newfoundland. Den del som presenteras i denna rapport innefattar en utvärdering av DNA från paraffininbäddad vävnad och en kandidatgensundersökning där association mellan genen *desmin* och sjukdomen DCM utvärderas.

DCM är en hjärtsjukdom som drabbar bland annat hundar och människor. Hjärtmuskeln blir förstorad eftersom hjärtat blir utvidgat och orkar inte pumpa ordentligt. Infektion kan vara en orsak till utveckling av DCM, men även andra orsaker finns. Sjukdomen kan i vissa fall ärvas och hos människa har man sett att det är en heterogen sjukdom med flera kända orsakande gener. Hos hund är sjukdomen homogen inom en ras och detta tillsammans med att stora regioner i hundens genom nedärvs tillsammans, vilket kallas ”linkage disequilibrium” (LD), gör att de passar bra för genetiska studier. Markörer i genomet jämförs mellan individer för att se vilka regioner som är lika i olika individer.

Vävnad inbäddad i paraffin renades från paraffin och sedan extraherades DNA. För att utvärdera kvalitén på DNA gjordes olika tester i PCR utan eller med ytterligare reningssteg såsom fenol/kloroform-rening och spinkolonner. Alla tester resulterade i att kvalitén på DNA från paraffininbäddad vävnad inte är så bra och att det är stor skillnad mellan olika klossar. Dock fanns inget tydligt samband mellan vilka klossar som gav bra eller dåligt DNA, men en möjlig förklaring kan vara att det gått olika lång tid från det att vävnaden tagits från en patient tills det att den bäddats in i paraffin. Kvalitén var ändå tillräckligt bra för att amplifiera en del fragment med mikrosatelliter, som är repetitioner i genomet som används som genetiska markörer. Markörerna låg runt genen *desmin* och genom att jämföra markörerna i sjuka och friska hundar kunde ingen association mellan genen och sjukdomen hittas. Dock utgjordes haplotypen endast av två markörer varav en som inte var så informativ, d.v.s. det var samma allel i de flesta individerna. En ny studie med kortare fragment med mikrosatelliter eller andra DNA-prover skulle behöva göras.

8. References

8.1 Literature

Akalu A., Reichardt J.K.V. (1999) ***A reliable PCR amplification method for microdissected tumor cells obtained from paraffin-embedded tissue***. Genetic Analysis: Biomolecular Engineering 15: 229-233

Alroy J., Rush J.E., Freeman L., Amarendhra Kumar M.S., Karuri A., Chase K., Sarkar S. (2000) ***Inherited infatile dilated cardiomyopathy in dogs: Genetic, clinical, biochemical, and morphologic findings***. American Journal of Medicine Genetics 95, 57-66

Brown T.A. (2002) ***Chapter 2: "Genome anatomies"*** & ***Chapter 5: "Mapping genomes"***. Genomes 2:nd ed., BIOS Scientific Publishers Ltd, Oxford

Coombs N.J., Gough A.C., Primrose J.N. (1999) ***Optimisation of DNA and RNA extraction from archival formalin-fixed tissue***. Nucleic Acids Research, vol. 27, no. 16, e12

Davies M.J. (1984) ***The cardiomyopathies: a review of terminology, pathology and pathogenesis***. Histopathology 8:363-393

Fatkin D., Graham R.M. (2002) ***Molecular Mechanisms of Inherited Cardiomyopathies.*** Physiol Rev. 82, 945-980

Franz W-M., Müller O.J., Katus H.A. (2001) ***Cardiomyopathies: from genetics to the prospect of treatment.*** The Lancet vol. 358, 1627-1637

Gibson G., Muse S.V. (2004) ***A primer of genome science. Second edition.*** Sinauer Associates, Inc Publishers. Massachusetts, USA.

Goebel H.H. (1995) ***Desmin-related neuromuscular disorders***. Muscle Nerve 18:1306-1320

Isola J., DeVries S., Chu L., Ghazvini S., Waldman F. (1994) ***Analysis of changes in DNA sequence copy number by comparative genomic hybridization in archival paraffin-embedded tumor samples***. The American Journal of Pathology 145(6): 1301-1308

Karlsson E.K., Salmon Hillbertz N., *et.al* (2007) ***Two-stage association mapping in dogs identifies coat color locus***. In preparation

Kumar V., Contran R.S., Robbins S.L. (2003) ***Chapter 11: "The heart"*** Robbins Basic Pathology 7th edition, Saunders, Philadelphia

Lazarides E. (1982) ***Intermediate filaments: A chemically heterogeneous, developmentally regulated class of proteins***. Annual Reviews of Biochemistry 51:219-50

Li D., Tapscoft T., Gonzalez O., Burch P.E., Quiñones M.A., Zoghbi W.A., Hill R., Bachinski L.L., Mann D.L., Roberts R. (1999) ***Desmin Mutation Responsible for Idiopathic Dilated Cardiomyopathy***. Circulation 100:461-464

Milner D.J., Weitzer G., Tran D., Bradley A., Capetanaki Y. (1996) ***Disruption of muscle architecture and myocardial degeneration in mice lacking desmin*** J. Cell Biol (Sep; 134 (5): 1255-70

Milner D.J., Taffet G.E., Wang X., Pham T., Tamura T., Hartley C., Gerdes A.M., Capetanaki Y. (1999) ***The absence of desmin leads to cardiomyocyte hypertrophy and cardiac dilation with compromised systolic function***. J Mol Cell Cardiol 31: 2063-2076

Page R.D.M., Holmes E.C. (1998) ***Chapter 3: "Genes: Organisation, Function and Evolution"***. Molecular Evolution –A Phylogenetic Approach, Blackwell Science Ltd, UK

Scheinin S., Capek P., Radovancevic B., Duncan J.M., McAllister H.A. Jr., Frazier O.H. (1992) ***The effect of prolonged left ventricular support on myocardial histopathology in patients with end-stage cardiomyopathy***. ASAIO Journal 38:M271-M274

Schuelke M. (2000) ***An economic method for the fluorescent labeling of PCR fragments***. Nature Biotechnology, vol. 18, Feb. 2000

Shi S-R., Cote R.J., Wu L., Liu C., Datar R., Shi Y., Liu D., Lim H., Taylor C.R. (2002) ***DNA Extraction from Archival Formalin-fixed, Paraffin-embedded Tissue Sections Based on the Antigen Retrieval Principle: Heating Under the Influence of pH***. The Journal of Histochemistry & Cytochemistry, vol 50(8): 1005-1011

Stabej P., Imholz S., Versteeg S.A., Zijlstra C., Stokhof A.A., Domanjko-Petric A., Leegwater P.A.J., van Oost B.A. (2004) ***Characterization of the canine desmin (DES) gene and evaluation as a candidate gene for dilated cardiomyopathy in Dobermann*** Gene 340,241-249

Stabej P. (2005) ***"Molecular genetics of dilated cardiomyopathy in Dobermann dog"***. Ph.D. thesis, Utrecht University, The Netherlands (referred to in Stabej *et al.* 2006)

Stabej P., Meurs K.M., van Oost B.A. (2006) ***Molecular Genetics of Dilated Cardiomyopathy*** The Dog and Its Genome, Cold Spring Harbor Laboratory Press 0-87969-742-3, 365-382

Sutter N.B., Eberle M.A., Parker H.G., Pullar B.J., Kirkness E.F., Kruglyak L., Ostrander E.A. (2004) ***Extensive and breed-specific linkage disequilibrium in Canis familiaris***. Genome Res. 14:2388-2396

Tidholm A. and Jönsson (1996) ***Dilated Cardiomyopathy in the Newfoundland: A Study of 37 Cases (1983-1994)***. J Am Anim Hosp Assoc 32:465-470

Tidholm A., Jönsson L. (1997) ***A retrospective study of canine dilated cardiomyopathy (189 cases)***. J. Am. Anim. Hosp. Assoc. 33:544-550

Tidholm A., Häggström J., Jönsson L. (1998) ***Prevalence of attenuated wavy fibers in the myocardium of dogs with dilated cardiomyopathy*** Journal of American Veterinary MedicalAssociation 212:1732-1734

Tidholm A., Häggström J., Jönsson L. (2000) ***Detection of attenuated wavy fibers in the myocardium of Newfoundlands without clinical echocardiographic evidence of heart disease***. Am J Vet Res, vol. 61, No. 3: 238-241

Tidholm A. and Jönsson L. (2005) ***Histologic Characterization of Canine Dilated Cardiomyopathy***. Vet. Pathol. 42:1-8

VWR bioMarke, the Market Source for Life Science (2006) ***High quality Genomic DNA from Animal Tissues and Using E.Z.N.A.$^{(R)}$ Tissue DNA Kits***. VWR International, issue 15 (fall)

Zhenlin Li, Alain Lilienhaum, Gillian Butler-Browneb and Denise Paulin ***Human desmin-coding gene: complete nucleotide sequence, characterization and regulation of expression during myogenesis and development*** *(Recombinant DNA; intermediate filaments; intron; exon; hamster; l phage vectors; promoter; actin; lariat; Al141 repeat; restriction map; vimentin)* Gene; 78 (1989) 243-254

8.2 Manuals

BD CHROMA SPIN^{TM} Columns User Manual, no. PT1300-1, ver. PR49719, BD Biosciences -Clontech

NanoDrop manual, ND-1000 Spectrophotometer V3.3 User's Manual

8.3 Internet references

Ensembl: ***http://www.ensembl.org***
Ensembl Dog (CanFam1): release 40, Aug 2006 and release 41, Oct 2006
Ensembl Pre Dog (CanFam2): v.38
Ensembl Dog (CanFam2): release 42, Dec 2006

NetPrimer: ***www.premierbiosoft.com/netprimer/index.html***

NCBI (Pubmed): ***http://www.ncbi.nlm.nih.gov***

Primer3: ***http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi***

TAG Copenhagen: ***http://www.tagc.com***

UCSC Genome Browser: ***http://genome.ucsc.edu***
Dog Assembly May 2005

Appendix 1. Sample list, with tissue weight, NanoDrop result and result of 600 bp PCR.

ID	from year	Case/Control	Sex	gram paraffin sample	ug DNA in solution	260/280	600 bp fragment in PCR	Yield Comparison*	
NF18	1994	control	Female	0,0174	66,2	1,86	No		
NF19	1995	case	Female				-		Missing block
NF20	2004	control	Male	0,0152	10,5	1,93	No		
NF21	2003	control		0,0203	34,1	1,96	No	1680	
NF22	1997	control	Male	-	71,1	1,89	No		
NF23	1997	case	Male	0,0190	27,3	1,91	Yes	1437	
NF24	1998	case	Female	0,0292	61,5	1,85	No		
NF25	1999	case	Male	0,0219	23,7	2,01	Yes	1082	
NF26	1995	case	Female	0,0130	33,9	1,86	No		
NF27	*only blood*	*case*	*Female*				-		No block
NF66	1993	case	Female	0,0310	22,7	1,89	No	732	
NF67	1999	control	Female	0,0220	16,6	1,76	No		
NF72	*only blood*	*case*					-		No block
NF73	2000	case	Female	0,0155	23,1	1,91	No		
NF74	1997	case	Male	0,0142	20,9	1,94	Yes	1472	
NF75	1996	case	Male	0,0270	27,8	1,91	No		
NF76	1996	case	Female	0,0212	32,4	1,97	No	1528	
NF77	1997	case	Female	-	44,9	1,87	No		
NF78	1997	case		0,0224	30,7	1,92	No		
NF79	1996	case		0,0181	31	1,95	No		
NF80	1999	case		0,0260	31,1	1,97	Yes	1196	
NF81	1995	case		0,0197	91,3	2,01	No		
NF82	1997	case	Female	0,0132	16,5	1,93	No		
NF83	1993	case		0,0110	26,4	1,88	No	2400	
NF84	1993	case	Male	-	56,9	1,82	No		
NF85	1993	case	Female	0,0183	24,9	1,76	No		
NF86	1992	case		0,0226	37,8	1,92	No		
NF87	1999	case		0,0233	8,4	1,79	Yes	361	
NF88	1994	case		0,0219	21,1	1,89	No		
NF89	1996	case		0,0170	20,1	1,94	No		
NF90	1994	case		0,0302	10,4	1,84	No		
NF91	1991	case		0,0255	22,8	1,92	No		
NF92	1995	case		0,0189	22,6	1,91	No	1196	
NF93	1995	case		0,0236	16,9	1,94	No		
NF94	1995	case		0,0256	67	1,93	Yes	2617	
NF95	1992	case	Female	-	46,5	1,88	No		
NF96	1992	case	Female	0,0140	21,6	1,96	No	1543	
NF97	1996	case	Male	0,0216	24,2	1,88	No		
NF98	2001	case	Male	0,0277	35,1	1,98	No		
NF99	2000	case		0,0195	14,9	1,72	No		
NF100	1998	case		-	31,1	1,92	Yes		
NF101	1989	case	Female	0,0157	18,3	1,96	No		
NF102	1993	case		0,0189	39,4	1,92	No	2085	
NF103	2001	case	Female	0,0158	25,3	1,94	No		
NF104	1996	case		0,0175	1,2	1,87	No		
NF105	1992	case	Male	0,0150	19,5	1,90	No		
NF106	1991	case		0,0188	42,3	1,94	No		
NF107	1991	case	Female	0,0250	65	1,99	No		
NF108	1999	case					-		Missing block
NF109	1998	case	Female	-	229,4	1,91	Yes		
NF110	1998	case	Female	0,0263	61	1,98	No		
NF111	1996	case		0,0142	12,6	1,91	No		
NF112	1996	case	Female				-		Missing block
NF113	1982	control		0,0220	27,6	1,87	No		
NF114	2001	control	Female	0,0135	37,8	2,01	Yes	2800	
NF115	2001	control		0,0222	39,2	1,91	Yes	1766	
NF116	2001	control	Female	0,0297	20,2	1,79	No		
NF117	1998	control		0,0274	19,3	1,85	No	704	
NF118	1998?	control		0,0193	37,9	1,92	Yes	1964	
NF119	1994	control		0,0235	17,8	1,91	No		
NF120	1991	control		0,0260	42,4	1,87	No		
NF121	1989	control		0,0187	23,3	1,90	No		
NF122	1992	control		0,0247	23,1	1,76	No		
NF123	1993	control		0,0158	22,4	1,92	No	1418	
NF124	1993	control		0,0131	29,5	1,90	No		
NF125	1990	control		0,0201	20,2	1,95	No		
NF126	1991	control		0,0303	116,8	2,00	No		
NF127	1992	control		0,0186	29,9	1,97	No		
NF128	1995	control		0,0173	205,5	2,02	No	11879	
NF129	1992	control		0,0177	20,2	1,83	No		
NF130	1994	control		0,0256	49,4	1,90	No		
NF131	1992	control		0,0211	25,2	1,85	No		
NF132	1995	control		0,0222	51,2	1,93	No		
NF133	2002	control		0,0271	25,2	1,72	Yes	930	
NF134	2002	control		0,0218	18,1	1,82	No		
NF135	2000	control		0,0263	35,2	1,93	No	1338	
NF136	2000	control		0,0229	57,2	1,98	Yes	2498	

* see figure in appendix 9

Appendix 2. Paraffin extraction protocols

a) Paraffin Extraction via Octane/Methanol

Reagents: *n-Octane*
100% Methanol

Harmful reagents,
use gloves and work in a hood.

1. Add 1 ml of octane to the sample in a 1.5-2 ml microcentrifuge tube.

2. Vortex vigorously for 10 sec (or until paraffin has detached from sample).

3. Add 100 µl of 100% Methanol.

4. Vortex vigorously.

5. Centrifuge at 10.000 rpm for 2 min. (Octane forms upper layer, and methanol with tissue forms lower layer).

6. Remove upper octane layer with fine tip transfer pipette.

7. Spin for 1 min at 10.000 rpm.

8. Remove residual octane layer.

9. Remove methanol that pellets with the tissue (let stand to dry;
CAUTION: Do not over dry as DNA may denature).

10. Start DNA extraction protocol

b) Xylene to extract paraffin
(suggested in E.Z.N.A. tissue DNA kit)

Xylene is harmful,
use gloves and work in a hood.

1. Put 30 mg tissue (about $2mm^3$) in a microcentrifuge tube.

2. Extract the paraffin by adding 1 ml Xylene. Mix carefully with vortex.

3. Centrifuge the tube at 10.000 x g (max. speed) for 10 minutes in room temperature. Throw away the supernatant (harmful waste) without touching the pellet.

4. Wash the pellet with 1 ml absolute ethanol to remove traces of Xylene. Centrifuge at 10.000 x g for 5 min in room temperature. Throw away the ethanol without touching the pellet.

5. Repeat the ethanol wash.

6. Air-dry the pellet in 37^oC for 15 min.

7. Start DNA extraction protocol or E.Z.N.A.-kit.

Appendix 3. Standard (salt) protocol for DNA extraction

DNA-preparation from muscle tissue

1. Mix
* about 50 mg tissue (homogenize)
* 300 µl prep- buffer
* 7,5 µl Proteinase K (stock 8mg/ml).

2. Incubate in 50^{o}C over night.

3. Add 80 µl saturated NaCl, vortex, centrifuge at 9000 rpm for 10 min.
Transfer the supernatant to a new tube.

4. Repeat step 3 until the solution is clear (3-4 times).

5. Add 800 µl 95% EtOH, turn the tube around, centrifuge at 13.000 rpm at least 45 min. Throw away the supernatant and let the pellet dry (37^{o}C).

5b. (added in preparation of all samples) Wash the pellet with 500 µl 70% EtOH (turn around the tube). Centrifuge at 13.000 rpm for 10 min and throw away the supernatant. (This removes traces of NaCl, as the 30% water in the solution binds the salt.)

6. Dissolve the pellet in 50 µl-100 µl 1 x TE.

Prep-buffer (for one sample = 300µl)
6 µl 5M NaCl
3 µl 1M Tris-HCl
15 µl 10% SDS
276 µl water

Appendix 4. PCR protocols

250bp- and 600bp- PCR (original version)
(600bp version were modified in effort to improve amplification)

Mix		**Program**		
	1 well			
PCR Buffer II (x10)	2µl	94°C	10 min	
$MgCl_2$ (25mM)	1.6 µl	94°C	30 sec	} 16 cycles
dNTP (20mM)	0.2 µl	69-54°C	30 sec	
DMSO	1 µl	72°C	45 sec	
Primer F (10 ng/µl)	0.6 µl			
Primer R (10 ng/µl)	0.6 µl	94°C	30 sec	} 38 cycles
AmpliTag Gold (5U/µl)	0.15 µl	54°C	30 sec	
H_2O	11.85 µl	72°C	45 sec	
Transfer	*18 µl/tube*			
		72°C	5 min	
		4°C	∞	
gDNA (25 ng/µl)	2 µl			

850bp- PCR (Des_ex456 and Des_ex7) (original version) (modified in some tests)

Mix		**Program**		
	1 well			
PCR Buffer II (x10)	2.5 µl	94°C	5 min	
$MgCl_2$ (25mM)	2.0 µl			
dNTP (20mM)	0.25 µl	94°C	40 sec	} 40 cycles
Primer F (10 ng/µl)	0.5 µl	56°C	40 sec	
Primer R (10 ng/µl)	0.5 µl	72°C	1 min	
AmpliTaq Gold (5U/µl)	0.20 µl			
H_2O	17 µl	72°C	5 min	
Transfer	*23 µl/tube*	4°C	∞	
gDNA (25 ng/µl)	2 µl			

Microsatellite PCR (Des1, Des2 and Des3)

Mix		**Program**		
	1 well			
PCR Buffer II (x10)	1.25 µl	94°C	5 min	
$MgCl_2$ (25mM)	1.25 µl	94°C	30 sec	} 10 cycles
dNTP (20mM)	0.125 µl	54°C	30 sec	
Primer F (10 ng/µl) (M13 (-21))	0.5 µl	72°C	30 sec	
M13 (-21) Primer (20ng/µl)	0.19 µl			
Primer R (10 ng/µl)	0.375 µl	94°C	30 sec	} 30 cycles
AmpliTaq Gold (5U/µl)	0.10 µl	52°C	30 sec	
H_2O	8.71 µl	72°C	30 sec	
Transfer	*12.5 µl/tube*			
		72°C	15 min	
		4°C	∞	
gDNA (25 ng/µl)	2 µl			

Appendix 5. Purification with phenol/chloroform (modified part of a DNA extraction protocol)

Phenol/chloroform protocol

1. Add an equal volume (as DNA sample) water saturated phenol/chloroform to the sample (removes protein and fat). Vortex for about 1 minute.
2. Centrifuge at 13.000 rpm in room temperature for 2 minutes.
3. Transfer the upper phase (containing water and DNA) to a new eppendorf tube. Measure (or estimate) the volume.
4. Add an equal volume chloroform+TE (remove phenol traces). Vortex and centrifuge as in step 2.
5. Transfer the upper phase (water and DNA) to a new eppendorf tube. Measure (or estimate) the volume.
6. Add 1/10 volume NaAC and 3 volumes ice cold absolute ethanol to precipitate the DNA.
7. Turn the samples a couple of times and let stand in -20^{o}C freezer for 2 hours.
8. Centrifuge at 13.000 rpm for 10 minutes.
9. Throw away the supernatant.
10. Wash the pellet with 70 % ethanol (same volume as the absolute ethanol earlier). Centrifuge at 13.000 rpm for 10 minutes.
11. Throw away the supernatant and let the pellet air-dry.
12. Dissolve the pellet in 50 µl-100 µl 1 x TE.

Appendix 6. Measurements for yield calculations. E.Z.N.A. samples were eluted twice (in elution buffer and water). DNA amount in both elutions are used for total yield calculation.

ID	New ID	Tissue weight (mg)	DNA-prep	Paraffin purification	DNA conc. (mean of three) (ng/uL)	DNA in sample (ng)	Yield (ng DNA/mg tissue)	Total Yield
Evaluation sample (NF115)	S.oct.1	12,1	Salt prep.	Octane/methanol	357,3	17865	1476,4	1466,4
Evaluation sample (NF115)	S.oct.2	14,6	Salt prep.	Octane/methanol	216,8	10840	742,5	742,5
Evaluation sample (NF115)	S.xyl.1	18,6	Salt prep.	Xylene	287,0	14350	771,5	771,5
Evaluation sample (NF115)	S.xyl.2	18,1	Salt prep.	Xylene	497,8	24890	1375,1	1375,1
Evaluation sample (NF115)	EZ.oct.wat.1	25,6	EZNA:1 (eluted with water)	Octane/methanol	94,5	9454	369,3	670,5
	EZ.oct.eb.1		EZNA:1 (second elution with elution buffer)		77,1	7710	301,2	
Evaluation sample (NF115)	EZ.oct.wat.2	13,5	EZNA:2 (eluated with water)	Octane/methanol	50,9	5086	376,7	641,0
	EZ.oct.eb.2		EZNA:2 (second elution with elution buffer)		35,7	3567	264,2	
Evaluation sample (NF115)	EZ.xyl.wat.1	25,2	EZNA:3 (eluted with water)	Xylene	58,7	5873	233,1	401,3
	EZ.xyl.eb.1		EZNA:3 (second elution with elution buffer)		42,4	4240	168,3	
Evaluation sample (NF115)	EZ.xyl.wat.2	11,0	EZNA:4 (eluted with water)	Xylene	91,1	9109	828,1	1098,1
	EZ.xyl.eb.2		EZNA:4 (second elution with elution buffer)		29,7	2970	270,0	
NF113	-	16,6	Salt prep.	Xylene	480,9	24044	1448,4	1448,4
NF116	-	19,0	Salt prep.	Xylene	294,1	14703	773,8	773,8
NF118	-	15,5	Salt prep.	Xylene	288,4	14422	930,4	930,4
NF122	-	15,3	Salt prep.	Xylene	396,9	19844	1297,0	1297
NF131	-	12,7	Salt prep.	Xylene	109,2	5459	429,8	429,8
NF133	-	14,8	Salt prep.	Xylene	129,9	6497	439,0	439
NF134	-	22,0	Salt prep.	Xylene	122,1	6106	277,5	277,5
NF135	-	17,8	Salt prep.	Xylene	198,1	9907	556,5	556,5

Appendix 7. The samples with PCR result sorted according to age of paraffin block.

Year	ID	600 bp	~300 bp (Des1)	~250 bp (Des2)	~100 bp (Des3)
1982	NF113			ok	ok
1989	NF101				
1989	NF121				
1990	NF125			ok	
1991	NF91			ok	ok
1991	NF106				
1991	NF107				ok
1991	NF120				ok
1991	NF126			ok	
1992	NF86			ok	ok
1992	NF95				ok
1992	NF96				ok
1992	NF105				ok
1992	NF122				
1992	NF127			ok	ok
1992	NF129				
1992	NF131				
1993	NF66				
1993	NF83				
1993	NF84				
1993	NF85				ok
1993	NF102				
1993	NF123				
1993	NF124				
1994	NF88				ok
1994	NF90				
1994	NF18				
1994	NF119			ok	ok
1994	NF130			ok	ok
1995	NF26				
1995	NF81				
1995	NF92				
1995	NF93				
1995	NF94	ok		ok	ok
1995	NF128			ok	ok
1995	NF132			ok	ok
1996	NF75			ok	ok
1996	NF76		ok	ok	
1996	NF79				
1996	NF89			ok	ok
1996	NF97				ok
1996	NF104				
1996	NF111				
1997	NF23	ok		ok	ok
1997	NF74	ok		ok	ok
1997	NF77				
1997	NF78			ok	ok
1997	NF82			ok	
1997	NF22			ok	ok
1998?	NF118	ok	ok	ok	ok
1998	NF24			ok	ok
1998	NF100	ok	ok	ok	ok
1998	NF109	ok	ok	ok	
1998	NF110	ok	ok	ok	
1998	NF117				ok
1999	NF25	ok	ok	ok	ok
1999	NF80	ok		ok	ok
1999	NF87	ok		ok	ok
1999	NF67				
2000	NF73			ok	ok
2000	NF99			ok	
2000	NF135				
2000	NF136	ok	ok	ok	ok
2001	NF98			ok	ok
2001	NF103			ok	ok
2001	NF114	ok	ok	ok	ok
2001	NF115	ok	ok	ok	ok
2001	NF116			ok	
2002	NF133	ok		ok	ok
2002	NF134				
2003	NF21			ok	ok
2004	NF20				

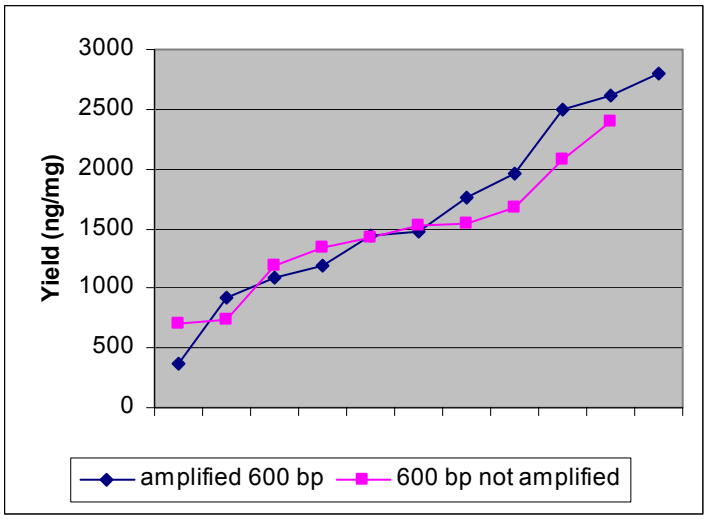
Appendix 8. Alleles and haplotypes for all tested individuals.
Question mark means uncertain haplotype (one allele is missing) and allele size in bold text means that it is measured twice.

	ID	Des1	Des1	Des2	Des2	Des3	Des3	Des2 & 3-haplotype	
Cases:	NF23			252	252	125	125	**A**	
	NF24			252	252	125	125	**A**	
	NF25	320	320	252	252	161	161	**C**	
	NF26								
	NF66								
	NF73			252	252	125	125	**A**	
	NF74			252	252	145	169	**B**	**D**
	NF75			252	252	145	169	**B**	**D**
	NF76	305	320	252	252				
	NF77								
	NF78			252	252	125	125	**A**	
	NF79								
	NF80			252	252	125	161	**A**	**C**
	NF81								
	NF82			252	252				
	NF83								
	NF84								
	NF85					125	125		
	NF86			252	252	125	161	**A**	**C**
	NF87			252	252	125	161	**A**	**C**
	NF88					125	125		
	NF89			**252**	**252**	125	125	**A**	
	NF90								
	NF91			252	252	125	161	**A**	**C**
	NF92								
	NF93								
	NF94			**248**	**252**	125	125	**E**	**A**
	NF95					125	125		
	NF96					161	169	C?	D?
	NF97					125	161		
	NF98			252	252	**161**	**161**	**C**	
	NF99			252	252				
	NF100	**320**	**320**	252	252	125	125	**A**	
	NF101								
	NF102								
	NF103			252	252	125	125	**A**	
	NF104								
	NF105					125	125		
	NF106								
	NF107					125	161		
	NF109	**314**	**320**	**248**	**252**				
	NF110	320	320	252	252				
	NF111								
Controls:	NF18								
	NF20								
	NF21			252	252	125	161	**A**	**C**
	NF22			**252**	**252**	145	161	**B**	**C**
	NF67								
	NF113			252	252	125	125	**A**	
	NF114	**305**	**320**	252	252	125	125	**A**	
	NF115	**305**	**320**	252	252	125	125	**A**	
	NF116			**252**	**252**				
	NF117					125	161		
	NF118	320	320	252	252	**161**	**169**	**C**	**D**
	NF119			252	252	125	125	**A**	
	NF120					125	161		
	NF121								
	NF122								
	NF123								
	NF124								
	NF125			252	252				
	NF126			**252**	**252**				
	NF127			252	252	125	161	**A**	**C**
	NF128			252	252	125	161	**A**	**C**
	NF129								
	NF130			**252**	**252**	125	125	**A**	
	NF131								
	NF132			252	252	125	125	**A**	
	NF133			252	252	125	125	**A**	
	NF134								
	NF135								
	NF136	305	320	252	252	125	125	**A**	

Microsat. 2 & 3-haplotypes:	
A:	-.252.125
B:	-.252.145
C:	-.252.161
D:	-.252.169
E:	*-.248.125*

Appendix 9. Comparison of yield from samples that could and could not amplify a 600 bp fragment.

	600bp	not 600bp
Yield:	361	704
(ng DNA /	930	732
mg tissue)	1082	1196
	1196	1338
	1437	1418
	1472	1528
	1766	1543
	1964	1680
	2498	2085
	2617	2400
	2800	11879
	18123	26503



The highest value for "600 bp not amplified" is above the scale.