The Genetic Basis of Canine Osteoarthritis

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy

by

Dylan Neil Clements

University of Liverpool

July 2007

Table of Contents

Declaration

No part of the work referred to in this thesis has been submitted in support of and application for another degree or qualification, or at any other University or Institute of Learning, with the exception of part of Chapter 5, where the data for the part of the experiments were submitted in support of the Master of Science degree obtained by Ms Lindsey Maccoux, and awarded by the University of Manchester. The differentiation of the contributions of each author is stated at the end of the manuscript.

Acknowledgements

This project was funded by the Biotechnology and Biological Sciences Research Council (student stipend) and the PetPlan Charitable Trust (consumables funding). As the recipient I am grateful to both for their generosity and vision in supporting canine genetic research. I would like to thank my supervisors, Professors Stuart Carter, Bill Ollier and John Innes for their help and advice throughout the project, and for giving me the opportunity to perform this work. Much of the work presented in this thesis was driven through advice and mentoring of Dr Philip Day, to whom I am greatly indebted for his technical assistance, inspiration and friendship.

Thanks to Ash Dunne, Shona Wood and Andrea Short for their help and friendship and I wish each of them every success in their future careers. I am also grateful for the healthy doses of reality and humour provided by Annette Barnes which were a welcome relief from the frustrations of failed experiments. Thanks also to Sue Bell and Di Isherwood for their help with the project and making me feel welcome at Liverpool. The study also could not have been completed without the provision of samples by surgeons affiliated to this project; John Ferguson, Steve Clarke, Mike Guilliard, Steve Butterworth, Malcolm Ness, Malcolm McKee, Toby Gemmill, Carlos Macias, Andrew Coughlan, Christopher May, Duncan Midgley are gratefully acknowledged in this regard.

I would like to thank my parents for giving me the opportunities in life, and the self determination to succeed which have lead me to this point in my career. Finally and most importantly, thanks to Jenny for her love, support, resilience and resolve throughout the project and without whom this work would never have been started let alone completed. This thesis is dedicated to her.

Structured Abstract

Osteoarthritis (OA) is a common debilitating disease of mammalian joints. Canine OA was classically understood to arise secondary to articular diseases, such as elbow dysplasia, hip dysplasia and cranial cruciate ligament rupture which produced a mechanical dysfunction of the affected joint. However it is now recognised that primary factors, such as genetics, govern the severity of OA for a given articular disease. In this study, two different aspects of the genetic basis to OA were investigated; gene transcription in diseased tissues and gene polymorphism frequencies in populations of dogs with diseases predisposing to OA.

Evaluation of the quality of extracted mRNA from canine joint tissues by microfluidic electrophoresis traces revealed that there were no differences in the quality of samples extracted using either an isopropanol or ethanol precipitation method. However, a significant proportion of RNA samples (32%) were identified as degraded, highlighting the importance of assessing RNA quality before usage.

In OA canine hip cartilage, there was an increase in the gene expression of structural matrix molecules (collagens and small leucine rich proteoglycans) and proteinases (matrix metalloproteinase 13, cathepsin -B and -D), with concurrent decreased expression of selected inhibitors or protease activity (tissue inhibitors of metalloproteinase-2 and -4) when compared to normal articular cartilage using quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) analyses. The general pattern of changes in matrix-associated gene expression was similar to that reported in naturally occurring human OA cartilage. Canine-specific oligonucleotide microarray gene expression profiling of a small sample set of normal and OA articular cartilage samples identified differential expression of a number of genes not previously associated with the disease. However, the high degree of heterogeneity observed in the expression profile data generated hampered subsequent data interpretation, and highlighted the limitations of expression profiling small sample sets with limited phenotype stringency.

Quantification of matrix-associated gene expression in OA elbow cartilage by RT-qPCR identified changes which were consistent with those identified in end stage hip OA cartilage, and which correlated with the radiographic severity of elbow OA for a number of genes (such as type I collagen alpha two chain, type III collagen alpha one chain and tissue inhibitor of matrix metalloproteinase 2). Matrix metalloproteinase expression in OA elbow trabecular bone was also identified to be increased when compared to normal trabecular bone, when quantified by RT-qPCR.

A general pattern of increased protease and extracellular structural matrix gene expression was identified in ruptured canine cranial cruciate ligaments (CCL) when compared to intact CCLs, with both RT-qPCR and oligonucleotide microarray. No significant differences were identified between the gene expression profiles of normal CCLs of a breed predisposed to CCL rupture (Labrador Retriever) when compared to a breed relatively resistant to CCL rupture (Greyhound), although a degree of risk-specific clustering was observed for expression profiles of genes which were

differentially expressed in CCL rupture. The expression profiles of ruptured canine CCLs were similar to those previously reported for ruptured human CCLs. A transcriptomic basis to breed specific risk for the development of canine CCL rupture was not identified.

Microarray data sets generated from normal and OA canine articular cartilage and normal and ruptured CCL were filtered to identify new reference genes for use in RT-qPCR experiments. One of the new reference genes (Mitochondrial ribosomal protein S7 [MRPS7]) demonstrated a high degree of stability across multiple articular tissues from normal and OA canine joints, as determined by multiple, different reference gene stability assessment algorithms, making it a potential universal reference gene for use in canine OA tissue studies.

Silica membrane spin columns provided the most consistent recovery of high quantities of genomic DNA (gDNA) from EDTA preserved and clotted blood samples without the co-extraction of PCR inhibitors, when compared to phenol-chloroform or modified salt precipitation methods of DNA extraction. Spectrophotometer quantification of extracted gDNA did not provide an accurate assessment of the functional gDNA quantity with phenol-chloroform extracted samples, because of protein contamination.

Single nucleotide polymorphisms (SNPs) were identified in twenty candidate genes and their allele frequencies evaluated in populations of Labrador Retrievers and Golden Retrievers with cruciate ligament disease, populations of Labrador Retrievers with elbow dysplasia and hip dysplasia and compared to general populations of Labradors Retrievers and Golden Retrievers. Significant associations were identified for the minor allele and haplotype frequencies of SNPs in interleukin 12B (*IL12B*) in Labrador Retrievers with elbow dysplasia, interleukin 4 (*IL4*) and interleukin 6 (*IL6*) in Labrador Retrievers with hip dysplasia, *IL4* and *IL12B* in Labrador Retrievers with cranial cruciate ligament rupture, and interleukin 10 (*IL10*) and Ankyrin repeat domain 10 (*ANKRD10*) in Golden Retrievers with cranial cruciate ligament rupture. A common genomic risk for the articular disease, or OA, was not identified between the two different breeds of dog evaluated, but common genomic risks were identified for different articular diseases within a single breed. A genetic basis to canine articular disease, or OA, was confirmed.

Abbreviations

A_{260}:A_{230} ratio = 260 nm to 230 nm absorbance ratio
A_{260}:A_{280} ratio = 260 nm to 280 nm absorbance ratio
AACT = Alpha 1-antichymotrypsin
ACE = angiotensin converting enzyme
ACL = anterior cruciate ligament
ACTB = Beta Actin
ADAM= a disintegrin and metalloprotease domain
ADAMTS4 = ADAM metallopeptidase with thrombospondin type 1 motif, 4
ADAMTS5 = ADAM metallopeptidase with thrombospondin type 1 motif, 5
AGC1 = aggrecan
Alpha1-antichymotrypsin (AACT)
ANXA2 = annexin 2
aRNA = amplified RNA
ASPN = Asporin
ATIC = 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase / IMP cyclohydrolase
B2M = beta-2-microglobulin
BGN = biglycan
BMP = bone morphogenetic protein.
bp = base pairs
C3 = complement component 3
CALM1 = calmodulin
CASP8 = caspase 8
CCL = cranial cruciate ligament
CCLR = cranial cruciate ligament rupture
cDNA = complementary DNA
C7orf28B = CG14980-PB
CFA = canine chromosome
CI = confidence interval
CILP = Cartilage intermediate layer protein
COL1 = type I collagen
COL10A1 = type X collagen, alpha 1 chain
COL1A2 = type I collagen, alpha two chain
COL2 = type II collagen
COL2A1 = type II collagen, alpha 1 chain
COL3 = type III collagen
COL3A1 = type III collagen, alpha 1 chain
COL5A1 = type V collagen, alpha 1 chain
COL9A3 = type IX collagen, alpha 3 chain
COMP = Cartilage oligometric protein
COX = cyclooxygenase
c-myc = myelocytomatosis viral oncogene homolog (Avain)
CRTM = Matrillin
cRNA = complementary RNA
CSPG2 = chondroitin sulphate proteoglycan 2
CT = mean threshold cycle

CTSB = cathepsin B
CTSD = cathepsin D
DCN = decorin
DF = degradation factor
DNA = deoxyribonucleic Acid
ECM = extracellular matrix
ED = Elbow dysplasia
EDTA = ethylenediamine tetra acetic acid
ERA = oestrogen receptor alpha
FCP = fragmentation of the medial coronoid process
FDR = false discovery rate
FN1 = fibronectin
FRZB = frizzled-related protein B
GAPDH = glyceraldehyde-3-phosphate dehydrogenase
GBP = Gu binding protein
gDNA = genomic DNA
GR = Golden Retriever
H-ras = Harvey rat sarcoma viral oncogene homolog
HIRP5 = HIRA interacting protein 5 isoform 2
IGF = Insulin-like growth factor 1
IL = interleukin
IL1R = Interleukin 1 Receptor
IL1RA = Interleukin 1 receptor antagonist
IMP = IMP cyclohydrolase
IQR = interquartile ranges
kb = kilobase
LR = Labrador Retriever
LUM = lumican
MAPK6 = mitogen-activated protein kinase 6
mb = megabase
MHC = major histocompatibility complex
MMP = matrix metalloproteinase
MMP = matrix metalloproteinase
mRNA = messenger Ribonucleic Acid
MRPS25 = mitochondrial 28S ribosomal protein S25
MRPS7 = mitochondrial ribosomal protein S7
NCOR = nuclear receptor co-repressor 2
NCK2 = cytoplasmic protein NCK2
NOS = nitric oxide synthetase
OA = osteoarthritis
OPG / TNFR11B = Osteoprotegrin / tumour necrosis factor receptor 11B
ORMDL1 = ORM1-like 2
PAPSS2 = phosphoadenosine 5'-phosphosulphate synthetase
PCR = polymerase chain reaction
PIAS1 = Gu binding protein
PTDSS1 = phosphatidylserine synthase I
qPCR= quantitative polymerase chain reaction
QTL = quantitative trait loci
RANK = receptor activator of nuclear factor-κB
RANKL = receptor activator of nuclear factor-κB ligand

RIN = RNA integrity number
RNA = ribonucleic Acid
RPL13A = ribosomal protein L13a
RPM = revolutions per minute
RR = ribosomal peak ratio
RT-PCR = quantitative reverse-transcriptase polymerase chain reaction
SLC26A2 = solute carrier family 26, member 2
SDHA = succinate dehydrogenase complex, subunit A
SNP = single nucleotide polymorphism
SOX9 = SRY (sex determining region Y)-box 9
Spec I = Ultrospec UV Spectrophotometer
Spec II = Nanodrop UV spectrophotometer
TBP = TATA box binding protein
TGF = transforming growth factor
TIMP = tissue inhibitor of metalloproteinase
TKT = transketolase
TNA = Tetranectin
TNC = tenascin C
TNF = tumour necrosis factor
TNFIAP6 = tumour necrosis factor interacting protein 6
TRAPPC2L = hematopoietic stem/progenitor cells 176
VDR = Vitamin D Receptor
VIM = vimentin
VZV = varriella zoster virus

Introduction

Osteoarthritis (OA) is a condition characterised by the destruction of articular cartilage, resulting in pain and dysfunction of the affected joint. Over time articular cartilage degenerates with fibrillation, fissures, ulceration, and eventual full thickness loss of the joint surface. Outgrowths of bone at the margin of the affected joints appear in later life, which cause joint pain and stiffness. OA is now recognised as probably being a group of overlapping distinct diseases, which may have different aetiologies but with similar biologic, morphologic, and clinical outcomes. The disease processes can involve the entire joint, with synovial membrane (1), infrapatella fat (2), ligament (3) and subchondral bone (4,5) also being affected, although research in OA has concentrated predominately on the pathogenesis of articular cartilage destruction

At present OA is the most commonly observed non-traumatic orthopaedic condition of dogs in the United Kingdom (6). Over 20% of dogs older than 1 year of age are estimated to be affected by OA (7). The three most common conditions resulting in canine OA are canine hip dysplasia, canine elbow dysplasia and canine cranial cruciate ligament (CCL) rupture. Each of these conditons leads to OA of the hip, elbow or stifle of affected dogs respectively. Canine hip dysplasia was first recognised in 1935, and is now understood to be a developmental trait characterised by instability of the hip joint, which leads to hip subluxation (8). Canine elbow dysplasia is a generic term encompassing a number of well defined phenotypes of the cubical joint, such as fragmentation of the medial coronoid process (FCP) (9), osteochondrosis dissecans of the medial part of the humeral condyle (9) and ununited anconeal process

(10). Canine cranial cruciate ligament (CCL) disease is a chronic degenerative condition which results from the progressive pathological failure and rupture of the canine CCL (11).

Primary and Secondary Canine OA

Canine OA may develop as an idiopathic primary event, or secondary to an identifiable initiating cause (for example, secondary to hip laxity with hip dysplasia, secondary to fragmentation of the medial coronoid process with elbow dysplasia, or secondary to stifle laxity with cranial cruciate ligament rupture). The role of genetic susceptibility to OA in dogs with elbow dysplasia or cranial cruciate ligament rupture is unknown. Differences in the breed tolerance threshold of passive laxity for the development of hip OA suggests that genetic differences can be involved in the severity of canine OA (12). Whilst the significance of primary versus secondary canine OA is unresolved, canine OA *per se* is likely to have a significant genetic background.

Human primary OA is recognised as developing earlier in onset, and with greater severity, than natural "wear and tear". In man, primary OA is the most prevalent form of the disease (13), although population studies of OA are often defined purely on a radiological basis, and therefore may include secondary forms (14), such as hip dysplasia which cannot necessarily be differentiated once OA develops. Even with secondary human OA, there are significant genetic influences on the severity of secondary OA which develops (15).

Canine hip dysplasia

Canine hip dysplasia is a characterised by instability of the hip joint, which leads to hip subluxation (8). Repeated articular trauma from hip subluxation results in the development of synovitis, articular cartilage wear (16), malformation of the femoral head, neck and acetabulum and pain with associated physical disability. The disease process involves all the tissues of the hip joint, including the subchondral bone (17), round ligament (18), joint capsule (19), synovial membrane (18), and periarticular muscles (20).

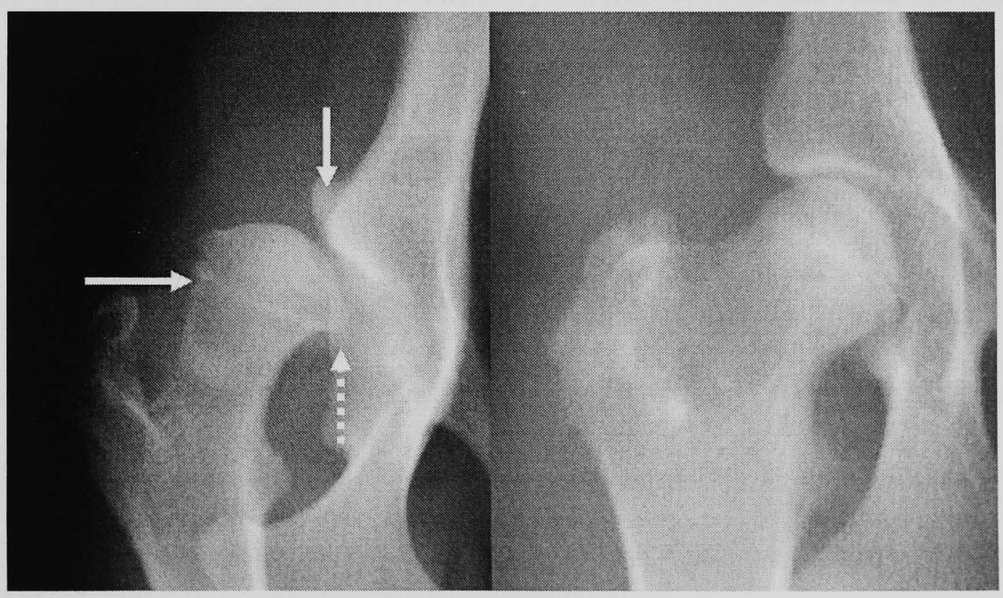
Canine hip dysplasia demonstrates a highly variable and dynamic phenotype both in affected individuals and across a dog population. Classically, the clinical presentation of hip dysplasia has a biphasic distribution within the canine population (21); young dogs are affected with the condition within the first year of life, with pain resulting from clinical subluxation of the hip, and older dogs have pain resulting from the development of OA of the hip as a result of coxofemoral incongruity, laxity, and subluxation (Figure 1). The incidence of hip OA in a dog breed can be directly related to the degree of hip joint laxity (12), and continues to progress linearly with age (22). The development of hip dysplasia in dogs is influenced by multiple factors including nutritional status (22,23), genetics (24) and hormonal factors (25).

The diagnosis of hip dysplasia is suspected upon clinical examination and confirmed by radiographic assessment of hip morphology. Multiple radiographic methods have been developed for detecting and quantifying abnormalities of the hip joint which indicate dysplasia. Conventionally features of the ventrodorsal extended hip radiograph are used to quantify different traits of hip dysplasia (26,27), such as static

hip laxity, morphology of the femoral head and acetabulum, joint congruency and osteoarthritic changes. Alternative radiographic measures, which quantify the laxity component of the disease through dynamic (28,29) or passive (30) stress radiography have also been developed. Stressed measures of hip laxity are the most sensitive method (31) for determining which individuals in a dog population which will develop OA over the long-term, but the ventrodorsal extended hip radiograph remains the more specific (31) and more commonly used method. It should be noted that the radiographic phenotype changes with age, as the disease (and associated radiographic changes) progress with time, which is reflected in the increase in both the radiographic features of OA (22) within individuals and radiographic (hip) scores across a dog population (32) as dogs become older.

Figure 1

Ventrodorsal extended hip radiograph of normal (right) and dysplasic (left) canine hips. The characteristic features of canine hip dysplasia, marked osteophyte formation (solid white arrows) and hip subluxation (dashed white arrow) can be observed on the dysplasic hip



Hip dysplasia affects most breeds of dog. Although the true prevalence of clinical hip dysplasia within individual breeds is unknown, there are estimations varying between 4.2% to 9.6% for clinical signs (33) and between 10% and 73% (8,34-36) for radiographic prevalence. The clinical importance of hip dysplasia, and associated OA of the hip joint, is highlighted by the fact that, in military working dogs, it is reported to be both a primary reason for rejection from training and the most common reason for ending active service (37,38).

At present, there is no medical or surgical treatment for canine hip dysplasia which can ameliorate or prevent the development of osteoarthritic changes of the affected joint. Management of affected individuals through anti-inflammatory medication, weight-optimisation and exercise restriction may be fruitful in the short term (39), but do not directly address the degenerative process affecting the joint. Surgical management may be required in cases not responding to medical management, through the use of corrective femoral or acetabular osteotomies, or arthroplastic procedures, such as total hip replacement (40) or femoral head and neck excision (41). Long-term medical management is expensive, whilst surgical management is both expensive and has the potential for significant morbidity.

Congenital hip dysplasia (also termed developmental dysplasia of the hip), is a human condition which demonstrates similar molecular, clinical and radiographic features to canine hip dysplasia. For example, features of the condition in both humans and dogs are similar, with clinical or radiographic hip joint laxity being a primary component of the disease, and hip joint laxity being strongly associated with the development of OA

in later life in both species (42). Furthermore the development of OA is also strongly associated with increasing age in humans (42), as has been described in dogs (22).

Although the clinical prevalence of congenital hip dysplasia is lower than in the dog, with an incidence of 0.1-1.8% of human births being affected by the disease (43), cross sectional analysis of humans populations indicates that the true radiographic prevalence of the disease is high (5.4-12.8%) (42). Congenital hip dysplasia in humans has a strong genetic background (43,44), with multiple familial aggregations and segregation analysis of familial pedigrees suggesting a two locus model (43), although the condition has not been extensively studied to date. Polymorphisms of the vitamin D receptor (*VDR*) and collagen type 2 alpha 1 gene (*COL2A1*) have been associated with the development of OA secondary to hip dysplasia (45). Even with secondary hip OA in humans, considerable genetic influences exist that affect the severity of the OA that develops (15,46).

In the dog, breed variations exist in the relative risk of developing hip dysplasia, with a higher frequency of disease observed in large and giant breeds, such as German Shepherd Dogs and Labrador Retrievers (47,48). Estimates of heritability for hip dysplasia in dogs vary widely, between 0.18 (49) and 0.74 (50). Furthermore, radiographic scores for the definition of hip phenotype also demonstrate heritability values of 0.23 (51) to 0.41 (52), and even higher heritability estimates of 0.50 to 0.61 are reported for radiographic measures of hip joint laxity (53). Benefits of selection of breeding canine populations with radiography for the assessment of the hip dysplasia phenotype have been widely described (51,54,55), and the degree of improvements in

a breed may be increased by heightening the stringency of selection of breeding animals on their radiographic phenotype (52).

The inheritance pattern of the canine hip dysplasia phenotype is complex (53), which suggests multiple major and minor quantitative trait loci contributing to trait expression. Initial investigations into the molecular genetic control of hip dysplasia in dogs have identified a number of major QTL that influence the phenotypic expression of hip joint laxity components of hip dysplasia (56,57). Interestingly, the QTL identified by different studies, utilising different dog populations (Portuguese Water Dog (56) and a Labrador Retriever / Greyhound pedigree (57)) were at different loci, which suggests that either different loci are responsible for the same disease in different breeds, or that the differences could be attributed to alternate methods of genotyping and statistical significance thresholds used (57). Clearly, the conflicting evidence over the likely genomic basis of canine hip dysplasia requires further investigation. To our knowledge, a candidate gene approach for the investigation of hip dysplasia has not been reported.

The hallmark of hip dysplasia, joint laxity, has been identified as an important risk factor in the development of degenerative joint disease of the canine hip joint (58). Breed differences exist in the degree of passive laxity which can be tolerated before the development of hip OA. This implies that genetic differences exist between dogs of different breeds for similar conditions that alter the phenotypic expression of OA (12). This has been confirmed by both studies of inheritance (59), and genomic studies which have identified QTLs which influence the expression of OA associated with hip dysplasia (60,61).

Canine elbow dysplasia

Canine elbow dysplasia is a term used to define multiple conditions of the canine elbow including fragmentation of the medial coronoid process (FCP) (9), osteochondrosis dissecans of the medial part of the humeral condyle (9) and ununited anconeal process (10). Each of these conditions may develop in isolation, or in combination with each other (62-64) and FCP is the most common condition in most (64), but not all dog populations (48). Each of these conditions results in the development of OA of the affected elbow joint.

FCP is characterised by fissuring and fragmentation of the cartilage and bone over the craniolateral aspect of the medial coronoid process of the ulna (Figures 2 and 3). Osteochondral fragments may remain in situ or may separate from the base of the coronoid process and become displaced (9). Cartilaginous 'kissing lesions' of the humeral condyle and secondary OA are commonly seen (9,65). Cartilage erosion over the medial coronoid process and the medial aspect of the humeral condyle can occur in the absence of discrete bony coronoid fragmentation (65,66), thus FCP probably represents a specific lesion within a wider spectrum of pathology affecting the coronoid process and medial compartment of the elbow joint which can be termed 'medial compartment disease' (MCD) (67).

Figure 2

A gross bone specimen indicating the site of fracture of the medial coronoid process of the ulna (dashed arrow).

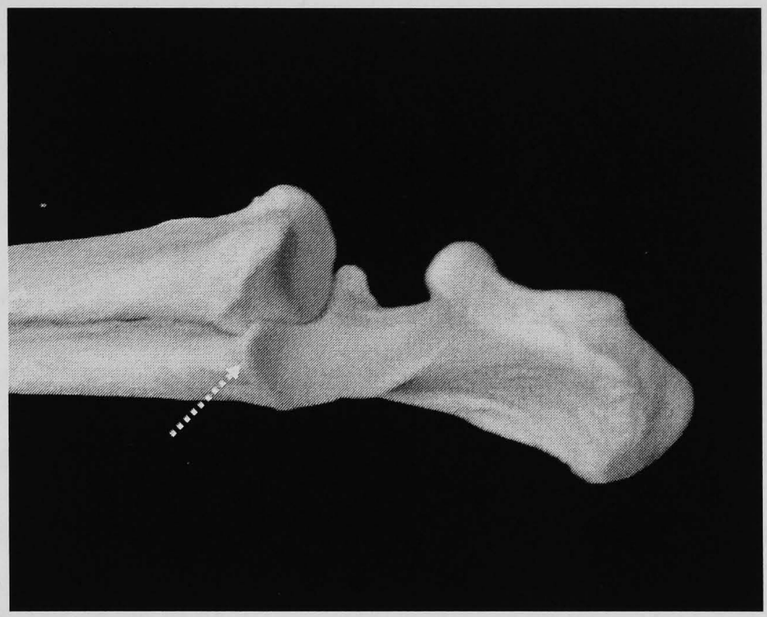
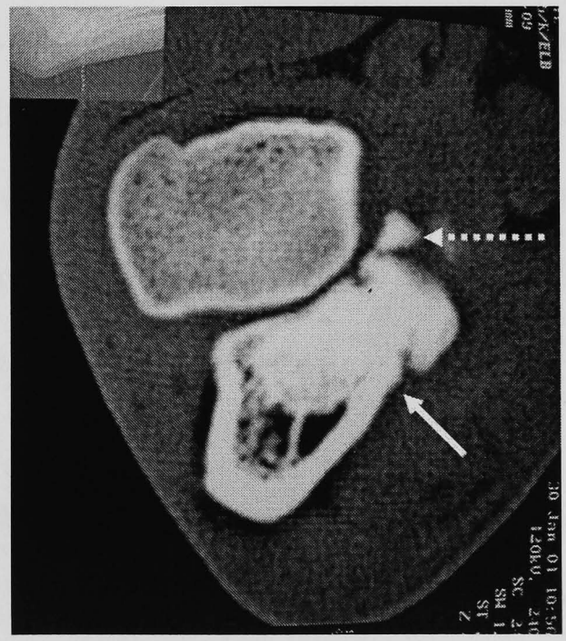


Figure 3

A saggital computed tomography view of the radius and ulna, demonstrating fragmentation of the medial coronoid process of the ulna (dashed white arrow) and sclerosis of the ulna (solid white arrow).



FCP was thought to represent a form of osteochondrosis (63). The medial coronoid process ossifies between 12 and 22 weeks (68) and may be susceptible to osteochondrosis during this period. Histological features consistent with

osteochondrosis have been reported in some dogs with FCP (9,69). However, other studies have not supported this theory and on the whole are more suggestive of an osteochondral fracture of the medial coronoid process (70). Guthrie and others (71) evaluated osteochondral fragments from 24 dogs with FCP. No microscopic evidence of osteochondrosis was found, and the histological picture was more consistent with a fibrous non-union. Fragmentation of the coronoid process has also been recognised in skeletally mature dogs and may also result secondary to trauma in some cases. Interestingly, adult dogs suffering traumatic coronoid process fracture which were treated by fragment excision appeared to have minimal progression of OA, suggesting that the spectrum of pathology observed with FCP in younger dogs is not solely due to the fragmentation alone (72).

The radiographic prevalence of elbow dysplasia has been reported to be between 2.9% and 17.8% (34,73), and the clinical prevalence has been between estimated between 4.0% and 5.0% in Labrador retrievers (74). The incidence of FCP alone in a population Labrador Retriever guiding dogs was reported to be 17.3% (75). Dogs can only be accurately assessed for FCP using computed tomography or direct visualisation. The expense of the former procedure and the invasiveness of the latter preclude their use routinely in clinical practice.

The majority of studies of elbow dysplasia in canine populations evaluated all the component conditions together. Dog populations are screened using a radiographic scoring system (International Elbow Working Group [IEWG] Scoring Scheme) (76) which uses the measurement of osteophyte size at multiple locations across the elbow joint to ascribe a score to a particular joint, but does not differentiate between the

different component conditions. A genetic basis to elbow dysplasia is suggested by the strong breed predispositions associated with the development of each of the components the diseases (48) though robust epidemiological data is lacking.

FCP appears to demonstrate a polygenic mode of inheritance (77). However, as most studies of the heritability of elbow dysplasia have used radiographic scoring for phenotype determination, and few radiographic scoring systems differentiate FCP from other component conditions of elbow dysplasia such as OCD and UAP. Therefore, although heritability estimates have been calculated for elbow dysplasia (0.10% to 0.77%) (73,78-81), the true values for FCP in isolation from UAP and OCD are unknown. Indeed, two studies suggest that FCP is inherited independently from OCD (77,82) and UAP (82), which supports the view that the heritability of FCP alone cannot be ascertained from studies utilising the radiographic assessment of elbow dysplasia *per se*.

At present, limited genomic data have been published regarding canine elbow dysplasia in peer reviewed literature. Salg and others (83) analysed a population of Labrador retrievers by pedigree and sibling pair analysis and reported that FCP was controlled by a major gene, with variable expression (male : female ratio 75%:25%). Previous epidemiological studies of FCP support the finding that this condition has a 3:1 male to female sex bias (78,84). The collagen genes were selected as candidate genes within the population of Labrador Retrievers studied, on the basis of their involvement in the bone formation and skeletal disorders in man. These included Type I collagen, alpha 1 chain (*COL1A1*), type I collagen, alpha 2 chain (*COL1A2*), type II collagen alpha 1 chain, type III collagen, alpha 3 chain (*COL3A1*), type V

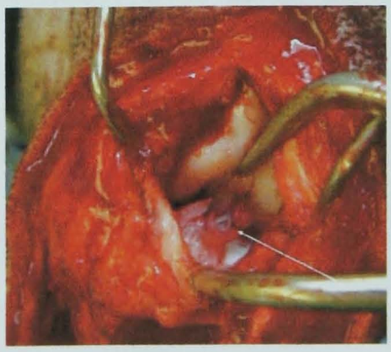
collagen, alpha 5 chain (*COL5A1*), type V collagen, alpha 2 chain (*COL5A2*), type VI collagen, alpha 3 chain (*COL6A3*), type IX collagen, alpha 1 chain (*COL9A1*), type IX collagen, alpha 2 chain (*COL9A2*), type IX collagen, alpha 3 chain (*COL9A3*), type X collagen, alpha 1 chain (*COL10A1*), type XI collagen, alpha 1 chain (*COL11A1*), type XI collagen, alpha 2 chain (*COL11A2*) and type XXIV collagen, alpha 1 chain (*COL24A1*) and the Vitamin D receptor genes. No significant deviation from 50% allele sharing between affected sib-pairs was observed, using variable number tandem repeat (VNTR) markers near the candidate genes. This indicated that none of these genes were associated with the development of FCP within the Labrador Retriever population studied. No estimations of statistical power were provided with the study, although only 34 sibling pairs were evaluated, which is a relatively small number, implying that these genes should not be completely discounted until further work confirms these findings.

Canine Cruciate Ligament Rupture

Rupture of the CCL is a devastating injury leading to stifle instability and the progressive development of OA (11,85,86) (Figure 4).

Figure 4

Open arthrotomy demonstrating the remnants of the ruptured cranial cruciate ligament.



Affected dogs may demonstrate acute or chronic lameness associated with OA of the affected joint and failure of the associated ligament. The disease often affects both stifle joints of an individual, with up to 60% of dogs affected developing disease in the contralateral stifle within 18 months of the initial diagnosis (87). Dogs may be affected by CCL rupture at any age (88), although breeds which are predisposed to CCL rupture tend to be affected earlier in their lifetime (11)

The cranial cruciate ligament is the primary ligamentous stabiliser of the stifle joint in dogs. The CCL acts to limit internal rotation (89), hyperextension (89), varus-valgus motion (90) and cranial tibial displacement (89) of the canine stifle joint, which makes it the anatomical equivalent of the human anterior cruciate ligament. The underling aetiology of CCL failure is presently unknown (91) although unlike the human condition it is almost never results from an traumatic event, hence the often chronic nature of the condition. A number of factors are thought to contribute to the development of CCL rupture, such as genetics (33,88,92), age related morphological changes in the ligament (93), activity levels (94), tibial plateau slope (66), stenosis of the intercondylar notch (95,96) and neuter status (88,97).

Cranial cruciate ligament rupture is of major economic and welfare importance to the canine population. The disease is one of the most common canine orthopaedic conditions (6), accounting for nearly 20% of veterinary presentations for canine lameness (98), and an estimated prevalence of disease being 3% in the general dog population (97). The estimated the prevalence of CCL rupture within the general Labrador Retriever population in the United Kingdom to be 6.6%, with 5% of dogs requiring surgery (74). Affected dogs require surgical treatment to address the stifle

instability, associated cartilage injury, although presently there is no treatment which prevents the development or progression of degenerative joint disease (99), or which can consistently return dogs to normal function (100). The economic impact of CCL rupture for owners of affected dogs was estimated to be over $1.3billion in the USA alone in 2003 (101).

Human anterior cruciate ligament (ACL) rupture usually occurs as the result of a traumatic event. However it is recognised that over 20% of patients with symptomatic knee OA have complete cruciate ligament tears, with no history of a traumatic event in half of the cases (102). Furthermore, joint laxity is hypothesised to be an important contributor to the pathogenesis of human knee OA, with laxity in the valgus-varus (103) and anterior-posterior (103,104) planes increasing with the severity of OA. Increases in anterior-posterior laxity are also recognised in breeds of dog susceptible to CCL rupture when compared to breeds protected from CCL rupture (91,105,106) which indicates that genetic factors may contribute to stifle laxity, CCL rupture and concomitant stifle OA.

Pathological changes to the ACL resulting in knee laxity may predispose human patients to knee OA. This hypothesis is supported by spontaneous animal models of OA which demonstrate an association between ligament laxity, specifically of the ACL, and the development of OA (107). Increases in anterior-posterior laxity are also recognised in breeds of dog susceptible to CCL rupture when compared to breeds protected from CCL rupture (91,105,106) which indicates that genetic factors may contribute to the development of CCL rupture.

At the molecular level increases of pro-matrix metalloproteinase -2 (*MMP2*) protein have been identified in normal ACLs of dogs with a high risk of CCL rupture (Labrador Retriever) compared to dogs with a low risk of CCL rupture (Greyhound) (91). Similar changes in gross ACL biomechanical properties (increasing laxity) have been related to molecular differences (increased *MMP2*) in the ACL in an animal model of spontaneous knee OA (107), further supporting the link between the development of knee OA, knee laxity and molecular changes in articular tissues.

Epidemiological studies have highlighted that dogs demonstrate a breed-associated risk to CCL rupture (88,108), with "at-risk" breeds such as the Labrador Retriever and Rottweiler demonstrating much higher levels of risk than "protected" breeds, such as the Greyhound. Dogs from breeds predisposed to ACL rupture have reduced ligament stiffness and reduce load to ultimate failure when compared to dogs from breeds with low risk of CCL rupture (91,105,106). This implies that the genetic susceptibility to the development of CCL rupture manifests itself through changes in the mechanical properties of the CCL. Difficulties in mapping the phenotype of CCL rupture are compounded by the fact that presently no diagnostic tests or screening schemes exist for this condition in canine populations. Furthermore, CCL rupture may not be clinically evident until late in life, which makes the accurate selection of controls difficult.

To date, only two studies have investigated the heritability of CCL rupture in dog breeds. Estimates of heritability of 0.31 in Boxers (33), and 0.27 in Newfoundland's (92) suggest there is a genetic component to the disease. However, both these studies probably underestimate the genetic component to CCL rupture, as affected dogs may

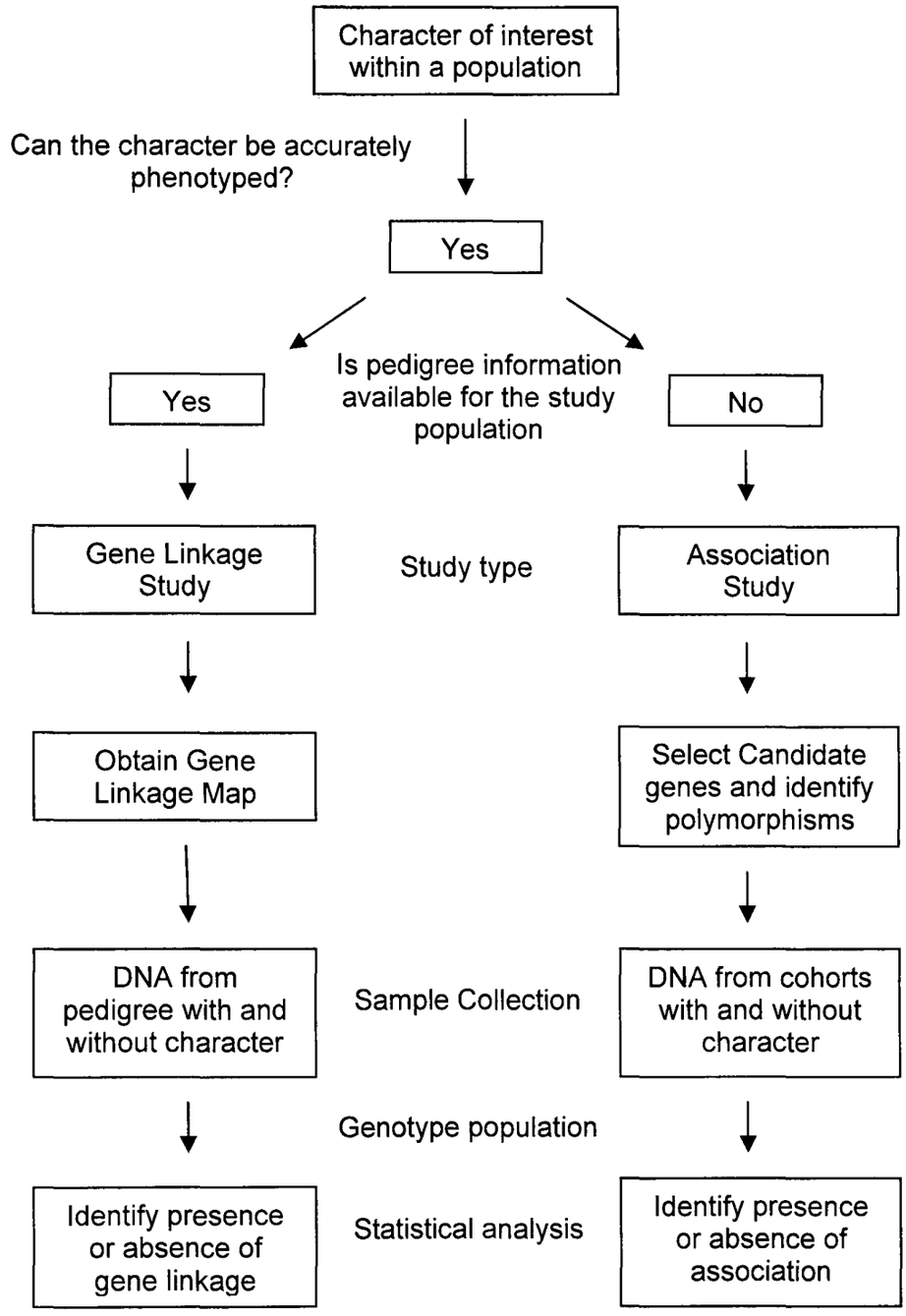
not demonstrate the phenotype until later life. Segregation analysis of the Newfoundland cohort predicted a major gene effect with a recessive pattern of inheritance, with the frequency of the recessive allele being 0.60 with partial penetrance of 51%. The preliminary results of a microsatellite screen of the Newfoundland pedigree indicated that CCL rupture is associated with chromosome 3 (109), although the exact location and strength of association have not been published.

Application of population genomics to canine OA

Genomic investigation of developmental diseases within dog populations requires either a study by association or gene linkage approach (Figure 5). In either case, studies are compromised by the variability of the phenotypic presentation of both the disease and the patients themselves (breeds). This certainly applies to studies examining canine hip or elbow dysplasia, where many different clinical, morphological and radiographic phenotypes exist for each diagnosis. The genetic heterogeneity which exists between and within breeds (110) further compounds the difficulty of studying canine genomic disease.

Figure 5

Basic overview of genomic investigation within populations.



Gene Linkage

Gene linkage maps are maps of known genetic loci across a genome, at known genetic intervals. As the physical distance between loci reduces, the less likely it becomes that genes causing a phenotypic trait will be subject to recombination during meiosis. Hence, it becomes the more likely that an allele at a specific loci will be transmitted with the causative gene(s), and the genetic trait. As such, the loci and gene are said to be in linkage. Each locus on a linkage map is genotyped in each individual within a known pedigree, for which phenotypic information is recorded for the trait being investigated. This allows a mathematical measurement of linkage to be made with the phenotype, and thus the loci in linkage with the phenotype can be identified.

Suitable canine gene linkage maps exist (111), and the use of a gene linkage approach with pedigree analysis allows the most accurate method for identification of genes involved with a phenotypic trait. However the process is time and labour intensive, requiring the genotyping of a large number of loci and the recording of a large amount of phenotypic information. The likelihood of obtaining a positive association with such a study is dependent on the quality of the pedigree and phenotypic information, and the strength (which reflects density or the number of loci investigated) of the linkage map used.

Polygenic disorders are difficult to elucidate using conventional linkage analysis, as the linkage maps available are frequently not powerful enough to detect an association with the multiple genes involved. Additionally, these studies provide suggestive evidence of linkage to relatively large chromosomal regions, and finer linkage maps with larger studies are required to pinpoint the genes responsible for a given disorder

(112). The controlling genes may have small or moderate effects on a trait or disease, thus requiring extremely large pedigree numbers to produce reliable results (113). Obtaining suitable pedigree sizes and numbers, with full phenotypic information, in canine disease studies is extremely time-consuming.

Studies of association

Studies of association are applicable for canine polygenic disorders where pedigree information is inaccurate or unavailable. Genetic polymorphisms may be identified near to, or within genes of interest (candidate genes), and then matched populations with and without disease may be screened for the polymorphisms and their associations tested by statistical means (113,114). These studies have the advantage of not requiring pedigree information for completion, so they can be rapidly set up, and screen fewer polymorphisms than linkage studies, allowing for time and cost savings. However, it should be noted that although genes with positive polymorphisms identified by association are not confirmed as being linked to the phenotype until they have been tested by gene linkage.

The success of genetic association studies in man, to date, has been limited with many successful association studies not being consistently repeatable (115). This has lead to recommendations for how study design should deal with population stratification, gene selection and quality measures for the genotyping analysis in this type of research (116). The major limitation of the association approach in human studies is the potential for spurious association (false positives) due to confounding variables, such as ethnicity (113).

Clearly, the same would be expected to be true of canine disease studies. However, low haplotype diversity exists within dog breeds, with 80% of chromosomes in a breed carrying two to four haplotypes. Furthermore a large degree of haplotype sharing is observed between different breeds of dog (117) suggesting that breed diversity may be less important, providing disease, rather than breed, specific polymorphic loci are evaluated. Additionally, dog breeding has resulted in extensive linkage disequilibrium, which is up to 100x greater than in humans (117). Hence a smaller number of markers should be required in canine gene association studies, and relatively small sample sizes should still produce strong associations, compared to human studies. Furthermore, the high linkage disequilibrium observed in dogs also implies that a small number of loci would be required in gene linkage maps, to obtain strong linkage, when compared to such studies in humans.

Labrador Retrievers show the lowest linkage disequilibrium, when compared to Akita, Bernese Mountain dogs and Pekingese (117). This is probably as a result of their popularity and broader founder population which will promote a greater degree of heterogeneity, although they still demonstrate a high degree of haplotype sharing. Clearly, the issue of breed specificity with regard to studies of canine diseases not specific to breeds, such as hip and elbow dysplasia, will remain unresolved until more information has been published on polymorphic allele frequencies within and between breeds.

There are two further variables which need to be considered when investigating canine disease using polymorphism association studies. Firstly, the sample sizes of cases and controls must be large enough to ensure that positive associations are of

reasonable power (80%) (118), which can be estimated from the allele frequencies in the different populations. Secondly, the quality of the control population must be high enough (i.e. their phenotype must be accurately determined) to prevent the failure of association (false negatives) purely on the basis of disease in the control population. Clearly, ethical issues exist regarding both sampling and phenotyping control populations for genotyping studies, and ethical frameworks regarding these points need to be determined at the point of study design.

Polymorphic loci used in association studies and linkage studies include microsatellites markers, variable nucleotide tandem repeats (VNTRs) and single nucleotide polymorphisms (SNPs). Microsatellite markers and VNTRs are short repeated lengths of sequence, which are polymorphic in the number of repeated elements that they contain. Different lengths (numbers of elements) of these repeated sequences are different alleles. Such markers are usually identified adjacent to genes, although they may occasionally be intronic or exonic in location. In contrast, SNPs are single nucleotide changes within the genome of which the most common allele occurs with less than 99% frequency in the population at large (119). The functional significance of SNPs within and around genes are that, in a coding region they may directly impact on the protein structure and function, in an intronic region they may alter splicing (120), and in the promoter region they may influence gene expression (121).

There are multiple advantages to studying SNPs for investigating genetic influences on disease. The large number of SNPs present within the genome may show distinctive patterns of linkage disequilibrium which may be utilised in genetic linkage

and direct association analyses. Allelic discrimination is relatively straight forward, and multiple methods of high through put genotyping exist (122). SNPs are less mutable than other types of polymorphism (123), which should make them more reliable for assessing linkage disequilibrium, allelic associations and co-segregation phenomena, as associations are unlikely to be confounded by mutation between generations (119). Thus, SNP identification in genes potentially provides a rapid and straightforward method of evaluating gene association with disease in canine populations.

Methods for genome wide SNP identification have been developed in human research and are already being use to identify genes and SNPs associated with human diseases (124,125). Population simulations estimate that roughly 500,000 SNPs are required in humans to provide genome wide linkage (126), but the more extensive linkage disequilibrium identified in dogs implies that far fewer SNPs would be required. Such technology is currently being developed for the canine genome (127), but at present candidate gene studies by association are the most viable method for identifying those genes likely to have an influence on susceptibility or outcome of canine diseases.

Candidate gene selection

Candidate genes for a disease are genes for which evidence exists showing that they may be related to that disease. Candidate gene selection is not an exact science. Valid methodologies for the selection for candidate genes include; gene position within a particular region of the human genome with evidence of linkage to disease (128), genes known to be involved in a physiological process relevant to disease (129),

increased or decreased gene expression in diseased *in vivo* tissue (130), *in vitro* tissue models, gene knockout experiments, the contribution of the gene product to homeostasis in normal tissue (131), or gene polymorphism(s) identified with a familial form of the disease (132). Once candidate genes have been identified, SNP identification provides a rapid and straightforward method of evaluating gene association with disease in canine populations.

Human OA candidate gene studies

A number of interlinked molecular pathways contributing to the degenerative process have been identified in osteoarthritic cartilage, such as those of cytokines (133), degradative enzyme production (134) and matrix synthesis (135). The majority of human association studies of OA evaluated components of and/or molecules affecting the extracellular matrix of articular cartilage; namely; the collagen proteins, other structural proteins, hormones, cytokines and growth factors.

The candidate gene approach has been applied to human hip, knee, hand and generalised OA, often with conflicting results. A summary of the findings to date for association studies of each gene class are described in Table 1, and the functions of each gene are describes as follows;

Table 1

Summary of positive and negative candidate gene association in studies of osteoarthritis in humans (NR = Not reported).

	Associations	
Genes	***Positive Association (reference No.)***	***No Association (reference No.)***
Collagens		
COL1	NR	(136-138)
COL2A1	(45,139-144)	(131,137,145-147,147)
COL9A1	(148,149)	(131,142,148,150)
COL9A3	(142)	NR
COL10A1	NR	(142)
COL11A1	NR	(131)
COL11A2	(148)	(131)
Extracellular matrix components		
CILP	(130)	NR
COMP	NR	(130,132)
CRTM	(131,151)	(152)
AGC1	(129,153)	NR
ASPN	(154)	NR
Hormones		
Oestrogen receptor-α	(130,155-157)	(136)
VDR	(144,158)	(136,137,145)
Cytokines		
IL1α	(159-161)	NR
IL1RA	(159-162)	NR
IL1B	(160-163)	NR
IL4R	(164)	NR
IL8	NR	(165)
IL10	(166)	(167)
Osteoprotegerin	(130)	NR (166)
Growth factors		
BMP2	(130)	NR
IGF1	(168)	(143)
TGFβ	(169)	NR
Other components		
ADAM12	(130)	NR
CD36 antigen	(130)	NR
COX2	(130)	NR
NCOR2	(130)	NR
ADAMTS3	(170)	(165)
Tetranectin	(130)	NR
α1-Antichymotrypsin	(130,171)	NR
TNFAIP6	(130)	(128)
ACE	(172)	NR
FRZB	(128,173)	NR
C3	(174)	NR
h-ras	(175)	(176)
SLC26A2	(177)	NR
PAPSS2	NR	(177)
CALM1	(178)	NR

The Collagens

Collagen is the predominant extracellular matrix protein of articular cartilage representing over 50% of the dry weight of articular cartilage. At least 16 different types of collagen exist, with 29 different chains, although all contain a characteristic triple helical structure. Collagen types 2, 5, 6, 9, 11 and 16 are the most commonly identified isoforms in articular cartilage.

Type I collagen is present in very small amounts in articular cartilage, and thus its limited role in the structure of the cartilage extracellular matrix. Type II collagen represents 90-95% of the total collagen in articular cartilage. Type IX collagen is present in small amounts in articular cartilage where it is found in association with the surface of Type II collagen fibrils, although its exact function remains unknown. Type X collagen is a short chain collagen expressed in hypertrophic cartilage, but only to a limited degree in articular cartilage. Type XI collagen is a long chain collagen which is present in the deep calcified zone of mature joints.

Other components of the extracellular matrix

Other extracellular matrix protein genes have been assessed as candidate genes for OA. Cartilage intermediate layer protein (*CILP*) is a non-collagenous protein of undetermined function, which is synthesised by chondrocytes (179). *CILP* synthesis is increased in early OA (180). Cartilage oligometric matrix protein (*COMP*) is a non-collagenous extracellular matrix protein, whose function is not entirely determined. *COMP* proteins mediate cell-matrix and matrix-matrix interactions, and possibly chondrocyte attachment (181). Expression of *COMP* is increased in the articular

cartilage of a mouse model of OA (182) and synovial levels are raised in osteoarthritic dogs (183). *COMP* mutations are the cause of other osteochondral dysplasias, such as psuedoachondrodysplasia and multiple epiphyseal dysplasia (184), which are associated with the early development of OA.

Matrillin (*CRTM*) is a non-collagenous protein expressed in developing cartilage, particularly epiphyseal cartilage (151). Aggrecan (*AGR*) is the primary proteoglycan constituent of cartilage extracellular matrix. This molecule is important in the proper functioning of articular cartilage because it provides a hydrated gel structure (via its interaction with hyaluronan and link protein) that endows the cartilage with load-bearing properties, and thus is an obvious candidate gene for genetic studies of OA.

Asporin is a extracellular matrix protein recently identified as belonging to the small leucine-rich proteoglycan (SLRP) family, which also contains decorin and biglycan (185). The exact function of Asporin is unknown; however it binds to *TGFβ in vitro*, and variations in the aspartic acid repeat functionally affect the responsiveness of chondrocytes to *TGFβ* (154). Asporin is abundantly expressed in osteoarthritic articular cartilage *in vitro* (185).

Hormones

Hormone receptors provide the cellular gateway for hormonal regulation of cellular function. Estrogens are associated with a protective effect on the development of OA in women (186). The prevalence of OA in postmenopausal women is much higher than in men (187), suggesting that oestrogen exerts a protective effect. Vitamin D is integrally involved with skeletal development and metabolism, and immune cell

development. Low intake and low serum levels of Vitamin D are both associated with an increased risk for progression of OA of the knee (188).

Cytokines

Several cytokines are involved in cartilage metabolism and synthesized by synovial cells and cartilage chondrocytes. The interleukins are cytokines which have a primary role in the development and progression of OA (133). Interleukin 1 (*IL1*) is believed to be an important catabolic cytokine of the osteoarthritic joint and can stimulate synthesis of a number of proteases, which result in the breakdown of the extracellular matrix. Interleukin 1 receptor antagonist (*IL1Ra*) competes with *IL1* for binding to the *IL1* receptors (*IL1R*) and can act as an inhibitor of cartilage loss. When the catabolic and anabolic activities of the cytokines are balanced, cartilage integrity is maintained. Where there is an imbalance favouring catabolism, however, cartilage destruction can proceed, resulting in OA. Hence, a proportion of the genetic susceptibility to OA may be encoded for by variation in the activity of interleukin genes. Interleukin 4 (*IL4*) is an active signalling molecule involved in the regulation of cartilage integrity by mechanical stimulation (189). Interleukin 10 (*IL10*) is an immunomodulatory cytokine primarily secreted by monocytes, and is expressed at increased quantities in OA synovium, although it may also play a role in chondrocyte metabolism (190).

Osteoprotegrin (*OPG*) is a member of the tumour necrosis receptor superfamily (Number 11B, also termed *TNFR11B*), which is secreted without a transmembrane domain. *OPG* binds *RANK* (receptor activator of nuclear factor-κB), a member of the TNF receptor family, expressed on the osteoclasts, and thus prevents interaction with its ligand, *RANK* Ligand (*RANKL*). *RANKL*, also know as osteoprotegrin ligand

(*OPGL*) is a cell membrane-anchored or soluble ligand for *RANK* expressed on the osteoblast / stromal cell surface, whose interaction with *RANK* stimulates osteoclastogenesis, which can be inhibited by *OPG* (191). The ratio of *OPG* and *OPGL* correlate strongly with indices of bone remodelling (histomorphometric data) in normal human cancellous bone (192). *OPG* expression is increased in osteoarthritic cartilage, and by *IL1* stimulation of chondrocytes *in vitro* (193).

Growth factors

Growth factors are important in the homeostatic regulation of cartilage, controlling functions such as chondrocyte integrin expression (194). Bone morphogenic proteins (BMPs) are potent growth and differentiation factors which belong to the transforming growth factor beta (*TGFβ*) superfamily. Exogenous *BMP2* increases proteoglycan and collagen synthesis and maintains the adult chondrocyte response *in vitro* (195), and is identified in OA chondrocytes and osteophyte tissue, but not in chondrocytes from healthy cartilage (196). Insulin like growth factor-I (*IGF1*) plays an important role in cartilage homeostasis. *IGF1* stimulates chondrocytes' proliferation and their synthesis of proteoglycan and *COL2*, and inhibits the endogenous catabolic activity of articular cartilage (197) *in vitro*. Transforming growth factor beta is a growth factor which may inhibit or stimulate articular cartilage synthesis, dependent upon experimental conditions (197).

Other genes associated with human OA

A number of other genes have been associated with the development of OA. *ADAM12* is a metalloprotease which regulates the formation of macrophage derived giant cells,

possibly by mediating the effects of 1,25-hydroxyvitamin D_3 on cell-cell fusion. Blocking *ADAM12* mRNA in osteoclast precursor cells results in a 50% decrease in giant cell formation (198), which may explain the association of mutations of this gene with both the presence and progression of osteophytosis.

CD36 is a Type I collagen receptor/ thrombospondin receptor, which is expressed primarily in mid zone chondrocytes, and expression is markedly increased in OA cartilage (199), although whether this is a cause or effect of OA change is unclear. Cyclooxygenase (*COX*) is a membrane bound haem protein which is expressed in the synovium and chondrocytes of OA cartilage. Increased prostaglandin E-2 (PGE_2, a pro-inflammatory mediator) synthesis by *COX2* in articular cartilage is a cellular response to activation by pro-inflammatory stimuli and an important component in the pathogenesis of arthritis. Osteoarthritic cartilage produces more PGE_2 than non-arthritic cartilage (200) and the synovium produces cyclooxygenase 2 (*COX2*) in osteoarthritic patients (201), although to a lesser degree than in immune mediated arthritis (202). Nuclear receptor co-repressor 2 (*NCOR2)* is a nuclear transcription factor under hormonal control, which is a silencing mediator for retinoid and thyroid hormone receptors.

Tetranectin (*TNA*) is a phosphorylated glycoprotein postulated to regulate mineral deposition within bone (203). Although the role of tetranectin in the pathogenesis of OA is currently unknown, it has been implicated in the impaired regulation of fibrinolysis associated with the inflammatory process in rheumatoid arthritis (204). Alpha 1-antichymotrypsin (*AACT*) is serine proteinase inhibitor which helps regulate diverse physiological processes such as coagulation, fibrinolysis, complement

activation, angiogenesis, apoptosis, inflammation, neoplasia and viral pathogenesis (205) and thus potentially can prevent the degradation of connective tissue components. Tumour necrosis factor interacting protein 6 (*TNFAIP6*) is also referred to as tumour necrosis factor secreted glycoprotein 6 (*TSG-6*), and functions to modulate the interaction between hyaluronan and cell surface receptor *CD44* (206). Tumour necrosis factor interacting protein 6 is expressed in the synovium and cartilage of osteoarthritic and rheumatoid joints (207), indicating that it may have a role in the pathogenesis of arthritic conditions.

Angiotensin converting enzyme (*ACE*) is responsible for converting angiotensin I to angiotensin II, which is a potent vasoconstrictor of the renin-angiotensin system, and also inactivates bradykinin, a vasodilator of the kallikrein-kinin system. Levels of kinin B2 receptors in synovium are up-regulated in osteoarthritic patients (208), thus indicating that there may be a potential link between the features of OA and *ACE* activity (172).

Secreted frizzled-related protein 3 (*sFRP3*) is a glycoprotein which antagonises the signalling of wingless (wnt) ligands through the frizzled membrane bound receptors, which control the primary activation of T cell factor/lymphoid-enhancing factor-dependent transcriptional activation. Joint patterning in embryogenesis (209) and bone formation (210) are determined by the wnt pathway.

Complement component 3 (C3) is a potent inflammatory mediator, and increased expression has been identified in the synovium of joints with rheumatoid arthritis (174) although no previous role in the pathogenesis of OA has been identified.

Harvey rat sarcoma viral oncogene homolog (H-*ras)* is an oncogene expressed in the synovium (175). Mutations in two sulphation related genes, phosphoadenosine 5'-phosphosulphate synthetase (*PAPSS2*) and solute carrier family 26, member 2 (*SLC26A2*) are responsible for severe chondrodysplasias in both the human and mouse. The role of abnormalities in the sulphation genes in OA is plausible, as proteoglycans such as aggrecan require the negative charge provided by sulphation to provide compressive resistance, and sulphation is reduced in osteoarthritic cartilage (211).

Calmodulin (*CALM1*) is a ubiquitous, calcium binding protein which regulates calcium signalling, and may be involved in collagenases and proteoglycanase activity (212). *CALM1* expression is increased in hip and knee osteoarthritic cartilage, compared to normal cartilage (178).

OA and Major Histocompatability Complexes (MHC)

Strong associations have been identified between MHC alleles and immune mediated arthritis (213). Associations have also been identified between MHC class II alleles and the development of OA (214,215). It has been hypothesised that the repeated association of the DR2 allele with OA suggests that DR2 may have a role in restricting immunological responses to the low-grade inflammation characteristic of OA (215). This allele may be in linkage disequilibrium with other DR alleles which are involved in the pathogenesis of OA or may predispose to T cell activation in other tissues, such as synovium or bone, involved in the pathogenesis of OA.

Rational for gene expression profiling to select candidate genes

The most successful disease association study of candidate genes in human OA to date was published by Valdes and others (130). By comparing the levels of mRNA expression for normal and OA affected synovium, and normal and OA affected cartilage in gene expression libraries (Incyte, Palo Alto, California), 54 genes were identified whose transcripts appeared to be differentially expressed in human osteoarthritic tissue. Twenty two of these genes were evaluated for intragenic SNPs by using public databases and published literature (n = 10) or by screening them for polymorphisms (n = 12). Seven genes showed statistical significance with the onset or progression of female knee arthritis, and eight showed nearly-significant association ($P < 0.07$). Additionally, significant associations were made between five genes with susceptibility traits and four genes with progression traits. Subsequently, a number of these associations have been identified as being reproducible in women alone (*ERα, BMP2*), men alone (*VDR*), and both women and men (*ADAM12, CILP* and *OPG*) (216). These results demonstrate the benefit of using expression profiling to select candidate genes for disease association studies.

A valid argument exists for the evaluation of candidate genes expressed in tissues other than cartilage for association studies of canine OA secondary to joint dysplasia. Given the role of laxity in the development of OA (58) other tissues such a joint capsule and the CCL should be evaluated. Likewise, the potential role of articular incongruency in the pathogenesis of some forms of elbow dysplasia (217) indicates that genes involved in the regulation of physeal growth should be evaluated in OA.

Gene Expression Profiling

The quantification of gene expression is a fundamental tool for investigating gene function in biological systems, particularly for elucidating pathological mechanisms at play in diseased tissues. Gene expression profiling is a broad assessment of mRNA expression, i.e. an evaluation of which genes are expressed within a tissue sample, and an assessment of their level of expression. Ideally any measurement should be accurate, repeatable, user friendly, cost efficient and provide a measure of the number of gene transcripts provided per cell. Many papers have been published discussing gene expression within osteoarthritic cartilage or synovium, thus making the selection of candidate genes on the basis of expression alone a laborious task.

Quantitative (Real-time) reverse-transcriptase polymerase chain reaction

Quantitative (real-time) reverse transcriptase-polymerase chain reaction (RT-qPCR) is currently considered the most accurate technique for quantifying gene expression. The principle of the process is straightforward; a polymerase chain reaction specific for the mRNA of interest is performed with the inclusion of a fluorometric dye binding to double stranded DNA, or a probe contaning a fluorometic dye whose sequence is complementary to part of the amplicon and is thus digested by the 3' exonuclease activity of the Taq polymerase during the PCR. After each PCR cycle, a measurement of the dye bound or released in each well is performed at the appropriate wavelength using a laser scanner. As the quantity of PCR product is doubled with each PCR cycle, the level of fluorescence detected should also be doubled. A given value of fluorescence, know as the threshold cycle (C_T) is ascribed to the level of fluorescence

where product is "detected". Consequently, each plate well has a C_T value determined, which is an extremely accurate reflection of the quantity of mRNA present.

With the publication of the canine genome (218), it is now possible to quantify individual mRNA expression of any canine gene in clinical osteoarthritic tissue (cartilage (219), synovium (220) or synovial fluid), or by using *in vitro* (221) or in-vivo (85) models of osteoarthritic processes using RT-qPCR. Although alternative measures of gene expression are available (such as oligonucleotide microarrays) for the quantification of mRNA transcripts in canine tissue, such as cartilage (222), RT-qPCR validation of the measures of gene expression is still required. The method provides accurate quantification of transcript number, good sensitivity over a wide range of transcript expression levels, and increasing high throughput capabilities. Several factors contribute to errors of variation in gene expression measurement, including issues relating to sample starting cell number and sample cell types, mRNA extraction protocol and handling techniques (223), mRNA quality (224,225), method of reverse transcription (226), PCR inhibition (227) and analytical detection chemistry method (223).

RNA Quality

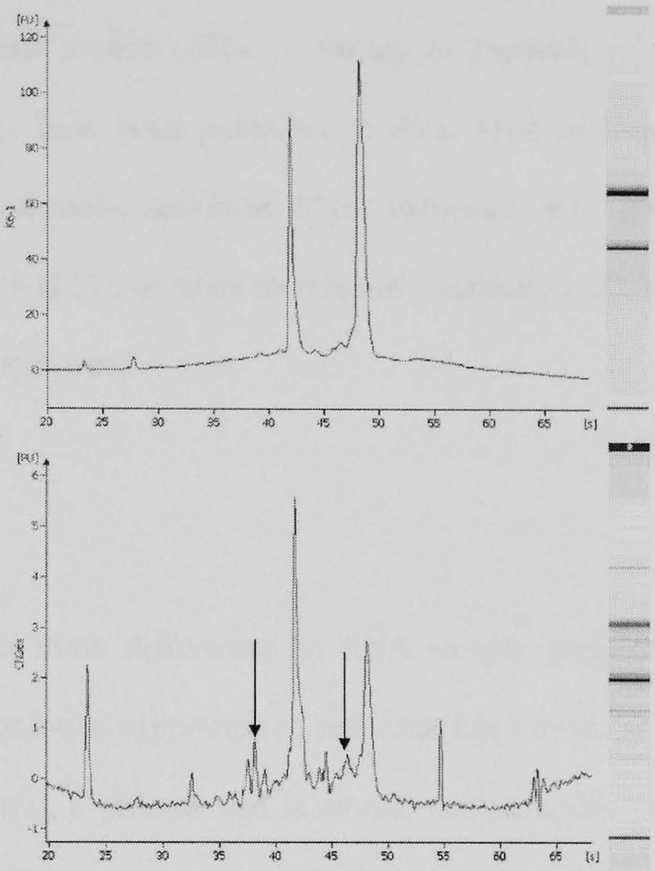
Assessment of the quality of RNA is a measure of RNA purity and integrity. RNA integrity is of critical importance because downstream gene expression profiling may be altered by changes in integrity (223,224). No gold-standard method exists to determine the quality of RNA extracted from tissue or cell cultures. Traditionally, visual assessment of an electrophoretic trace has been used as an identifier of RNA quality. Subsequently, the ribosomal band (28S:18S) ratio has also been used as an

identifier of RNA quality, with a ratio > 2.0 indicative of high-quality RNA. A number of alternative metrics have been used to quantify RNA quality, such as the 260 nm to 230 nm absorbance ratio (A_{260}:A_{280} ratio), (228) and ethidium bromide (229) or SYBR green dye stained agarose gel electrophoresis.

As low amounts of RNA are recovered from articular cartilage, quality assessment is ideally performed with the minimum amount of sample necessary. The recent development of microfluidic capillary electrophoresis has allowed the assessment of RNA quality (225) with low volumes of sample (1 μl) through direct trace observation and automated calculation of the 28S:18S ratio (Figure 6). Analysis of RNA integrity can also be performed by use of computational software based analysis of the electrophoretic trace (224,230). The DF (230) is a figure calculated by use of a mathematic model examining degradation peak signals present in the lower molecular weight range and comparing them with ribosomal peak heights. A lower number denotes a higher quality sample. The RIN is an algorithm that calculates RNA integrity from the electrophoresis trace, by evaluation of features such as the height of the 18S peak, the ratio of the area of the ribosomal bands compared with the total area of the electropherogram, and the ratio of the fast area of the electropherogram to the total area of the electropherogram (231). A comparison of the RIN and DF of human tissue samples reported that the RIN produces the most reliable data (224).

Figure 6

The microfluidic capillary gel electrophoresis traces from samples of high (top) and low (bottom) quality RNA. The 28S ribosomal RNA peak height is reduced in size compared to the 18S ribosomal RNA peak, and multiple small RNA degradation peaks (black arrows) are also present in the low quality sample.



The ideal measure of RNA integrity is to use a 3' and 5' PCR (224). However, this requires the both the reverse transcription and qPCR measurement of the mRNA sample, which is more costly and time consuming, particularly if a sample is determined to be of insufficient quality to be used.

RNA Extraction from Articular Cartilage

Extraction and purification of RNA from articular cartilage is problematic. The tissue is relatively acellular, usually only available is small quantities (with clinical canine samples often < 100 mg in wet-weight), and contains a large amount of proteoglycan in the extracellular matrix (232). A variety of methods of RNA extraction from articular cartilage have been published to date. Most methods use liquid nitrogen dismembration, phenol-chloroform RNA extraction and caesium trifluoroacetate ultracentrifugation (219), or silica membrane purification (233) with or without (234) isopropanol precipitation.

Normalisation

To accommodate these differences in RNA sample preparation and analysis, the measurement of relative expression of transcript has evolved a means to control these variables employing a process that is termed normalisation (235). Normalisation of real-time RT-qPCR data is classically performed through the selection of a calibrant internal control gene, known as a reference gene or "house-keeping" gene. Conceptually, an ideal gene selected as an internal reference should have a constant level of expression across the tissue or cell samples used throughout the experiment, and should not exhibit altered expression with diseased or, control tissues, or indeed experimental conditions (236). Initially, ubiquitously applied reference genes were sought that could be applied across many tissue and experimental types (237). However, recent studies have shown that the expression stability of some of the commonly used reference genes, such as *B2M*, *GAPDH* and *ACTB* is not constant for all tissues or disease states (236,238).

Reference Gene Selection

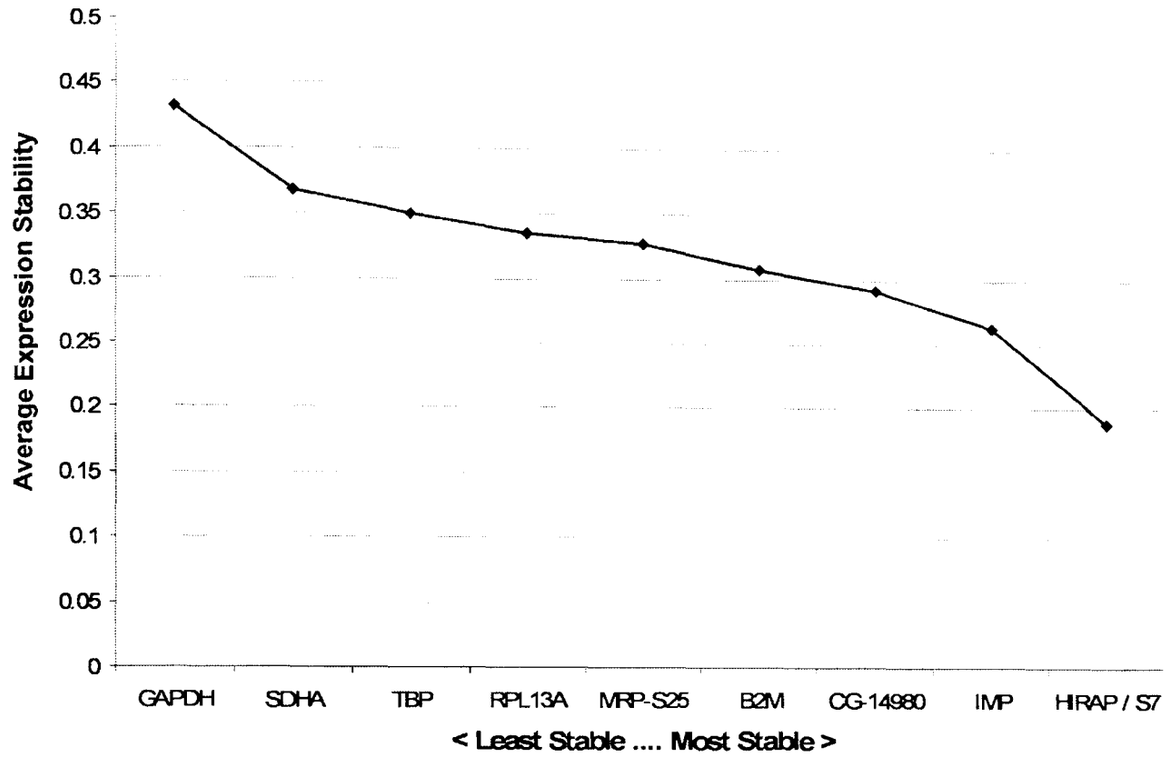
The identity of the most stable reference genes for target gene expression normalisation can be determined by evaluating data from real-time RT-qPCR statistical algorithms, such geNorm (238), Global Pattern Recognition (239), Bestkeeper (240), Normfinder (241) or equivalence tests (242).

The principal of the geNorm algorithm is that from an initial group of candidate reference genes tested across all the types of tissue studied and the experimental conditions, the expression ratio of the two reference genes that display the most similar expression identified these genes as the best choice to monitor variation in test gene expression (238). Global Pattern Recognition is a statistical algorithm which compares the expression or each gene to every other gene used in the comparison, similar to analysis of variance (ANOVA) but with exclusion of nonsensical data (e.g. threshold cycle (C_T) values of 40, where no amplification has taken place) (239). The Bestkeeper algorithm measures the geometric mean of reference gene crossing point values, to determine the optimal reference gene for use in a samples set (240). Equivalence testing is the mathematical determination of the standard deviation of differences in expression values between samples being compared (242). The Normfinder algorithm uses a model-based approach to the estimation of expression variation, which takes into account variation across sub-groups and avoids the artificial selection of co-regulated genes (241).

The use of these algorithms allows the identification of reference genes which are most stably expressed across different tissues or cells, or within the same tissue or cells within different diseases (238,243) (Figure 7). However, the identification of

new reference genes from microarray data sets, within a particular tissue type, has been demonstrated to provide more "stable" reference genes than those conventionally used (241,244-246), as determined using stability algorithms. Microarray data can be stratified on the basis of fold changes in expression (245), the variance of expression (241,246) or integrative correlations (244). Candidate genes can then be selected from stratified data, and frequently demonstrate expression stabilities greater than conventionally used reference genes (241,244,245). However microarray data has yet to identify a new reference which shows consistent stability across multiple different tissue or cell types, and / or disease situations. Therefore, a ubiquitous reference gene suitable for normalisation of gene expression of all experiments probably does not exist, but the identification of new reference genes to improve in reference gene stability is still important to reduce error in individual RT-qPCR experiments.

Figure 7
A graph demonstrating the stability of different reference genes (lower value indicates greater stability) in a tissue, as determined by the geNorm algorithm.



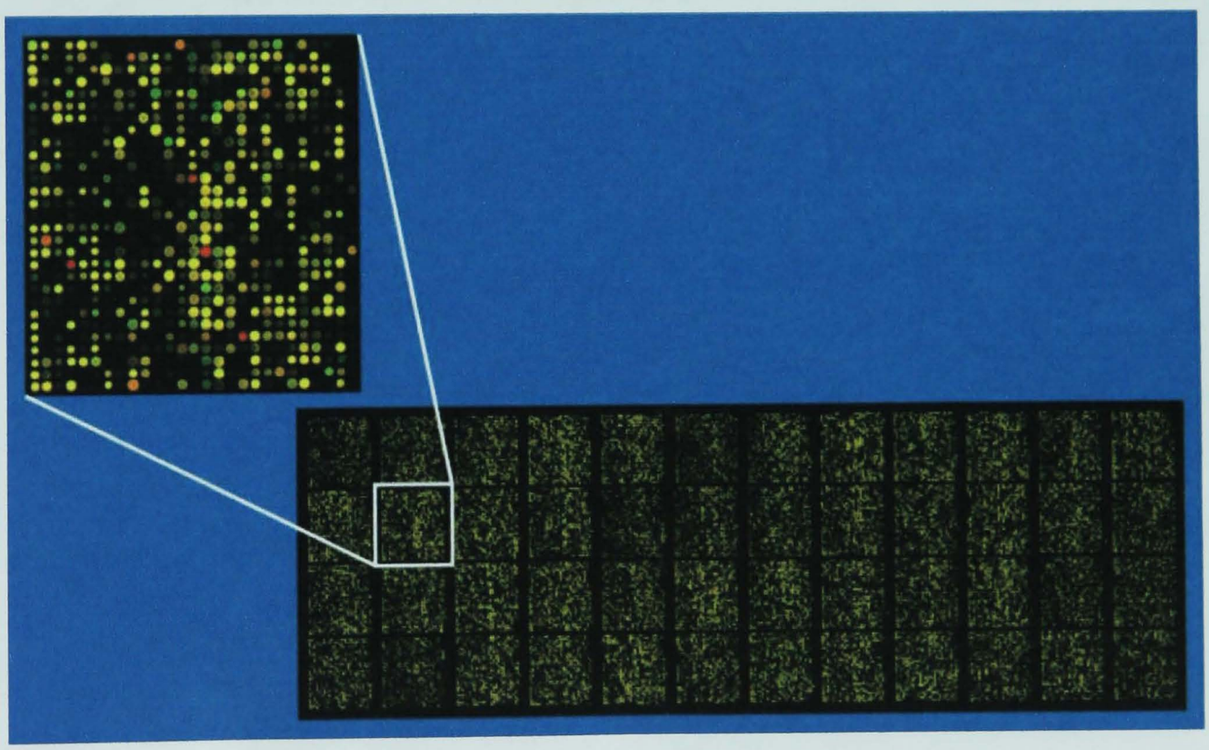
Whole Genome Microarray

The development of micoarrays, utilising either synthetic DNA oligonucleotide or cloned complementary DNA (cDNA) sequences has allowed the simultaneous evaluation mRNA expression of up to tens of thousands of genes within a single tissue sample, or even within an individual cell. The principle of the technique is straight forward; each sequence is "spotted" onto a glass slide or custom designed "chip". Tens of thousands of different DNA spots can be located on the same slide, or chip, in a known order. Messenger RNA is labelled with a dye or fluorophor, and hybridized to the array. As the spot sequences are complimentary to a specific mRNA sequence they should bind to that spot. The slide is then "read" using a laser scanner,

which quantifies the degree of fluorescence present at each spot, which in turn should be proportional to the quantity of transcript present in a sample. Whole genome microarray screens allow the assessment of expression of all known genes in a genome and are now available and utilised in canine OA research (222), whilst gene expression can be further dissected through the determination of splice variants of individual genes through the use of exon arrays.

Figures 8

A 40,000 spot microarray slide, with a small area of the slide highlighted. Note the differential fluorescence of different spots, indicated by different spot colour intensity (Bright yellow- high fluorescence, black, no fluorescence, red= control spot). Image from (247), unrestricted permission to use.



Microarray analysis of OA Cartilage

Cartilage is particularly applicable to the use of microarray techniques because it consists of a single cell population (chondrocytes) therefore gene expression levels can be attributed to this cell population alone (248).

Expression profiles of normal, early degenerate, and end stage osteoarthritic cartilage have been evaluated using a human cancer array (ClonTech Cancer 1. Array) (248). Expression profiling was checked with quantitative PCR for *COL1, COL2, COL3, AGC1*, β-actin (*ACTB*), and *GAPDH*. Type II and type III collagen expressions were up-regulated in late disease, as assessed by microarray and qPCR. Aggrecan expression was not changed by either assessment. The expression of *ACTB* was variable by either measurement, and *COL1A2* expression was up-regulated in late osteoarthritic cartilage, as measured by microarray, but was not changed as measured by qPCR.

A total of 68 genes were up or down-regulated in OA cartilage compared to normal cartilage samples in this landmark study. Genes involved with cartilage metabolism, anabolism and catabolism, were identified by this means. These may be regarded as candidate genes for human studies of OA by association, and include; *COL1A2, COL2A1, COL3A1*, type 6 collagen, alpha 1 chain (*COL6A1*), the proto-oncogene c-myc, biglycan (*BGN*), bone morphogenic protein 3 (*BMP3*), α2 macroglobulin ($\alpha^2 m$), frizzled motif associated with bone development (*FRZB*), interleukin 6 receptor alpha (*IL6Rα*), MAX dimerization protein 3 (*MAD3*), matrix metalloproteinase -2 (*MMP2*), -3 (*MMP3*), -11 (*MMP11*), Tissue Inhibitor of Metalloproteinase 4 (*TIMP4*), Tenascin (*TNC*), tumour necrosis factor receptor 1 (*TNFR1*) and Ubiquitin. Interestingly, SNPs in the genomic sequence of a number of these genes (*COL2A1* and *FRZB*) have been reported to demonstrate associations with OA in population studies (detailed previously).

A follow up study by the same group evaluated a much larger number of articular samples (n = 78) from normal, early degenerative and late stage OA using a custom-made cDNA array covering over 4000 genes (249). As previously reported a large number of matrix associated genes were differentially expressed in late stage OA cartilage. These included genes previously reported as being up-regulated, such as collagens (*COL1A2, COL2A1, COL3A1, COL5A1* and *COL9A3*) and non-collagenous proteins *BGN, CILP, COMP*, lumican [*LUM*], secreted protein, acidic, cysteine-rich [ostenectin, *SPARC*] and tenascin C [*TNC*]). The expression of the major transcription factor responsible for the chondrocytes phenotype, SRY (sex determining region Y)-box 9 (*SOX9)* (250) was decreased in end stage OA, which identifies a possible mechanism for the change in cell phenotype observed in end stage OA. Likewise, many of the genes involved in oxidative damage defence, such as Glutathione peroxidise (*GPX3*), superoxide dysmutase (*SOD2* and *SOD3*) and thioredoxin-interacting protein (*TXNIP*) were down regulated in late stage OA, suggesting an increased risk to oxidative stress damage to end stage chondrocytes. Interestingly, little difference was observed between the expression profiles of normal (healthy) and early OA articular cartilage.

Using a different complimentary DNA microarray chip, Zhang and others (251) identified 131 genes up-regulated in severely osteoarthritic human cartilage. Many of the genes up-regulated were the same as those identified by Aigner, with a number of notable additions such as Interleukin 1 (*IL1*), Interleukin 1 receptor antagonist (*IL1RA*), decorin [*DCN*], osteopontin, and beta-2-macroglobulin (*β2M*).

The results of *in vitro* cell culture experiments evaluating chondrocyte expression using microarray technology have been published (251), however doubt has been cast as to their importance, compared to assessment of *in vitro* tissue samples (252). Comparisons of the level of gene expression between cell culture and *in vitro* samples have demonstrated increased expression of similar genes, although the level of their expression can be widely different (252). Thus, although a degree of heterogeneity in expression profiling results should be expected when using clinical tissue samples, the results obtained may ultimately be more meaningful than those obtained from cell culture bases studies.

Summary

On the basis of literature review, a large number of genes are suitable for analysis in case - control studies for canine OA secondary to joint dysplasia. Narrowing down the list of potential genes, through the use of methods such as expression profiling may identify the gene polymorphisms which associate with disease phenotypes. However, given the limited sample sizes and numbers available for canine genotyping and expression profiling studies, critical variables such as sample extraction technique should also be optimised. The success and repeatability of a candidate gene study depends on both the quality of the phenotypic data provided, and the quantity of samples available.

Thesis aims and outline

This study aimed to characterise the transcriptome of canine osteoarthritis (OA) articular tissues, and to relate this to genomic changes. The transcriptomic basis of canine OA was investigated by generating expression profiles of diseased hip and elbow articular cartilage, ruptured cranial cruciate ligament and and elbow trabecular bone, using the RT-qPCR and a canine specific oligonucleotide microarray. RNA extraction from OA articular cartilage was optimised using microfluidic capillary electrophoresis and new reference genes for use in RT-qPCR experiments were identified from the oligonucleotide microarray data sets. Finally, DNA polymorphisms were identified in a group of candidate genes, using both *in silico* and *in vitro* methods, and tested for association with diseases in a case-control cohort study.

Chapter 1

Assessment of the use of RNA quality metrics for the screening of articular cartilage specimens from clinically normal dogs and dogs with osteoarthritis

Dylan N. Clements[1,2],

Anne Vaughan-Thomas[1],

Siriporn Peansukmanee[1],

Stuart D. Carter[1],

John F. Innes[1],

William E.R. Ollier[2],

Peter D. Clegg[1].

[1] Connective Tissue Research Group, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX.

[2] Centre for Integrated Genomic Medical Research, University of Manchester, The Stopford Building, Oxford Rd, Manchester, M13 9PT. UK.

Published in the *American Journal of Veterinary Research* (2006) **67**, 1438-1444

ABSTRACT

Introduction

RNA quality is important for maintaining the consistency of downstream measurements. We compared two methods of RNA purification using different quality metrics, and identified the most useful metric for quality assessment of RNA extracted from articular cartilage from dogs with osteoarthritis (OA).

Materials and Methods

RNA was extracted, from 40 articular cartilage specimens from the femoral heads of 3 clinically normal dogs and 37 dogs with OA, by 2 purification methods. Quality metrics of each sample were determined and recorded by use of a UV spectrophotometer ([Spec I] to determine the 260 nm to 280 nm absorbance [A_{260}:A_{280}] ratio), a second UV spectrophotometer ([Spec II] to determine A_{260}:A_{280} and A_{260}:A_{230} absorbance ratios), and a microfluidic capillary electrophoresis analyzer (to determine the ribosomal peak ratio [RR], degradation factor [DF], and RNA integrity number [RIN]). Metric results were compared with visual analysis of the electropherogram to determine the most useful RNA quality metric.

Results

No differences between the two methods of RNA purification were identified with quality metrics. RNA extracted from unaffected (normal) cartilage was of higher quality than that extracted from affected (osteoarthritic) cartilage, as determined by the RIN and Spec II A_{260}:A_{230} ratio. The RIN and RR were the most sensitive metrics for determining RNA quality, whereas the DF was most specific. A significant proportion (32%) of RNA samples extracted from osteoarthritic articular cartilage specimens was determined as being of low quality.

Conclusions

No single metric provided a completely sensitive and specific assessment of the quality of RNA recovered from articular cartilage. Microfluidic electrophoresis trace analysis can be used to objectively analyse RNA quality from canine articular cartilage samples.

INTRODUCTION

Transcriptomics, the quantification of RNA, is an import tool for research into gene function within biological systems (253). Techniques such as quantitative PCR (254) and microarray (248) enable the quantification of RNA transcripts in tissues, such as cartilage (255), bone (256) and muscle (257).

Quality of RNA is a measure of RNA purity and integrity. RNA integrity is of critical importance because downstream gene expression profiling may be altered by changes in integrity (223,224). No gold-standard method exists to determine the quality of RNA extracted from tissue or cell cultures. Traditionally visual assessment of an electrophoretic trace has been used as an identifier of RNA quality. Subsequently, the ribosomal band (28S:18S) ratio has also been used as an identifier of RNA quality, with a ratio > 2.0 indicative of high-quality RNA. A number of alternative metrics have been used to quantify RNA quality, such as the A_{260}:A_{280} ratio (228), A_{260}:A_{280} ratio, and ethidium bromide (229) or SYBR green dye stained agarose gel electrophoresis.

As low amounts of RNA are recovered from articular cartilage, quality assessment is ideally performed with the minimum amount of sample necessary. The recent development of microfluidic capillary electrophoresis has allowed the assessment of RNA quality (225) with low volumes of sample (1 μl) through direct trace observation and automated calculation of the 28S:18S ratio. Analysis of RNA integrity can also be performed by use of computational software based analysis of the electrophoretic trace (224,230). The degradation factor (DF) (230) is a figure calculated by use of a

mathematic model examining degradation peak signals present in the lower molecular weight range and comparing them with ribosomal peak heights. A lower number denotes a higher quality sample. The RNA integrity number (RIN) is an algorithm that calculates RNA integrity from the electrophoresis trace, by evaluation of features such as the height of the 18S peak, the ratio of the area of the ribosomal bands compared with the total area of the electropherogram, and the ratio of the fast area of the electropherogram to the total area of the electropherogram (231). A comparison of the RIN and DF of human tissue samples reported that the RIN produces the most reliable data (224).

Extraction and purification of RNA from articular cartilage is problematic. The tissue is relatively acellular, usually only available is small quantities (with clinical canine samples often < 100 mg in wet-weight), and contains a large amount of proteoglycan in the extracellular matrix (232). A variety of methods of RNA extraction from articular cartilage have been published to date. Most methods use liquid nitrogen dismembration, phenol-chloroform RNA extraction and caesium trifluoroacetate ultracentrifugation (219), or silica membrane purification (233) with or without (234) isopropanol precipitation.

The aims of the study reported here were to firstly investigate the benefit of isopropanol precipitation on the quality of RNA extracted from the articular cartilage from clinically normal dogs by use of different RNA quality metrics. Each RNA quality metric method was selected on the basis of the sample volume (i.e. 1 μL) required to complete the test. We also assessed the relationship between RNA quality metrics to identify whether the results from one method infer those of another.

Finally, a large number of RNA extractions from articular cartilage of clinically normal dogs and dogs with osteoarthritis (OA) were evaluated by use of low volume RNA quality metrics to determine the value of quality information, compared with that of conventional quality assessment (i.e. visual analysis of the electrophoretic trace). We hypothesised that the RNA quality metrics evaluating integrity could not be inferred from metrics evaluating purity (i.e. absorbance ratios). We also hypothesized that RNA quality metrics evaluating integrity could be used to differentiate samples determined as being of high or low RNA quality by visual analysis of an electrophoretic trace.

MATERIALS AND METHODS

Articular cartilage from clinically normal dogs

Extraction of RNA was performed by use of methods similar to those described by Reno (258) and Flannery (234). Articular cartilage from the femoral heads of 3 clinically normal crossbred dogs was harvested as previously described (259) by sharp dissection and stored in a storage reagent (RNAlater, Qiagen , Crawley, UK.). Articular cartilage specimens were cut into 1 mm^3 fragments and pooled. Aliquots of between 90 to 100 mg were dried, weighed, and snap frozen in liquid nitrogen. Each aliquot was separately pulverized for 2 minutes at 2000 revolutions per minute (rpm) in a liquid-nitrogen cooled dismembrator (Braun Mikro-Dismembrator Vessel, B. Braun Biotech International GmbH, Melsungen, Germany) and snap frozen in a 1.5 ml centrifuge tube. A 1 ml aliquot of monophasic solution of phenol and guanidine isothiocyanate reagent (Trizol solution, Invitrogen Ltd, Paisley, UK) was added directly to the powdered cartilage, mixed, and warmed to room temperature

(approximately 20°C) and allowed to stand for 30 minutes. Each sample was transferred to a 1.5 ml microcentrifuge tube and centrifuged at 12,000 x *g* for 10 minutes at 4°C. The supernatant was placed in a new 1.5 ml microcentrifuge tube and the cell and tissue debris discarded. Following the addition of 0.2 ml of chloroform, samples were vortexed for 15 seconds and allowed to sit at room temperature for 10 minutes then centrifuged at 12,000 x g for 15 minutes at 4°C. The upper aqueous phase was then removed. The RNA was purified by use of 1 of 2 methods (below). Once purified, RNA samples were stored at –80°C until analysis.

RNA Purification Method 1

The upper aqueous phase was mixed with 0.5 ml of isopropanol and stored at –70°C overnight, followed by centrifugation for 10 minutes at 4°C and 12,000 × *g*, removal of the supernatant, and re-suspension in 70% ethanol. The pellet was centrifuged at 7,500 x g for 5 minutes at 4°C. The 70% ethanol wash and centrifugation step was repeated, after which the supernatant was removed and the pellet re-suspended in 30µl of RNase free water. Contaminating genomic DNA was then reduced by performing DNase digestion (RQ1 RNase-Free DNase, Promega Corporation, Madison, USA). The RNA solution was then further cleaned using mini-columns and reagents according to the manufacturer's protocol (RNeasy, Qiagen), and eluted in 30 µl of RNase free water. The elution step was repeated with the elutant to maximize the amount of RNA eluted.

RNA Purification Method 2

The upper aqueous phase was removed and mixed with an equal volume of 70% ethanol. The solution was then cleaned using mini-columns and reagents according to

the manufacturer's protocol (RNeasy, Qiagen) that included an on-column DNA digestion with an RNase-free DNase set (RNase-Free DNase Set, Qiagen). Final elution was in 30 µl of RNase free water. The elution step was repeated with the elutant to maximize the amount of RNA eluted.

Articular cartilage from dogs with OA

Extraction of RNA was performed by use of methods similar to those described by Reno (258) and Flannery (234). Articular cartilage from the femoral heads of 37 dogs with OA was obtained at surgery (i.e. femoral head and neck excision, or total hip arthroplasty) and immediately (< 15 minutes) placed in storage reagent (RNAlater, Qiagen). Extraction of RNA from articular cartilage was performed on all the specimens from dogs with OA by use of method 1 or 2. Once purified, RNA samples were stored at –80°C until analysis.

Spectrophotometer measurements

Two UV spectrophotometers (Spec I and Spec II, respectively) were used to measure RNA quality. For measurement of RNA quality by use of Spec I (Ultrospec 2000, Pharmacia Biotech, Cambridge, UK), 1 µl of each sample was dissolved in 39 µl of RNase free sterile water and analyzed. The RNA concentration and Spec I A_{260}:A_{280} ratio were recorded, and the concentration adjusted to account for the dilution.

For measurement of RNA quality by use of Spec II (NanoDrop ND-1000, Labtech International Ltd, East Sussex, UK) 1 µl of each sample (undiluted) was evaluated. The RNA concentration, Spec II A_{260}:A_{280} ratio, and Spec II A_{260}:A_{230} ratio were

recorded. All samples analyzed had an RNA concentration of > 20 ng/μl, as determined by use Spec II.

Electrophoresis

For each sample, 1 μl (undiluted) was analyzed (Agilent 2100 Bioanalyser, Agilent Technologies UK Ltd, West Lothian, UK). The ribosomal band ratio (28S:18S), as calculated by a software program (260) from the area below each peak (i.e. RR), was recorded. Further RNA quality metrics, DFs (230) and RIN (224), were calculated by use of freely available software programs (260,261) from the electrophoretic traces produced. Visual assessment of the electrophoresis traces was performed independently by 2 blinded observers experienced in analyzing RNA quality. Each sample was designated high or low quality RNA. No descriptors were provided. For samples where the 2 observers did not agree (n = 2), a third blinded observer assessed the trace and the sample was assigned a designation in accordance with the majority view.

Statistical analysis

The means, standard deviations (SDs), interquartile ranges (IQR), and 95% confidence interval (95% CI) were calculated for each of the methods used, and different methods compared by the *t* test statistic, as a normal distribution was assumed because the values were obtained from the same tissue by use of the same method. Comparisons between tissues from unaffected (normal) and affected (osteoarthritic) joints were performed with the Mann Whitney U test, as values from the same tissue extracted by different methods were pooled. Comparisons between samples of high and low RNA quality were performed with the Mann Whitney U test.

Correlations between different methods of RNA quality assessment were analyzed by use of Spearman correlation coefficients. Sensitivity and specific values were determined for each test by use of arbitrary cut off values selected between the lower value of the 95% CI, for samples determined as being high quality by visual assessment, and the upper value of the 95% CI, for samples determined as being low quality by visual assessment. Significance was determined at a value of $P < 0.05$. All data analyses were performed by use of a software program (Minitab v14.1, Minitab Ltd, Coventry, UK).

RESULTS

Articular cartilage from clinically normal dogs

No significant difference was identified between the quality of RNA extracted by the different methods (i.e. Spec I A_{260}:A_{280} [$P = 0.064$]; Spec II A_{260}:A_{280} [$P = 0.800$]; Spec II A_{260}:A_{230} [$P = 0.149$]; RR [$P = 0.507$]; RIN [$P = 0.681$]; DF [$P = 0.872$]; Figure 1). No significant correlations were identified between the variables measured (Table 1).

Articular cartilage from dogs with OA

Significant positive correlations were identified (Table 2) between Spec II A_{260}:A_{280} and Spec II A_{260}:A_{230} ($P = <0.001$), RR and Spec II A_{260}:A_{280} ($P = 0.016$), RIN and Spec I A_{260}:A_{280} ($P = 0.001$), RIN and Spec II A_{260}:A_{280} ($P = 0.025$), RIN and Spec II A_{260}:A_{230} ($P = 0.017$), and RIN and RR ($P = 0.001$). Significant negative correlations were identified between DF and Spec I A_{260}:A_{280} ($P = 0.026$), DF and RR ($P = < 0.001$), and DF and RIN ($P = 0.004$).

Twelve samples were considered to be of low RNA quality and 25 samples were considered to be of high RNA quality by use of visual assessment of the electrophoretic trace. The Spec I A_{260}:A_{280} ratio could not be calculated for 13 samples, 5 of which were of low quality and 8 of which were high quality. The DF could not be calculated for 7 of the samples, all of which were considered to be of low quality (as judged by the RIN and visual assessment of the electrophoretic trace), and a RIN could not be calculated for 8 samples, 6 of which were of high quality and 2 of which were low quality. The results were stratified into 2 groups (i.e. low and high RNA quality; Figure 2). Values (median [IQR]) for each metric were as follows: Spec I A_{260}:A_{280} ratio (low quality, 1.39 [0.98 to 1.70]; high quality, 1.87 [1.72 to 2.03]), Spec II A_{260}:A_{280} ratio (low quality, 1.43 [1.02 to 1.82]; high quality, 1.91 [1.65 to 2.025]), Spec II A_{260}:A_{230} ratio (low quality, 0.31 [0.11 to 0.58]; high quality, 0.68 [0.40 to 0.93]), RR (low quality, 0 [0 to 0]; high quality, 0.8 [0.65 to 1]), DF (low quality, 27.45 [12.92 to 34.17]; high quality, 5.91 [5.01 to 7.64]), and RIN (low quality, 1.1 [1.0 to 2.6]; high quality, 7.65 [7.0 to 8.0]). Significant differences were identified between all quality metrics (i.e. Spec I A_{260}:A_{280} ratio [$P = 0.049$], Spec II A_{260}:A_{280} ratio [$P = 0.021$], Spec II A_{260}:A_{230} ratio [$P = 0.029$], RR [$P < 0.001$], DF [$P = 0.001$], and RIN [$P = 0.013$]). Significant ($P = 0.05$) differences in quantity were also identified between samples of low (27.3 ng/µl [26.0 to 47.5 ng/µl]) and high quality (43.1 ng/µl [29.3 to 175.2 ng/µl]).

Comparisons between groups

A significant difference in the quality of RNA from articular cartilage specimens from clinically normal dogs, compared with articular cartilage specimens from dogs with OA, was identified by the Spec II A_{260}:A_{230} ratio ($P = 0.031$) and RIN ($P = 0.002$),

but not by the Spec I A_{260}:A_{280} ratio (P = 0.275), Spec II A_{260}:A_{280} ratio (P = 0.661), RR (P = 0.1876), and DF (P = 0.155). No significant (P = 0.982) difference in quantity was found between articular cartilage specimens from clinically normal dogs (median, 51 ng/μl; IQR, 26 to 68 ng/μl) and articular cartilage specimens from dogs with OA (median, 35.1 ng/μl; IQR, 27 to 86 ng/μl).

RNA quality of articular cartilage from dogs with OA

Six of 25 articular cartilage specimens from dogs with OA determined as being high in RNA quality could not have a RIN determined, yet all of these samples had a DF indicating high quality (i.e. value < 10). Conversely, 8 of 12 articular cartilage specimens from dogs with OA of low RNA quality could not have a DF calculated, whereas in all samples, with the exception of 2 low RNA quality samples, it was possible to ascribe a RIN value.

Sensitivity and specificity of RNA quality metrics

The sensitivity and specificity of each metrics were compared with visual assessment of the electrophoretic trace (Table 3). The RIN and RR provided the most sensitive method for determining high RNA quality, whereas the DF was most specific. The Spec I A_{260}:A_{280}, Spec II A_{260}:A_{280}, and Spec II A_{260}:A_{230} ratios were found to be moderately informative with regard to RNA quality. By use of an algorithm (Figure 3), 36 of 37 samples could be ascribed to groups agreeing with the visual assessment of quality.

DISCUSSION

Comparisons of the two RNA extraction methods are insightful, even though there appeared to be little difference in the quality of the RNA extracted. Isopropanol precipitation of RNA allows separation of RNA from proteoglycans that partially co-purify with RNA following phenol-guanidine thiocyanate-chloroform extraction (262), and the isopropanol precipitation step essentially allows an extra cleaning procedure. However, our results indicate that there was no significant difference in RNA quality between the 2 methods for RNA purification of articular cartilage specimens from clinically normal dogs. On the basis of our results there appears to be no benefit in performing the additional step of isopropanol precipitation during RNA extraction of articular cartilage.

Extraction of RNA from articular cartilage specimens of dogs with OA was characteristically more degraded as determined by the RIN and Spec II A_{260}:A_{230} ratio than that from unaffected articular cartilage. A number of factors will contribute to these findings. Firstly samples may not have been collected and stored consistently in the optimal manner, as most were taken during a surgical procedure, where sample collection is not a priority. Secondly, markedly osteoarthritic tissue contains a proportion of cells which are apoptotic (263), and thus likely to contain degraded or degrading RNA. More rapid degradation of RNA has been observed in canine tumour specimens, compared with unaffected tissue (264). Thirdly, the articular cartilage specimens from clinically normal dogs were pooled, which would reduce the variability of the metrics assessed, thus direct comparison may not be strictly valid. Finally, the RIN metric could not determine the RNA quality of 6 samples of high

quality, and thus there is a bias in using this metric for assessment of articular cartilage specimens from dogs with OA. A comparison of the methods of RNA purification on the quality of RNA purified from osteoarthritic articular cartilage tissues was not done as no difference was found between the methods of extraction in unaffected (normal) cartilage specimens in the preliminary work, as well as the inherent variability of clinical sample quality.

No significant correlations were observed among RNA quality metrics with articular cartilage specimens from clinically normal dogs. This was probably a result of the small sample size (n = 6 values for each metric). Strong correlations were observed between the DF and the RIN and RR with articular cartilage specimens from dogs with OA. This may have occurred because each of these values was generated from the same electrophoretic trace. The RIN had significant correlations with all other metrics, whereas the RR was significantly correlated with the Spec II $A_{260:280}$ ratio, and DF was weakly positively correlated with the Spec I $A_{260:280}$ ratio. The reason for this was not clear, but probably reflects the fact that the RIN could be calculated for most of the high and low RNA quality samples. A strong positive correlation was also found between the Spec II $A_{260:280}$ ratio and the Spec II $A_{260:230}$, probably because the same value for A_{260} is used in both calculations.

A major advantage of using cultured chondrocytes for expression analysis, rather than articular cartilage specimens from dogs with OA, is that perfect quality RNA (with no loss of integrity, as determined by quality metrics) can be obtained. Monolayer cultured chondrocytes lack the large volume of extracellular matrix present in affected (osteoarthritic) cartilage specimens, most notably proteoglycans that are likely to

interfere with RNA extraction profiles. Additionally, cell recovery can be performed directly into a solution for RNA recovery from the culture, without the requirement for pulverisation. Although the culture of chondrocytes produces a "cleaner" system, their phenotype may become altered in culture (265). Comparisons of the level of gene expression between cell culture and *in vitro* samples have shown that although similar genes may be up-regulated, the level of their expression can be widely different in comparison to evaluation of in-vivo material (252).

Quantity of RNA was not significantly different between articular cartilage specimens from clinically normal dogs and those from dogs with OA. The quantities of RNA produced were all >20 ng/μl, which, although < 50 ng/μl, as recommended by the manufacturer for RNA quality assessment by electrophoresis, has been shown to produce RIN values and DFs that strongly correlate with visual assessment of quality (224). The quantity of RNA measured in affected (osteoarthritic) articular cartilage specimens of low quality was significantly lower than that measured in specimens of high quality. Quantity may contribute to the differences in RNA quality metrics identified between affected cartilage specimens of low and high quality. However, quantity should have no effect on the ultimate relative measurement of expression (266), whereas quality does (224), thus the determination of RNA quality metrics is still important.

A RIN and DF value could not be ascribed to 9 and 8, respectively, of the articular cartilage specimens from dogs with OA. This is somewhat higher than previously published results assessing these tools (224), and reflects the difficulty in obtaining RNA of sufficient quantity and quality in a clinical setting. For affected

(osteoarthritic) articular cartilage specimens, RNA quality determination required calculation of RIN and DF values. Values of RIN were generated more consistently for low quality samples, whereas DFs were generated more consistently for high quality samples.

Although gene expression profiles from partially degraded RNA samples have a high degree of similarity when compared with matching intact samples (267), results of other studies have shown that a significant decline can occur in the relative expression of genes in RNA samples of poor quality (223,224). Proposals for the minimum information about which should be presented with microarray experiments (268) request that the additional details, such as quality metrics of the original sample, (for example capillary electrophoresis), can be added to control the quality of data produced.

Significant differences were observed among RNA quality metrics, when the articular cartilage specimens from dogs with OA were stratified into high and low quality. Theoretically there is a tendency for visual assessment of the electrophoretic trace to select for integrity, rather than purity, as the trace generated by the analyzer does not identify protein contamination, unless it is bound to RNA. This is also true for agarose gel assessment of RNA with ethidium bromide. However, metrics assessing RNA purity did differentiate the samples when stratified by visual assessment of the electrophoretic trace, indicating that the tendencies for protein contamination and loss of integrity are linked.

CONCLUSIONS

On the basis of our results, we propose that the RR, RIN, and DF metrics are used to assess the quality of RNA obtained from canine cartilage. Although visual assessment of the electrophoretic trace allows a gross evaluation of RNA quality, use of graded metrics, such as the RIN, avoids individual error, and provides a graded or scaled output. By use of the algorithm presented, 97% of articular cartilage specimens from dogs with OA could be ascribed quality values that agreed with the visual analysis of the electrophoretic trace. Clearly, every effort should be made in the clinical setting to optimize articular cartilage specimen collection and storage, to maximize specimen quality, however some degree of integrity loss is to be expected, compared with cell culture systems. We recommend using multiple metrics for the accurate assessment of clinical articular cartilage specimens.

Figure 1

The microfluidic capillary gel electrophoresis traces from samples of high (middle) and low (right) quality RNA, with the marker scale (left, nt = approximate nucleotide number]).

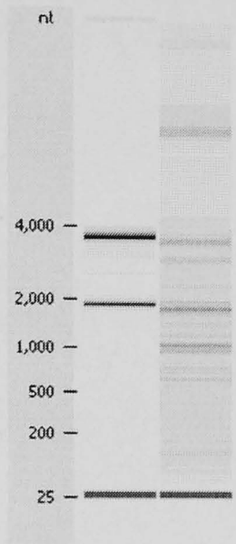


Figure 1

Quality metrics of 2 methods of RNA purification from articular cartilage specimens from clinically normal dogs. All measurements are in arbitrary units (ratios, RIN, or DF) and the values for each triplicate are displayed. M1 = Method 1. M2 = Method 2.

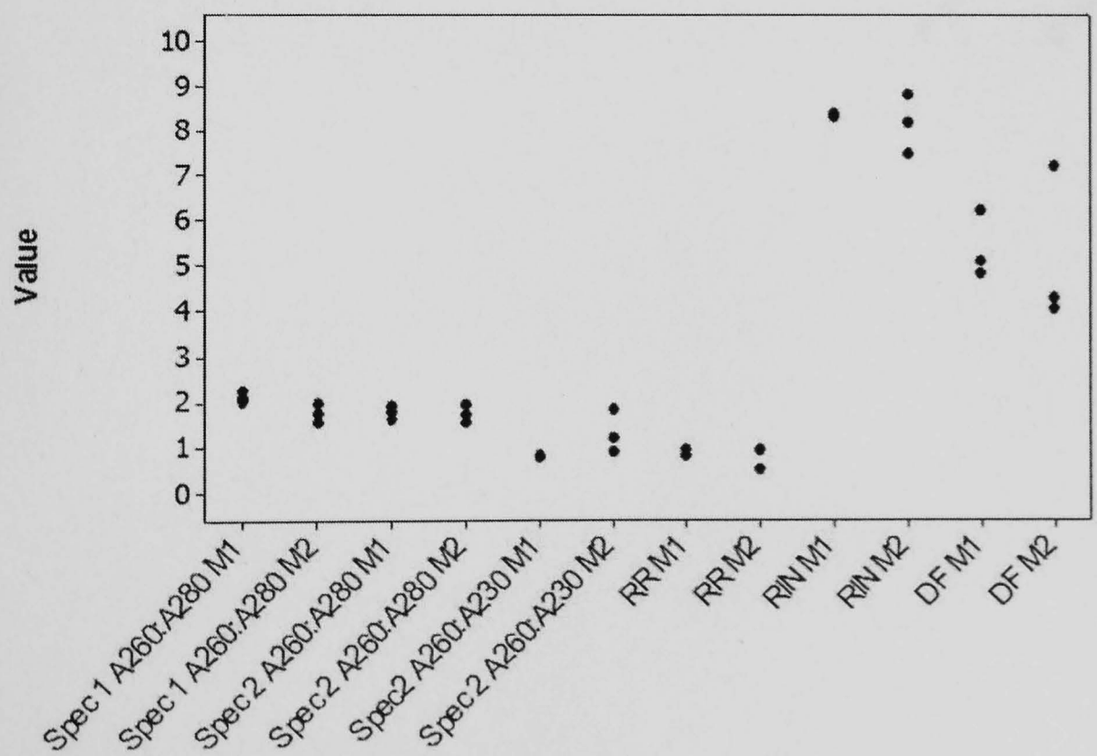


Figure 2

Box plots (A and B) of quality metrics of RNA purified from articular cartilage specimens from dogs with osteoarthritis. Box indicates IQR, line within box indicates median, whiskers indicate 95% CI, and asterisks indicate outliers. Low and high quality were differentiated by visual assessment of the electropherogram. All measurements are in arbitrary units (ratios, RIN or DF). Low = low quality. High = high quality.

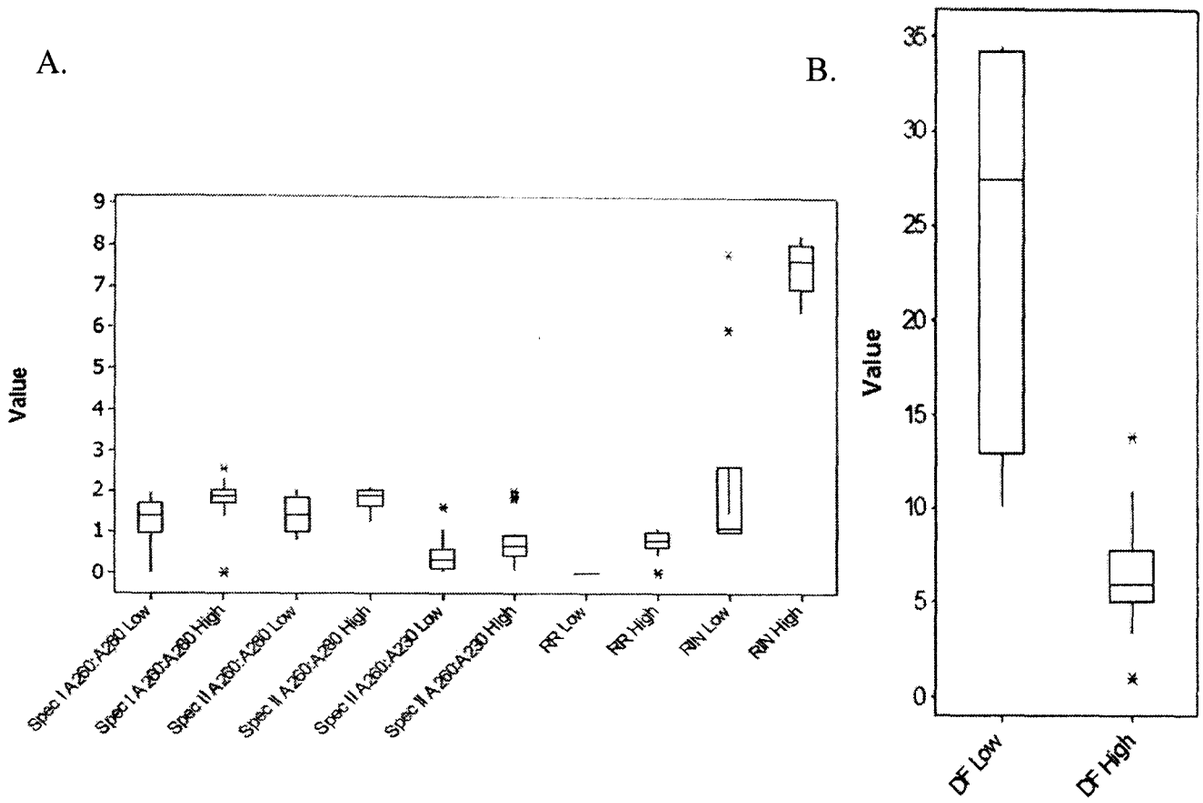


Figure 3

Proposed algorithm for determining high quality (HQ) samples from low quality samples (LQ) by use of DF, RIN, and RR.

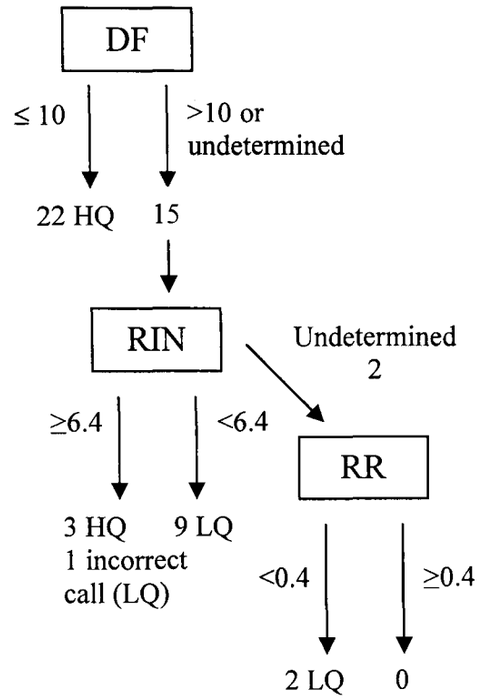


Table 1

Spearman correlation coefficients for different metrics used to assess RNA quality of articular cartilage specimens from clinically normal dogs

Variables		Metrics used to assess RNA quality				
		Spec I $A_{260}:A_{280}$	**Spec II** $A_{260}:A_{280}$	**Spec II** $A_{260}:A_{230}$	**RR**	**RIN**
Spec II $A_{260}:A_{280}$	Correlation coefficient	0.543				
	P value	0.266				
Spec II $A_{260}:A_{230}$	Correlation coefficients	–0.580	0.348			
	P value	0.228	0.499			
RR	Correlation coefficients	0.169	0.338	0.051		
	P value	0.749	0.512	0.923		
RIN	Correlation coefficient	0.232	–0.058	–0.294	0.600	
	P value	0.658	0.913	0.572	0.208	
DF	Correlation coefficient	–0.086	–0.657	–0.580	–0.507	–0.464
	P value	0.872	0.156	0.228	0.305	0.354

Table 2

Spearman correlation coefficients for different metrics used to assess RNA quality of articular cartilage specimens from dogs with osteoarthritis

Variables		Metrics used to assess RNA quality				
		Spec I A_{260}:A_{280}	**Spec II A_{260}:A_{280}**	**Spec II A_{260}:A_{230}**	**RR**	**RIN**
Spec II A_{260}:A_{280}	Correlation coefficient	0.389				
	P value	0.074				
Spec II A_{260}:A_{230}	Correlation coefficients	0.240	0.745			
	P value	0.282	<0.001			
RR	Correlation coefficients	0.278	0.461	0.297		
	P value	0.211	0.016	0.132		
RIN	Correlation coefficient	0.730	0.487	0.515	0.669	
	P value	0.001	0.025	0.017	0.001	
DF	Correlation coefficient	–0.523	–0.171	–0.124	–0.681	–0.761
	P value	0.026	0.471	0.602	<0.001	0.004

Table 3

Sensitivity and specificity of different metrics for prediction of high quality RNA, compared with visual assessment of the electrophoretic trace

	Metrics used to assess RNA quality					
Variables	**Spec I A_{260}:A_{280}**	**Spec II A_{260}:A_{280}**	**Spec II A_{260}:A_{230}**	**RR**	**RIN**	**DF**
Value	1.730	1.720	0.440	0.400	6.4	10.00
Sensitivity	81%	72%	72%	92%	100%	88%
Specificity	93%	86%	82%	96%	95%	100%

Chapter 2

Analysis of normal and osteoarthritic canine cartilage mRNA expression by quantitative polymerase chain reaction

Dylan N. Clements[1,2],

Stuart D. Carter[1],

John F. Innes[1],

William E.R. Ollier[2],

Philip J.R. Day[2].

[1]Connective Tissue Research Group, c/o Department of Veterinary Pathology, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX.

[2]Centre for Integrated Genomic Medical Research, The Stopford Building, University of Manchester, Oxford Road, Manchester, M13 9PT, UK

Published in *Arthritis Research and Therapy* (2006) **8,** R158

ABSTRACT

Introduction

The molecular basis to mammalian osteoarthritis (OA) is unknown. We hypothesised that the expression of selected proteases, matrix molecules, and collagens believed to have a role in the pathogenesis of OA, would be changed in naturally occurring canine OA cartilage when compared to normal articular cartilage.

Materials and Methods

Quantitative (real-time) reverse transcriptase-polymerase chain reaction assays were designed measuring the expression of selected matrix molecules (collagens and small leucine-rich proteoglycans), key mediators of the proteolytic degradation of articular cartilage (metalloproteinases, cathepsins), and their inhibitors (tissue inhibitors of matrix metalloproteinases). All data were normalised using a geometric mean of three housekeeping genes, and the results subjected to power calculations and corrections for multiple hypothesis testing.

Results

We detected increases in the expression of *BGN*, *COL1A2*, *COL2A1*, *COL3A1*, *COL5A1*, *CSPG2*, *CTSB*, *CTSD*, *LUM*, *MMP13*, *TIMP1*, and *TNC* in naturally occurring canine OA. The expression of *TIMP2* and *TIMP4* was significantly reduced in canine OA cartilage. The patterns of gene expression change observed in naturally occurring canine OA were similar to those reported in naturally occurring human OA and experimental canine OA.

Conclusions

We conclude that the expression profiles of matrix-associated molecules in end-stage mammalian OA may be comparable but that the precise aetiologies of OA affecting specific joints in different species are presently unknown.

INTRODUCTION

Osteoarthritis (OA) is the most common debilitating disease of mammalian joints. Clinical OA has been estimated to affect 12.1% of the human population aged 25 to 74 (269), whereas clinical OA affects up to 20% of the canine population at large (7). Canine OA usually develops secondary to an identifiable initiating cause (for example, secondary to hip dysplasia (16)), although it can be experimentally induced (85). Experimental models provide controlled and reproducible development of OA (270), but only the study of naturally occurring disease allows experimental findings to be directly related to the clinical presentation with absolute certainty. The relatedness of the pathogenesis of a common disease, such as OA, in two different species has not been characterised (271).

At present, the precise mechanisms underlying the molecular pathogenesis of OA are unknown. Quantification of gene expression is a fundamental tool for investigating gene function in biological systems, particularly for elucidating pathological mechanisms at play in diseased tissues. Quantitative reverse transcriptase-polymerase chain reaction (RT-qPCR) is currently considered the most accurate technique for quantifying gene expression. With the publication of the canine genome (218), RT-qPCR assays can now be readily designed for the measurement of canine gene expression. Although canine-specific oligonucleotide microarrays are available for the quantification of mRNA transcripts in canine tissue, such as cartilage (222), quantitative RT-qPCR validation of the results produced is still required.

Articular cartilage is composed of chondrocytes embedded in an extracellular matrix (ECM). The structural strength of the matrix is provided by collagens such as type II collagen (*COL2*), type VI collagen (*COL6*), type IX collagen (*COL9*), type XI collagen (*COL11*), and type XVI collagen (*COL16*), with *COL2* accounting for 90% to 95% of the collagen composition of the ECM. Other than water, the major non-collagenous component of articular cartilage is aggrecan (*AGC1*); smaller components include the small leucine-rich proteoglycans such as biglycan (*BGN*), chondroitin sulphate proteoglycan 2 (*CSPG2*), decorin (*DCN*), lumican (*LUM*), and tenascin C (*TNC*). The proteolytic degradation of normal and osteoarthritic cartilage matrix is performed by proteases such as the matrix metalloproteinases (*MMPs*) (272), members of the *ADAMTS* (a disintegrin and metalloproteinase with thrombospondin-like motif) family (or 'aggrecanases') (273), and lysosomal proteases (such as cathepsins) (274). Tissue inhibitors of matrix metalloproteinases (*TIMPs*) are naturally occurring inhibitors of *MMP* and *ADAMTS* function (275). The authors are unaware of any publications documenting the change in expression of structural ECM and protease collagens in the articular cartilage of dogs with naturally occurring OA. We hypothesised that the expression of selected proteases, matrix molecules, and collagens would be modulated in naturally occurring canine OA.

MATERIALS AND METHODS

Cartilage samples

Osteoarthritic articular cartilage was harvested from the femoral heads of dogs that had end-stage naturally occurring OA secondary to hip dysplasia (n = 15, mean age 2.7 years [range 1 to 12 years], mean weight 28.2 kg [range 25 to 36 kg]) and which

were undergoing routine surgical treatment of the disease (total hip replacement). In all cases, severe clinical and radiographic signs associated with OA of the affected joint necessitated surgical treatment of the disease. Articular cartilage was harvested from the area surrounding the central cartilage erosion usually observed on the canine OA hip (16). Normal articular cartilage was harvested without visual evidence of hip dysplasia or OA from the femoral heads of dogs, which had been euthanatized for reasons unrelated to joint disease (n = 13, mean age 3.3 years [range 1 to 11 years], mean weight 26.2 kg [range 15 to 40 kg]). Articular cartilage was obtained from the same site of the femoral head in the control dogs as it was in diseased dogs. Cartilage samples were immediately immersed in RNAlater™ (Ambion Ltd., Huntingdon, UK) at room temperature for 24 hours before being stored at -20°C until use, in accordance with the manufacturer's instructions.

RNA extraction from articular cartilage

Tissue samples were removed from RNAlater™ and total RNA was extracted using phenol/guanidine HCl reagents (Trizol™, Invitrogen Ltd, Paisley, UK) and isolated as previously described (258,276). An on-column DNA digestion step was included (RNase-Free DNase Set; Qiagen Ltd, Crawley, UK). Final elution of the total RNA was performed using 30 µl of RNase-free water and repeated to maximise the amount of RNA eluted.

RNA quality assessment

The concentration of total RNA of each sample was quantified by using a spectrophotometer (NanoDrop Technologies, Wilmington, USA). RNA integrity was analysed by evaluating the capillary electrophoresis trace (Agilent 2100 Bioanalyser;

Agilent Technologies, Santa Clara, CA, USA) of the sample by using the RNA integrity number (RIN) algorithm (224), degradation factor (DF) (230), and ribosomal peak ratio. The sample was determined to have minimal or no loss of integrity (RIN > 6.4 and/or DF < 10 and/or a ribosomal ratio > 0.4) and thus deemed suitable for use in the following experiments in accordance with a previously developed quality algorithm (276).

Synthesis of cDNA

Each sample was normalised to a concentration of 20 μg/μl, using RNase-free water, and reverse transcription was performed using 10 μl RNA (200 μg total RNA) with oligo-dT_{12-18} and Superscript II reverse transcriptase (Invitrogen Ltd). After reverse transcription, the template was diluted with 500 μl RNase/DNase-free water. cDNA was stored at -80°C until later use in quantitative PCR.

Quantitative PCR

Transcript sequences were obtained from the Ensembl canine genome database (277), with cross-reference to the National Centre for Biotechnology Information (Bethesda, MD, USA) (278). Where possible, assays were designed in areas of sequence showing 100% homology between predicted and verified sequences. Primer and probe sequences were designed using online design software (279). To enhance the probability of transcript-specific PCR, selected amplicon systems were designed so that the last six to seven bases of a 3′ primer or the probe crossed an exon-exon boundary. When this was not possible, the primers were designed to be hybridised on different exons, with an intronic sequence greater than 1,100 base pairs, to maintain specificity for mRNA. Some assays could be designed within only a single exon, and

thus a genomic DNA assay was also designed to determine whether genomic contamination was present. BLAST (Basic Local Alignment Search Tool) searches were performed for all primer sequences to confirm gene specificity.

Genes were selected for assay on the basis of their importance to cartilage homeostasis or pathology as derived from a literature review of naturally occurring human OA and experimental canine OA and from the results of a preliminary canine-specific whole genome microarray study, using a small number of samples. Assays were designed for quantification of expression of five collagen genes (type I collagen, alpha 2 chain [*COL1A2*], type II collagen, alpha 1 chain [*COL2A1*], type III collagen, alpha 1 chain [*COL3A1*], type V collagen, alpha 1 chain [*COL5A1*], and type IX collagen, alpha 3 chain [*COL9A3*]), seven ECM genes (*AGC1*, *BGN*, *CSPG2*, *DCN*, *LUM*, and *TNC*), an intermediate filament (vimentin), proteases and their inhibitors (*ADAMTS5*, cathepsins B [*CTSB*], cathepsin D [*CTSD*], *MMP13*, *TIMP1*, *TIMP2*, and *TIMP4*), and genomic DNA. Assays for four reference genes (glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], TATA box binding protein [*TBP*], ribosomal protein L13a [*RPL13A*], and succinate dehydrogenase complex, subunit A [*SDHA*]) (Table 1) were also designed. The reference genes used were selected from a panel of reference genes by applying a gene stability algorithm (238). Primers were synthesised by MWG Biotech (London, UK). Locked nucleic acid fluorescence resonance energy transfer probes with a 5′ reporter dye FAM (6-carboxy fluorescein) and a dark quencher dye were synthesised by Roche Diagnostics Ltd (Lewes, West Sussex, UK).

The quantitative (real-time) PCR assays were all performed in triplicate using a TaqMan™ ABI PRISM 7900 SDS (Applied Biosystems, Foster City, CA, USA) in 384-well plate format. Each assay well had a 10 µl reaction volume consisting of 5 µl 2X PCR master mix with Uracil N-Glycosylase (Universal PCR Mastermix; Applied Biosystems), 0.1 µl each of 20 µM forward and reverse primers, 0.1 µl of 10 µM probe (Exiqon; Roche Diagnostics Ltd), and 4.7 µl of sample cDNA (templates) or water (negative controls).

The amplification was performed according to a standard protocol with 10 minutes at 50°C followed by 40 cycles of 95°C for 1 minute and 60°C for 15 seconds, as recommended by the manufacturer (Applied Biosystems). Real-time data were analysed by using the Sequence Detection Systems software, version 2.2.1 (Applied Biosystems). The detection threshold was set manually at 0.05 for all assays. Standard curves were generated for each assay, to confirm that all assays were generated within acceptable limits (efficiency $93\% > x > 107.4\%$) and R^2 values ($R^2 > 0.98$) (with the exception of the genomic contamination assay, in which efficiency was lower, but the detection of any transcript was deemed unacceptable).

Data analysis

The weights and ages of the patients were normally distributed and thus compared with the calculation of means and Student *t* tests. The weight of the articular cartilage samples and quantity of RNA extract were compared using median values and Mann-Whitney *U* tests because the data were not normally distributed.

Real-time data were analysed by generation of mean threshold cycle (C_T) values from each transcript in triplicate. Geometric means (238) were calculated for the combined three reference genes (*GAPDH*, *TBP*, and *RPL13A*) and used to calculate the $\Delta\Delta C_T$ (delta-delta C_T) values and the relative amount of each target gene (280). A fourth reference gene (*SDHA*) was not included as a reference gene, because it was found to have differential expression between normal and OA samples, even when included as part of the normalisation calculation. The upper detection limit of dynamic range generated from the standard curves was used as a cut-off point, above which real-time data were discarded (that is, included in the statistical analyses as zero/no transcript present).

Data were compared with the calculations of means, standard deviations, and fold changes from normal and paired two-tailed *t* tests (body weight and age) performed in a spreadsheet program (Microsoft Excel 2003; Microsoft Corporation, Redmond, WA, USA) and the calculation of graphs, 95% confidence intervals (CIs) of the mean, and Mann-Whitney *U* tests (to compare the amount of each target) performed in a statistical analysis software package (Minitab version 14.1; Minitab Ltd., Coventry, UK). One-sided power calculations were performed, assuming normality from the two samples with unequal variance and using a freely available web-based program [22]. Significance was established at $P < 0.05$ and a robust statistical analysis was assumed to have a power value greater than or equal to 80%. Data were checked for errors due to multiple hypothesis testing by using the Benjamini and Hochberg false discovery rate (FDR) (281).

RESULTS

There were no significant differences between the ages (mean control 3.3 years [± 3.2 years, range 1 to 12 years], mean OA 2.7 years [± 3.1 years, range 1 to 11 years], P = 0.768) or body weights (mean control 26.2 kg [± 8.0 kg, range 15 to 32 kg], mean OA 28.3 kg [± 3.8 kg, range 23 to 36 kg], P = 0.109) of the dogs in the diseased and control groups. There was no significant difference between the weight of the cartilage samples (median control 103 mg [range 45 to 260 mg], median OA 92 mg [range 40 to 192 mg], P = 0.817) or the quantity of RNA extracted, as determined by spectrophotometer (median control 35 ng/µl [range 26 to 339 ng/µl], median OA 42 ng/µl [range 22 to 247 ng/µl], P = 0.788).

Expression values are presented in Table 3. Two genes were determined to have significant down regulation (*TIMP2* and *TIMP4*) in canine OA cartilage. One gene was determined to be significantly down regulated (*SDHA*) but with a low power value (72%); this gene was excluded after FDR correction. Ten genes were determined to be significantly up regulated in the OA samples (*BGN*, *COL3A1*, *COL5A1*, *CSPG2*, *CTSB*, *CSTD*, *LUM*, *MMP13*, *TIMP1*, and *TNC*). Furthermore, in OA, three genes were determined to be up regulated (*COL1A2*, *COL2A1*, and *COL9A3*) but with low power values (74%, 78%, and 63%, respectively) and one gene was excluded after FDR correction (*COL9A3*).

No amplification of genomic DNA was observed for any of the samples. The average standard deviation for the triplicates in each assay was 16.9% (range 7.3% to 37.9%), indicating that all assays were reproducible. Eleven of the 2,592 data points were

removed because they were assumed to be aberrant (markedly different from the other two values in the triplicate). All 'no template' control wells (n = 864) revealed no signal. Fold gene expression changes are illustrated in Figures 1 and 2, with all data normalised to the mean of the control values (with a fold change of 0 being no change, a fold change of 1 meaning a doubling of expression, and a fold change of -0.5 meaning a halving of expression). Statistical and power calculations are reported in Table 3.

DISCUSSION

Quantitative (real-time) RT-PCR is the most sensitive technique for the determination of mRNA transcript number (223). To maximise the precision of our data, we included only mRNA samples that had been determined as being of high quality (using an algorithm determined by previous work (276)), because mRNA degradation can affect assay performance (223). Assays were optimised within specific limits of efficiency, and the dynamic range of each assay was determined, used, and presented with the expression data. Additionally, we corrected our results for multiple hypothesis testing (reducing the opportunity for making a statistical type II error) and present power values, allowing an interpretation of the strength of each significant up- or down regulation.

If variables such as the methods of mRNA extraction, RNA quality assessment, reverse transcription, assay design, measurement of genomic contamination, standard curve data generation, reference gene selection, and data normalisation were presented in the 'Materials and methods' and 'Results' sections of manuscripts using

quantitative PCR, more appropriate comparison of results between different studies could be made. The geometric mean of three reference (housekeeping) genes was used in this study to reduce the variability associated with the use of a single reference gene. Geometric mean methodology has been validated as a more accurate normalisation technique than that using a single reference gene, if the reference genes are selected through the use of a stability algorithm (238), although in this study one of the genes identified by the algorithm (*SDHA*) was not stabily expressed (Table 3).

Gene expression varies with both the site of cartilage harvest (282) and the degree of cartilage degeneration (283) in the OA joint. We attempted to minimise this variability by using end-stage OA, age- and weight-matched samples, and stringent RNA quality control. A relatively high degree of heterogeneity (large 95% CIs) was observed in the level of gene expression measured from the clinical samples in this study, even existing between samples within the same group. This may reflect differences in dog age and/or breeds or variation in the time from surgical removal to collection in the preservative fluid. The analysis of additional samples or the phenotyping and selection of samples through histological grading may have increased the statistical powers of each of these differences observed, as the severity of OA measured by histology (Mankin score) correlates with a reduction in the expression of *COL2* and *AGC* (284).

Cell culture-based biological systems provide a more controlled methodology for evaluating gene expression when compared with *in vivo* tissue. For example, increased cell numbers can be obtained, breed and age factors can be eradicated, and the absence of ECM facilitates the extraction of higher quality of mRNA (224). This

is particularly true for studies of smaller mammals such as the dog, in which clinical samples of osteoarthritic cartilage may be less than 100 mg in size. However, cell-based models may differ in both gene expression profiles (252) or cell phenotype (285) with *in vitro* tissue. Ultimately, our understanding of the molecular pathogenesis of OA requires relating changes observed with *in vitro* experimentation to those identified from clinical tissue.

The paucity of literature reporting changes in gene expression observed in naturally occurring canine OA implies that often this is not easy to quantify. In part, this reflects the difficulties associated with the use of clinical tissue samples, as noted above, and the fact that the technology required to enable the economic evaluation of gene expression across large groups of tissue samples is only just becoming available. Indeed, we were limited by sample quantity, quality, and cost and needed to rationalise our list of genes selected for evaluation, as discussed previously.

We document marked elevation of expression in genes encoding for collagen synthesis in the articular cartilage of dogs with end-stage OA, which concurs with the findings in early experimental canine OA (286-289). *COL1A2*, *COL3A1*, and *COL5A1* are characteristically synthesised by cells with a fibrocartilaginous phenotype (290) as frequently seen in cartilage repair.

The increased expression of *BGN*, *CSPG2*, *CTSB*, *LUM*, *MMP13*, and *TNC* is consistent with previous studies of expression of these genes in both naturally occurring human (251,255,291-293) and experimental canine OA (286-288). The biological significance of fold changes in gene expression between control and OA

samples is unknown in the absence of additional data such as gross, radiographic or histological scoring, or protein quantification. Likewise, the changes in gene expression documented do not specify whether these changes are causal or simply associated with the development of pathology in the OA joint.

We documented decreases in the expression of *TIMP2* and *TIMP4* and an increase in the expression of *TIMP1* in canine OA cartilage. The decrease in *TIMP4* expression was consistent with expression profiles of human OA cartilage (255), although *TIMP1* expression has been documented as being decreased and *TIMP2* expression has been documented as being unchanged in human OA (255). Direct comparison of gene expression levels with those measured in other joints and/or in different species may be of limited value because the underlying aetiologies to the development of OA may differ. However, the evaluation of structural matrix components and proteases affecting those components is still of considerable interest. Furthermore, it is becoming increasingly apparent that the end-stage pathology characterising canine OA mimics that described for human OA (5).

CONCLUSIONS

On the basis of the results we present, the gene expression of selected matrix molecules and key mediators of the proteolytic degradation of articular cartilage is changed in end-stage, naturally occurring OA of the canine hip. The patterns of gene expression change are broadly similar to those reported in experimental canine stifle OA and naturally occurring human OA.

Figure 1

*Graph illustrating the means and 95% confidence intervals (CIs) of the gene expression profiles. To normalise values, the mean of each control group has been used to normalise and produce fold changes in expression. The results of the COL9A3 transcript are omitted because the 95% CIs were very high. *Significant difference.*

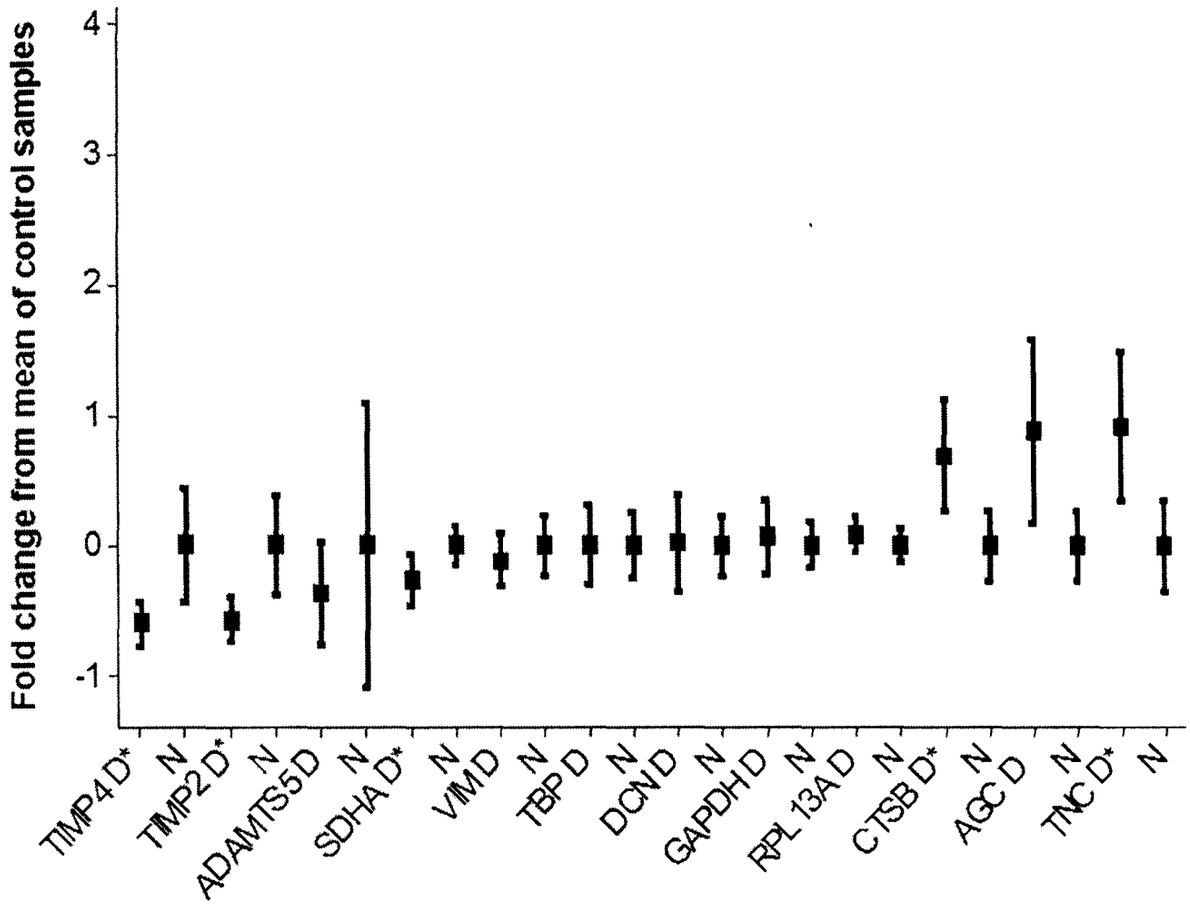


Figure 2

*Graph illustrating the means and 95% confidence intervals (CIs) of the gene expression profiles. To normalise values, the mean of each control group has been used to normalise and produce fold changes in expression. The results of the COL9A3 transcript are omitted because the 95% CIs were very high. *Significant difference.*

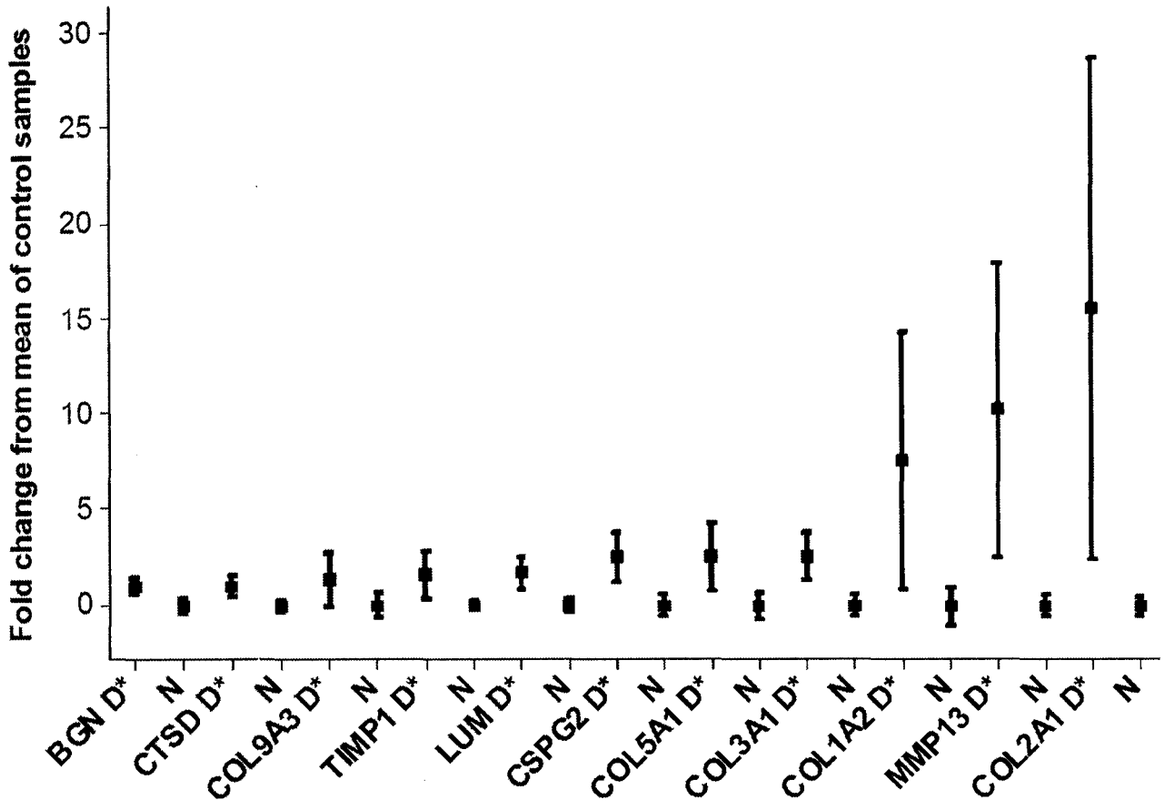


Table 1

A list of primer and probe sequences for the genes evaluated.

Gene	*Forward*	*Reverse*	*Probe*
ADAMTS5	TGGGTTCCCAAATATGCAG	CTGTCCCATCCGTCACCT	CTGGGAGA
AGC1	GGGACCTGTGTGAGATCGAC	GTAACAGTGGCCCTGGAACT	AGGAGCTG
BGN	CAGAACAACGACATCTCAGAGC	TCACCAGGACGAGAGCGTA	CTCCACCA
COL1A2	CTATCAATGGTGGTACCCAGTTT	TGTTTTGAGAGGCATGGTTG	GCCTGCTG
COL2A1	CTGGTGAACCTGGACGAGAG	ACCACGATCACCCTTGACTC	CCTCCTGG
COL3A1	GGATGGTGGCTTCCAGTTT	CCAGCTGGACATCGAGGA	GCTGCCTG
COL5A1	AACCTGTCGGATGGCAAGT	CAGTCCAAGATCAAGGTGACAT	CAGCATCC
COL9A3	CGAGGTGCCTCAGGTGAC	ACCCAGCTCTCCTTTGTCC	GAGACCAG
CSPG2	TGGATGGTTTTAATACGTTCAGG	GCCGTAGTCACACGTCTCTG	CTGCCTTC
CTSB	CGGCCTTCACCGTGTACT	GTGACGTGCTGGTACACTCC	CTTCCTGC
CTSD	GGTCCACATGGAGCAGGT	TATGAGGGAGGTGCCTGTGT	TGGGCAGC
DCN	CGCTGTCAGTGCCATCTC	GGGGGAAGATCTTTTGGTACTT	TCCAGTGT
GAPDH	CTGGGGCTCACTTGAAAGG	CAAACATGGGGGCATCAG	CTGCTCCT
Genomic	AACCCTCAAAGATGAGGTTTAGC	ACTCTGGGATCACGCATGT	CTGCCTTC
LUM	ACCTGGAAATTCTTTTAATGTATCATC	CGGTATGTTTTTAAGCTTATTGTAGGA	TGCTGGAG
MMP13	CCGCGACCTTATCTTCATCT	AACCTTCCAGAATGTCATAACCA	AGAGGCAG
RPL13A	CTGCCCCACAAGACCAAG	GGGATCCCATCAAACACCT	CCAGGCTG
SDHA	GGTGGCACTTCTACGACACC	ATGTAGTGGATGGCGTCCTG	CTGGCTGG
TBP	TCCACAGCCTATCCAGAACA	CTGCTGCTGTTGTCTCTGCT	CTGGAGGA
TIMP1	TGCATCCTGCTGTTGCTG	AACTTGGCCCTGATGACG	CCCAGCAG
TIMP2	ATGGGCTGTGAGTGCAAGAT	CACTCATCCGGAGACGAGAT	CTGCCCCA
TIMP4	GCAGAGAGAAAGTCTGAATCATCA	GGCACTGTATAGCAGGTGGTAA	TGTGGCTG
TNC	TGGATGGGACAGTCAAGGA	GCTCAGCTCTGCCAGGTTA	CCACCTCC
VIM	TACAGGAAGCTGCTGGAAGG	CCTCAGGTTCAGGGAAGAAA	GAGCAGGA

Table 2

The dynamic range, standard curve slope, R^2 value, and efficiency of each polymerase chain reaction assay.

Assay	*Lower detection limit (C_T value)*	*Upper detection limit (C_T value)*	*Standard curve slope*	*R^2 value*	*Efficiency*
ADAMTS5	26.0	35.9	-3.32	0.99	100.2
AGC	18.5	34.7	-3.29	0.99	101.5
BGN	20.8	34.8	-3.49	1.00	93.3
COL1A2b	17.4	33.5	-3.30	1.00	101.0
COL2A1	22.7	32.2	-3.22	1.00	104.6
COL3A1	16.5	33.0	-3.33	1.00	99.9
COL5A1	23.2	33.1	-3.31	1.00	100.5
COL9A3	26.3	32.7	-3.22	1.00	104.8
CSPG2	21.4	34.3	-3.25	1.00	103.2
CTSB	19.7	32.6	-3.24	1.00	103.3
CTSD	24.1	34.2	-3.29	1.00	101.5
DCN	19.0	31.9	-3.25	1.00	103.0
GAPDH	22.7	35.2	-3.27	0.99	102.3
Genomic	16.8	40.0	-4.42	1.00	68.3
LUM	19.9	33.7	-3.48	1.00	93.9
MMP13	26.1	36.3	-3.36	0.98	98.6
RPL13A	18.6	32.1	-3.36	1.00	98.6
SDHA	21.6	34.6	-3.26	1.00	102.5
TBP	16.5	30.0	-3.39	1.00	97.4
TIMP1	22.6	33.1	-3.48	1.00	93.7
TIMP2	21.8	32.1	-3.43	1.00	95.7
TIMP4	29.5	35.8	-3.16	0.99	107.4
TNC	20.1	33.0	-3.26	1.00	102.5
VIM	15.8	32.7	-3.35	1.00	98.8

Table 3

Change in gene expression, mean $2^{-\Delta\Delta CT}$ values, significance and power of comparisons between normal and OA canine articular cartilage.

Gene	*Number of values included in the analysis*	$2^{-\Delta\Delta CT}$ *normal*	$2^{-\Delta\Delta CT}$ *OA*	*Fold change in expression (diseased versus normal)*	*Mann-Whitney* U *test P value*	*Power*
TIMP4	27	0.109	0.043	-0.608	0.0094	0.859
TIMP2	28	3.959	1.664	-0.580	0.0020	0.844
ADAMTS5	16	0.031	0.019	-0.551	0.8478	0.175
SDHA	28	0.323	0.234	-0.275	0.0476	0.722
VIM	28	32.742	28.909	-0.117	0.5493	0.195
TBP	28	0.106	0.106	0.001	0.8178	0.051
DCN	28	73.034	74.253	0.017	0.5190	0.059
GAPDH	28	1.548	1.648	0.064	0.9633	0.105
RPL13A	28	7.048	7.722	0.096	0.3814	0.275
CTSB	26	0.280	0.476	0.698	0.0060	0.886
AGC	28	0.082	0.155	0.887	0.1670	0.778
TNC	28	2.700	5.205	0.927	0.0099	0.886
BGN	28	15.511	30.984	0.998	0.0043	0.976
CTSD	28	0.148	0.295	0.999	0.0066	0.944
COL9A3	27	0.231	0.546	1.365	0.0304	0.633
TIMP1	28	0.551	1.468	1.663	0.0008	0.853
LUM	28	1.635	4.476	1.738	0.0015	0.991
CSPG2	27	0.079	0.279	2.530	0.0005	0.981
COL5A1	28	0.615	2.188	2.555	0.0069	0.887
COL3A1	26	10.573	37.867	2.581	0.0011	0.982
COL1A2b	28	0.805	6.941	7.621	0.0043	0.737
MMP13	26	0.014	0.161	10.322	0.0010	0.857
COL2A1	27	1.412	23.583	15.705	0.0001	0.779

Chapter 3

Microarray Analysis of Gene Expression in Osteoarthritic Canine Articular Cartilage

Dylan N. Clements[1,2],

Stuart D. Carter[1],

William E.R. Ollier[2],

Philip J.R. Day[2,3].

[1]Connective Tissue Research Group, c/o Department of Veterinary Pathology, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX.

[2]Centre for Integrated Genomic Medical Research, The Stopford Building, University of Manchester, Oxford Road, Manchester, M13 9PT.

[3]ISAS - Institute for Analytical Sciences, Bunsen-Kirchhoff-Str. 11, 44139 Dortmund, Germany

Not submitted for publication

ABSTRACT

Introduction

Osteoarthritis (OA) is a common and debilitating disease of dogs which frequently affects the canine hip joint. The development of species-specific oligonucleotide microarray technology allows genome wide analysis of gene expression in tissues, such as articular cartilage, and thus may provide further insights into the molecular events which underlie the development of hip OA in dogs.

Material and Methods

Articular cartilage from the femoral heads of five dogs with end stage OA and five dogs with normal articular cartilage was harvested, the messenger RNA (mRNA) extracted, double amplified, labelled and hybridised to a 44000 gene canine whole genome oligonucleotide microarray. Selected genes were also quantified using the reverse transcriptase quantitative real-time PCR (RT-qPCR).

Results

In the OA articular cartilage, a total of 2866 transcripts were differentially expressed when compared to the normal articular cartilage. The expression of 1956 transcripts was identified as being increased in OA cartilage and the expression of 910 transcripts were identified as being decreased in OA cartilage. A number of genes involved in progenitor cell activity demonstrated increased expression in OA cartilage, and selected genes involved in preventing apoptosis and cell senescence demonstrated reduced expression in OA cartilage. No genes were differentially expressed on OA cartilage when the data was corrected for multiple hypothesis testing.

Discussion

Gene expression profiling of OA articular cartilage identified a number of genes not previously associated with the disease. However, the high degree of heterogeneity observed in the expression profile data generated from both oligonucleotide microarray and RT-qPCR hampered subsequent interpretation. The study highlights the limitations of expression profiling small sample sets with limited phenotype stringency.

INTRODUCTION

Osteoarthritis (OA) is the most common disease of mammalian joints, and is characterised by articular cartilage wear and degeneration, resulting in pain and dysfunction of the affected joint. Osteoarthritis commonly affects the canine hip, where it occurs most often in association with a dysplasia of the joint. Canine hip dysplasia is characterised by laxity of the affected joint, leading to hip subluxation (8), synovitis, articular cartilage wear (16), malformation of the femoral head, neck and acetabulum and pain with associated physical disability. The disease invariably results in the development and progression of OA of the affected joint. Estimation of the prevalence of clinical hip dysplasia varies between 4.2% to 9.6% for clinical signs (33) and between 10% and 73% (8,34-36) for radiographic prevalence. Canine hip dysplasia is both the primary reason for training rejection and the most common reason for ending active service of military working dogs (37,38), which highlights both the financial and welfare importance of this disease to owners of affected dogs.

Quantification of gene expression within normal and diseased tissues provides information regarding the molecular mechanisms which characterise a disease. With the publication of the canine genome (218), global analysis of gene expression can be performed using canine-specific oligonucleotide microarrays (222). Such technology has already been used to determine the gene expression pathways which underlie cartilage degeneration in human OA (249), chondrocyte de-differentiation *in vitro* (294) and the response of chondrocytes to exogenous cytokines (295), and the response of cartilage explants to loading (222).

The importance of naturally-occurring disease for the evaluation of the molecular mechanisms which underpin disease is highlighted by the fact that *in vitro* cell based models of OA may differ in both cell phenotype (285) and the associated gene expression profiles (252) when compared to *in vivo* tissue. We have previously reported on the expression profiles of selected matrix-associated genes in naturally occurring canine OA secondary to hip dysplasia using the reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) (296) which is the current gold standard method of mRNA transcript quantification. In this manuscript, we report the results of global gene expression profiling of canine osteoarthritic articular cartilage.

MATERIALS AND METHODS

Sample Collection

Articular cartilage was harvested from the femoral head of five skeletally mature Labrador Retrievers undergoing total hip replacement for the treatment of end-stage OA secondary to canine hip dysplasia (mean age 3.3 years +/-4.8 years [range 1-12 years], mean weight 31.1 kg +/-2.1 kg [range 28-33.5], one male, one male neutered, one female, two female neutered). Normal canine hip articular cartilage was harvested by sharp dissection and from the femoral head of five Labrador Retriever dogs (mean age 3.4 years +/-4.2 years [range 1-11 years], mean weight 28. 2kg +/-2.8 [range 25-32 kg], one male, one male neutered, two female, one female neutered) without any macroscopic evidence of hip OA and which were euthanatized for reasons unrelated to orthopaedic disease. Articular cartilage was harvested from the area surrounding the central cartilage erosion usually observed on the canine OA femoral head (16), and from the same position in normal dogs. There were no statistically significant

differences in the age or weights of the patients ($P = 0.444$, $P = 0.136$ respectively), as determined by Mann-Whitney U test statistics.

Articular cartilage was obtained from the same site of the femoral head in the control dogs as it was in diseased dogs. Cartilage samples were immediately immersed in RNAlater™ (Ambion Ltd, Huntingdon, UK) at room temperature for 24 hours before being stored at -20°C until use, in accordance with the manufacturer's instructions.

RNA extraction and quality assessment

Articular cartilage samples were removed from RNAlater and total RNA was extracted using phenol / guanidine HCl reagents (Trizol, Invitrogen Ltd, Dorset, UK) and isolated, as previously described, with the inclusion of an on-column DNA digestion step (258,276). Total RNA of each sample was quantified using a spectrophotometer (NanoDrop technologies Ltd, Utah, USA) and RNA integrity was evaluated with a capillary electrophoresis trace (Agilent 2100 Bioanalyser, Agilent Technologies, California, USA) of each sample using the RNA integrity number [RIN] algorithm (224), the Degradation Factor [DF] (230) and ribosomal peak ratio in accordance with a previously developed quality algorithm (276). All RNA samples were determined to be of high quality, and therefore deemed suitable for downstream use to determine mRNA expression in tissues.

RNA Amplification

A set quantity (200 μg) of Messenger RNA (mRNA) was amplified for each sample using a commercially available kit (Ambion T7 MEGAscript high yield transcription kit, Ambion (Europe) Ltd, Huntingdon, Cambridge, UK) as previously described

(297). A second round mRNA amplification was performed using MessageAmp aRNA Amplification kit (Ambion [Europe] Ltd), as described by the manufacturer. The amplified RNA (aRNA) was quantified using a spectrophotometer. For full details of the RNA amplification procedure see the supplementary material at the end of this chapter

aRNA Labelling

Two micrograms of aRNA was labelled with Cyanine-3dCTP (Cy3) or Cyanine-5dCTP (Cy5), using a fluorescent dye labelling kit (Agilent Technologies UK Ltd, South Queensferry, UK) as described by the manufacturer. Fluorescent dye incorporation was determined using a spectrophotometer, ensuring that >750 ng cRNA was labelled, and that the label incorporation was > 8pmol per μg RNA. Samples were stored at -80°C until use.

Microarray hybridization and slide reading

750 ng of both Cy3 and Cy5 cRNA was fragmented and hybridized to a canine-specific, custom designed, whole genome 44219 spot 60mer oligonucleotide microarray chip (298) at 65°C for 17 hours using the manufacturer's protocol (Agilent Technologies UK Ltd). Slides were washed according to the manufacturer's instructions, read using an Agilent DNA Microarray Slide Reader, and fluorescence data extracted by employing the Agilent Feature Extraction 8.5 software (Agilent Technologies UK Ltd). DNA spots were automatically located and subtracted from the intensity of the local backgrounds. Where intensities of the spots were below set thresholds, data was discarded from further analysis. Spots were flagged if they exhibited poor hybridization signals or when they were saturated.

Data Normalisation and Statistics

Data were imported into Genedata Expressionist Analyst (Genedata AG, Basel, Switzerland), and the Cy3 and Cy5 fluorescence weighted linear least squares regression (LOWESS) (299). Ultimately 41519 spots (93.8%) coding for transcripts were considered acceptable for application in data analysis. Expression data were then exported into Excel 2003 and comparisons between groups were achieved using paired Student t-tests. Correction for multiple hypothesis testing was performed using the Benjamini and Hochberg false discovery rate programme (FDR) (281). Corrected *P* values were calculated by dividing the true *P* value by the individual correction factor and multiplying by 0.05.

The gene annotation for each transcript was checked by manual search of the microarray oligonucleotide sequence using the basic local alignment search tool (BLAST) (278). Transcript function was then further annotated by evaluation of the gene description in Entrez Gene and Pubmed (278). Finally, each annotated transcript was checked for associated function in OA by searching the major publications documenting the microarray analysis of gene expression in human osteoarthritic articular cartilage (249,300), and Pubmed (278).

Clustering

The normalised microarray data for fifty genes differentially expressed between OA cartilage and normal cartilage, with complete annotation, were loaded into a gene clustering software program (Cluster, Eisen Labs (301)). Data was log transformed and genes centred to the mean. Hierarchical clustering of differentially expressed genes was then performed for arrays and genes using Spearmans Rank Correlation

and complete linkage link clustering. Clustering of genes and arrays were visualised with publicly available software (TreeView, Eisen Labs (301)). Genes whose expression profiles were linked as demonstrated by the clustering algorithm were checked for pathway linkage by searching a pathway database (302).

RT-qPCR Assays

The expression profiles of 20 selected matrix associated genes (five collagen genes (type I collagen, alpha 2 chain [*COL1A2*], type II collagen alpha 1 chain [*COL2A1*], type III collagen alpha 1 chain [*COL3A1*], type V collagen alpha 1 chain [*COL5A1*], and type IX collagen alpha 3 chain [*COL9A3*]), seven ECM genes (aggrecan [*AGC1*], biglycan [*BGN*], chondroitin sulphate proteoglycan 2 [*CSPG2*], decorin [*DCN*], lumican [*LUM*], and tenascin C [*TNC*]), an intermediate filament (vimentin [*VIM*]), proteases and their inhibitors (A disintegrin and metalloproteinase with thrombospondin type 1 motif, 5 [*ADAMTS5*], cathepsin B [*CTSB*], cathepsin D [*CTSD*], matrix metalloproteinase-9, [*MMP9*], -13 [*MMP13*], Tissue Inhibitor of Metalloproteinase 1, [*TIMP1*], -2 [*TIMP2*] and -4 [*TIMP4*]), four reference genes (glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], TATA box binding protein [*TBP*], ribosomal protein L13a [*RPL13A*], and succinate dehydrogenase complex, subunit A [*SDHA*]) and genomic DNA were generated from the original cDNA samples (un-amplified). The reference genes used were selected from a panel of reference genes by applying a gene stability algorithm (238). The full details of these validated assays have been previously published (296). Spearmans rank correlation (ρ) between the fold up-regulation or down-regulation of each gene by the microarray analysis and the RT-qPCR results was performed using an online calculator (303).

RESULTS

Microarray Results

In the OA articular cartilage, a total of 2866 transcripts were differentially expressed when compared to the normal articular cartilage, with 1956 transcripts demonstrating increased expression and 910 transcripts demonstrating reduced expression ($P < 0.05$). However, correction of the data using the FDR determined that the expression of none of the transcripts were significantly increased or decreased. The top 50 annotated transcripts whose expression was changed by 2 fold or more are listed in Table 1, with their putative functions, and the un-corrected t-test *P* values.

A visual "heat" map illustrating the differential expression of the 50 annotated transcripts in each cartilage samples is presented in Figure 2, with the clustering patterns of genes and arrays. One diseased cartilage sample (HD1X) demonstrated a microarray gene expression profile which was more closely related (as determined by the clustering algorithm) to normal articular cartilage samples than to diseased articular cartilage samples. None of the genes clustered were determined to be on the previously defined pathways.

RT-qPCR Results

Five genes were differentially expressed in OA cartilage when compared to normal cartilage. The expression of four genes (*COL1A2, COL3A1, CSPG2* and *MMP13*) were increased (9.1, 1.2, 3.6 and 16.4 fold respectively), and the expression of one gene (*TIMP2*) was decreased (-1.0 fold). When corrected for multiple hypothesis testing using the FDR, none of the genes remained differentially expressed. The

expression profiles for the same genes as determined by microarray identified four genes to be differentially expressed in OA cartilage when compared to normal cartilage. Three different genes (*AGC, CTSD* and *TNC*) were identified to have significantly increased expression (0.6, 1.8, 1.0 fold respectively) when compared to normal cartilage, and one gene (*TIMP2*) was demonstrated to have a significantly reduced expression (-2.88 fold). Only one gene (*TIMP2*) demonstrated differential gene expression in OA articular cartilage when compared to normal cartilage, as determined by both RT-qPCR and microarray (when comparing data not corrected for multiple hypothesis testing). When corrected for multiple hypothesis testing using the FDR, none of the genes remained differentially expressed. Correlation between the mean fold expression changes of all the genes as determined by microarray and real-time PCR were strong (ρ=0.779, P = 0.00002).

DISCUSSION

The microarray analysis of canine OA hip articular cartilage revealed the change in expression of a number of genes previously not associated with the condition, but which may provide potential insight into the molecular mechanisms which govern the disease. For example, *ALS2CR2* is a gene involved in cell cycling which protects against IL-1β mediated apoptosis in human embryonic kidney cells (304). Thus, the reduction in its expression which we report is consistent with the increase in apoptosis recognised in OA cartilage (305). Similarly the *TERF2IP* gene, which helps prevent the dys-regulation of telomere length and structure (306), was shown to have reduced expression. Loss of telomere length is also recognised as a feature of OA

chondrocytes (307), and is thus consistent with the pattern of gene expression we identified in OA cartilage.

Two of the genes (*P2RX5* and *MSI*) demonstrating increased expression in OA articular cartilage, have been associated with the cellular differentiation pathways in studies of several different human tissues. MSI has been identified as a marker of stem cells, progenitor cells and differentiating cells in human fetal brain (308), and *P2RX5* receptor activity is responsible for early and late cellular differentiation in keratinocytes (309), with concomitant reduced expression of inhibitors of signal transduction, such as *MTUS* (310). The clustering algorithm indicated that expression profiles of both these genes were closely linked. Progenitor cells are recognised in mature articular cartilage (311), thus markers of progenitor cell activity would be anticipated to be active in the OA joint. Increased expression of other phenotypic markers of chondrocyte differentiation, such as *COL1A2*, which was identified by RT-qPCR in this and previous studies (296) of canine OA cartilage, suggest the change in OA chondrocyte phenotype to a more fibroblastic cell type, although this is disputed by some authors (312).

The down regulation of a number of ribosomal proteins, such as the 40S ribosomal protein *RPS2* and the 60S ribosomal protein *RPL36*, other RNA splicing associated proteins such as *PFRP8* and *SNRP*, and other regulators of transcription such as *ESF1* (313) and *FUBP1* (314) suggest a general reduction in gene transcription in OA cartilage. Likewise a reduction in the kinetchore associated protein *SMC1* (315) suggests a reduction in mitosis in the OA cartilage, which is again consistent with the increasing senescence of chondrocytes in OA (307).

Expression profiles of canine hip cartilage have previously demonstrated significant changes in expression of multiple matrix associated genes (*BGN, COL1A2, COL2A1, COL3A1, COL5A1, CSPG2, CTSB, CTSD, LUM, MMP13, TIMP1, TIMP2, TIMP4* and *TNC*) (296) in OA when compared to normal cartilage. In the limited sample set we evaluated in this study by microarray (5 OA versus 5 control cartilage samples), differential gene expression of five of these genes was identified by RT-qPCR, and three of these genes (with the addition of *AGR*) by the microarray profiles, which highlights the problems of using limited number of samples in diseases demonstrating a high level of heterogeneity. Similarly, although there was strong correlation between microarray and RT-qPCR quantification of the fold changes in expression of the 26 genes assayed, only one gene (*TIMP2*) demonstrated uncorrected statistical significance in both the microarray and RT-qPCR profiles.

The difference between the identity and number of genes with significant differences in expression can be ascribed to multiple potential causes, within this experiment. Firstly, visual examination of the microarray expression profile results, as highlighted by heat map, identified multiple samples having different patterns of expression. In particular, one disease sample demonstrated a pattern of gene expression which was in closer proximity to the normal articular cartilage samples than those from dogs with OA. Samples were only selected on the basis of breed, the site of articular cartilage retrieval (16,282), and sample RNA quality (276). Other factors known to affect cartilage gene expression such as patient age (316), joint radiographic score (317,318) or tissue histological score (284), were not controlled for. Hence, although all OA samples were taken from the same femoral head position (16) in clinical cases

requiring total hip replacement, different degrees of disease severity (and thus change in gene expression) may have been a major factor in mRNA expression.

Quantification of gene expression in a larger number of canine hip cartilage samples by RT-qPCR (296) demonstrated a degree of heterogeneity in measured articular cartilage gene expression, which highlights the problems of assaying end-stage OA tissue samples without additional phenotyping. Similar large variations in the range of fold changes of gene expression have also been observed in studies utilising clinical human OA cartilage, when small numbers of samples are evaluated (300). The cost of double amplification and oligonucleotide microarray analysis severely limited the number of replicates which could be performed in this study. The number of genes demonstrating differential expression, when uncorrected for multiple hypothesis testing, were within the expected range of previous microarray gene expression studies using the same number of experimental groups, albeit with paired samples of intact and damaged cartilage from individuals (300), rather than from different (control) dogs. For greater resolution of the nuances of gene expression, and avoiding the potential false positives, between normal and end/late stage OA cartilage requires the profiling of vastly greater number of samples (249). However, it should be noted that we have generated much more meaningful expression profile data using the same method and sample numbers with a different articular tissue (canine cranial cruciate ligament [CCL]) (319) in a different disease (CCL rupture).

The false discovery rate of microarray studies are determined by the proportion of truly differentially expressed genes, the distribution of the true differences, the measurement variability and sample size; but only sample size can be controlled by

the investigator. Consequently, small sample numbers dramatically limits the sensitivity of the standard statistical test (*t* test) to detect the truly differentially expressed genes, when not corrected. For example, the false discovery rate (number of genes which are reported as differentially expressed, but are not truly differentially expressed) can vary between 60 and 95%, when evaluating the *P* value alone (320). The sensitivity of the test (the proportion of genes which are truly differentially expressed and reported as such by the test) is also unacceptably low, and rises as the corresponding FDR rises, which cannot be fully resolved even by using larger numbers of arrays (320). Reducing the corresponding significance level does not improve the FDR greatly with small sample numbers unless the proportion of truly differentially expressed genes declines, and the genes considered are limited to those demonstrating greatest fold changes (as we have done in this study), although FDR assessment is still a much better assessment than *P* value alone. The FDR can be reduced. In this experiment, we set the FDR at 0.05, which makes the assumption that we expected a maximum of 5% of differentially expressed genes to be false positives. The limitations of the FDR in correcting this type of data are that some of the genes may be co-regulated- and thus their expressions being correlated, rather than truly random, and secondly that the null distribution of the statistic may not be truly normal, because of the limited number of samples evaluated (321).

When the data was looked at without FDR correction, the differential expression of certain genes would appear to be highly unusual. For the increased expression of example olfactory receptor *OFR7A17* in OA cartilage is highly unusual and suggests either that either it is a false positives (thus highlighting the benefit of the correction for multiple hypothesis testing of such large data sets), that this gene possesses a

unique function in OA cartilage hitherto not described, or that this gene is expressed by "accident", reflecting a degree of aberrant gene expression in end-stage OA cartilage. At present, the first explanation would appear to be the most likely, although the quantification of these genes by a more sensitive and accurate measure, such as RT-qPCR would have confirmed or refuted this assertion.

Current pathway analysis tools do not include the majority of genes which have been annotated from genome sequencing projects, because only a very limited number of pathways have been identified and annotated. Thus, identifying the functional significance of many apparently unrelated genes with limited function information can be both exceedingly time consuming and unrewarding. None of the genes identified as having changes in expression between normal and OA cartilage was subsequently identified to be on the same pathway in this study, which is probably a reflection of the wide variation in gene expression observed between samples, and the restricted power of pathway analysis tools to link genes about which there is such limited information.

CONCLUSIONS

A number of interesting and previously un-reported gene expression changes were noted in canine OA hip cartilage. However, the lack of strong statistical significance, and the limited number of technical replicates severely hindered the interpretation of the results, making meaningful conclusions difficult. Large numbers of better phenotyped samples may have produced more convincing results.

Table 1

Genes up- or down-regulated (2 fold or more) in canine hip OA cartilage when compared to normal hip cartilage, as determined by microarray analysis.

Gene ID	*Full Gene Name*	*Ref Seq Number*	*Function*	*Fold Change Disease vs Normal (+/- Standard Deviation)*	*Uncorrected P Value*
TRIOBP	TRIO and F-actin binding protein	XM_538384	Cytoskeletal organisation and motility	3.8 (+/-1.7)	0.002105
PTPRJ	Protein tyrosine phosphatase, receptor type, J	XM_540737	Transmembrane signalling	3.6 (+/-1.5)	0.006033
BLNK	B-cell linker	XM_543943	Intracellular signalling cascade	3.3 (+/-1.3)	0.004417
KRT21B	Type II keratin Kb21	XM_543659	Intermediate filament	3.3 (+/-1.0)	0.001731
MSI1	Musashi homolog 1	XM_849159	Post transcriptional gene regulation - RNA binding	3.2 (+/-1.3)	0.005023
P2RX5	Purinergic receptor P2X, ligand-gated ion channel, 5	XM_548343	Ligand gated membrane ion channel	3.0 (+/-1.2)	0.007693
OR7A17	Olfactory receptor, family 7, subfamily A, member 17	XM_848171	Signal transduction	2.8 (+/-1.0)	0.004417
RIMS3	Regulating synaptic membrane exocytosis 3	XM_844171	Exocytosis	2.6 (+/-0.8)	0.004477
TNRC4	Trinucleotide repeat containing 4	XM_857425	Transcription / splicing	2.1 (+/-0.7)	0.007869
NDOR1	NADPH dependent diflavin oxidoreductase 1	XM_548355	Electron transport	2.1 (+/-0.2)	0.005858
MCM9	Minichromosome maintenance complex component 9	XM_541221	Possibly Initiation of DNA replication	2.0 (+/-0.7)	0.002143
PPP3CA	Protein phosphatase 3, catalytic subunit, alpha isoform	XM_535672	Protein phosphorylation	-2.0 (+/-0.8)	0.001539
ANKRD2	Ankyrin repeat domain 2	XM_532483	Structural	-2.0 (+/-1.0)	0.005063
MTUS1	Mitochondrial tumour suppressor 1	XM_532829.	Control of cellular proliferation	-2.1 (+/-0.9)	0.001556
NF1	Neurofibromin 1	XM_537738	Regulation of signal transduction	-2.1 (+/-1.0)	0.002071
REV1	DNA Repair protein REV1	XM_538458	DNA Replication	-2.1 (+/-0.9)	0.00159
RPS2	Ribosomal protein S2	XM_844921	Ribosome (translation)	-2.2 (+/-0.9)	0.004822
BMS1	BMS1 homolog, ribosome assembly protein	XM_534956	Unknown	-2.3 (+/-1.2)	0.001886

Gene ID	*Full Gene Name*	*Ref Seq Number*	*Function*	*Fold Change Disease vs Normal (+/- Standard Deviation)*	*Uncorrected P Value*
PRNP	Prion protein	XM_542906	Copper and microtubule binding (membrane glycoprotein)	-2.3 (+/-1.4)	0.003377
AKT3	v-akt murine thymoma viral oncogene homolog 3	XM_547496	Signal transduction	-2.3 (+/-1.1)	0.007758
THRB	Thyroid hormone receptor, beta	XM_857597	Thyroid hormone receptor	-2.4 (+/-1.4)	0.007092
GCC2	GRIP and coiled-coil domain containing 2	XM_848883	Membrane protein	-2.4 (+/-1.8)	0.007722
FGL2	Fibrinogen-like 2	XM_533109	Extracellular matrix	-2.5 (+/-0.5)	0.007127
PRPF8	Pre-mRNA processing factor 8 homolog	XM_863390	mRNA processing	-2.5 (+/-1.6)	0.002763
SON	SON DNA-binding protein	XM_852093	DNA binding (regulation of transcription)	-2.8 (+/-1.5)	0.000669
SNRP	Small nuclear ribonucleoprotein	XM_536165	Regulation of transcription	-2.8 (+/-0.9)	0.000077
DST	Dystonin	XM_861733	Cytoskeletal (adhesion)	-2.8 (+/-1.1)	0.000852
FUS	Fusion (involved in t(12;16) in malignant liposarcoma)	XM_851770	Unknown	-2.8 (+/-1.3)	0.000247
TIMP2	Tissue inhibitor of matrix metalloproteinase-2	AF188489	Extracellular protease inhibition	-2.9 (+/-1.2)	0.002662
RPL36	Ribosomal protein L36	XM_538108	Translation (ribosomal protein)	-2.9 (+/-1.9)	0.003313
SERBP1	SERPINE1 mRNA binding protein 1	XM_536673	Regulation of mRNA stability	-3.0 (+/-2.0)	0.000748
RSBN1	Round spermatid basic protein 1	XM_540397	Unknown	-3.0 (+/-3.3)	0.007744
PRSS3	Protease, serine, 3	XM_547173	Extracellular protease	-3.1 (+/-2.6)	0.004796
TERF2IP	Telomeric repeat binding factor 2, interacting protein	XM_536776	Regulation of transcription	-3.2 (+/-3.0)	0.003013
PNP	Purine nucleoside phosphorylase (Inosine phosphorylase)	XM_532617	Transferase activity	-3.2 (+/-2.3)	0.002795
ESF1	ESF1, nucleolar pre-rRNA processing protein, homolog	XM_844119	Regulation of transcription	-3.2 (+/-2.8)	0.006126
ALDH5A1	Aldehyde dehydrogenase 5 family, member A1	XM_545368	Metabolic process (Mitochondrial dehydrogenase)	-3.3 (+/-3.3)	0.003797
ALS2CR2	Amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 2	XM_846559	Cell cycling	-3.3 (+/-3.3)	0.007478
LARP2	La ribonucleoprotein domain family, member 2	XM_533293	Unknown	-3.3 (+/-2.9)	0.005888

Gene ID	*Full Gene Name*	*Ref Seq Number*	*Function*	*Fold Change Disease vs Normal (+/- Standard Deviation)*	*Uncorrected P Value*
FOXO1A	Forkhead box O1A	XM_534487	Regulation of transcription	-3.5 (+/-2.5)	0.005663
NDUFAB1	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	XM_536932	Mitochondrial electron transport	-3.7 (+/-3.5)	0.001025
SLC38A	Solute carrier 38, member 2	XM_543722	Cell membrane transport	-3.9 (+/-3.0)	0.000318
PSMC2	Proteasome (prosome, macropain) 26S subunit, ATPase, 2	XM_533103	Proteolysis	-4.0 (+/-4.6)	0.001212
FUBP1	Far upstream element (FUSE) binding protein 1	XM_862753	Regulation of transcription	-4.0 (+/-3.3)	0.005836
RTN4	Reticulon 4	BK003959S2	Endoplasmic recticulum transport	-4.0 (+/-4.3)	0.00178
AKAP1	A kinase (PRKA) anchor protein 1	XM_861511	Regulation of cAMP signalling	-4.7 (+/-6.0)	0.004939
SMC1A	Structural maintenance of chromosomes 1A	XM_531868	DNA repair and chromosome organisation	-7.4 (+/-9.1)	0.004969
ATP11B	ATPase, Class VI, type 11A	XM_535816	Membrane ATPase (Ion transport)	-11.3 (+/-9.5)	0.00174
ZSWIM2	Zinc finger, SWIM-type containing 2	XM_535994	Metal / ion binding	-21.4 (+/-1.7)	0.000709
AHCTF1	AT hook containing transcription factor 1	XM_537228	Nucleopore assembly	-25.2 (+/-18.7)	0.001462

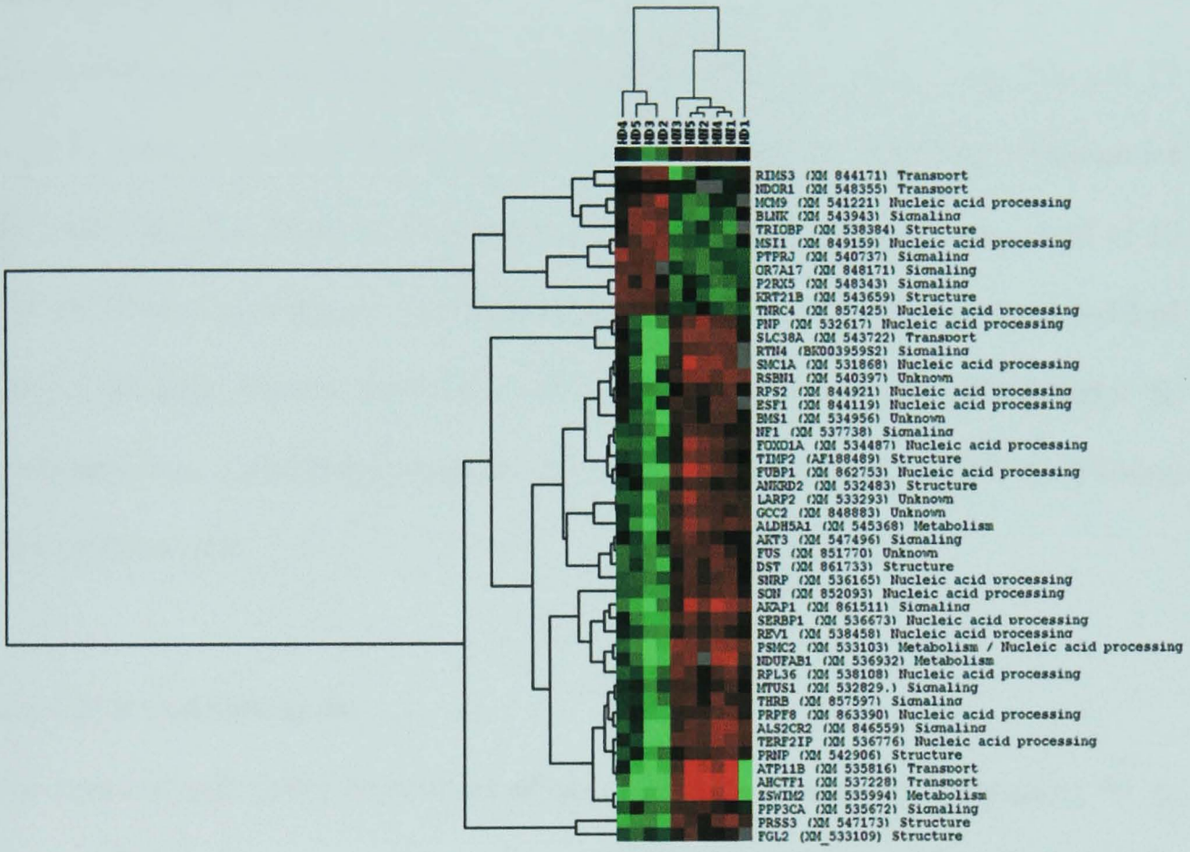
Table 2

Fold change in gene expression of selected genes in normal and osteoarthritic articular cartilage, as determined by microarray and RT-qPCR.

Gene	*Microarray Fold Change (+/-Standard Deviation)*	*t-test P Value*	*RT-qPCR Fold Change (+/- Standard Deviation)*	*t-test P Value*
ADAMTS5	-0.13 (+/-0.26)	0.527	-0.96 (+/-0.64)	0.457
AGC	0.97 (+/-0.32)	0.048	0.63 (+/-1.21)	0.289
BGN	0.20 (+/-0.78)	0.736	0.51 (+/-0.5)	0.158
COL1A2	7.01 (+/-0.74)	0.125	9.09 (+/-8.66)	0.049
COL2A1	0.56 (+/-0.28)	0.170	9.32 (+/-9.02)	0.050
COL3A1	0.37 (+/-0.86)	0.215	1.15 (+/-0.73)	0.022
COL5A1	0.08 (+/-0.47)	0.759	1.77 (+/-2.43)	0.160
COL9A3	0.29 (+/-0.65)	0.624	2.35 (+/-4.19)	0.250
CSPG2	0.96 (+/-0.29)	0.444	3.62 (+/-1.87)	0.004
CSTB	0.25 (+/-0.08)	0.079	0.57 (+/-0.61)	0.123
CSTD	0.59 (+/-0.14)	0.020	0.64 (+/-0.9)	0.188
DCN	0.17 (+/-0.47)	0.645	-0.19 (+/-0.77)	0.656
GAPDH	0.37 (+/-0.31)	0.297	0.06 (+/-0.5)	0.830
LUM	0.17 (+/-0.4)	0.724	1.11 (+/-1.22)	0.089
MMP13	4.56 (+/-0.32)	0.347	16.41 (+/-15.33)	0.044
MMP9	2.49 (+/-0.24)	0.050	56.06 (+/-121.62)	0.333
RPL13A	0.09 (+/-0.29)	0.775	0.17 (+/-0.08)	0.065
SDHA	-0.02 (+/-0.22)	0.875	-0.44 (+/-0.32)	0.102
TBP	0.27 (+/-0.65)	0.959	-0.06 (+/-0.53)	0.831
TIMP1	0.25 (+/-0.6)	0.724	0.63 (+/-0.57)	0.057
TIMP2	-2.88 (+/-0.31)	0.003	-1.05 (+/-0.42)	0.046
TIMP4	-0.05 (+/-0.17)	0.637	-1.67 (+/-0.25)	0.218
TNC	1.80 (+/-0.36)	0.045	0.77 (+/-0.86)	0.087
VIM	-0.01 (+/-0.54)	0.976	-0.03 (+/-0.33)	0.896

Figure 1

Heat map and hierarchical clustering plot of differentially expressed genes and arrays from canine hip OA cartilage (suffixed HD) when compared to normal canine hip cartilage (suffixed HN). (Colour coding; Green= Decreased expression, red= increased expression, black= no-change in expression, grey=missing value).



Supplementary Material

RNA Amplification Methods

First Strand Synthesis

Ten microlitres of RNA (total amount 200 ng) was incubated with 1 µl of 100 µM T7 oligo dT promoter primer (Invitrogen) for 10 minutes at 70°C, and then cooled on ice for 2 minutes. A mastermix, containing 2 µl of 0.1M DDT (Invitrogen), 1 µl of 10 mM dNTP mix (Invitrogen), 1 µl Ribonuclease inhibitor (RNAsin, Promega) and 1 µl (200U) Molony Murine Leukemia Virus Reverse Transcriptase (Superscript II, Invitrogen) was added to each sample. The mixture was incubated at 42°C for 2 hours, then cooled on ice.

Second Strand Synthesis

One hundred and thirty microlitres of second strand mastermix, containing 91 µl RNase / DNase free water, 30 µl second strand buffer, 3 µl of 10 mM dNTP mix (Invitrogen), 1 µl (10U) DNA ligase (Invitrogen), 4 µl (40U) *Escherichia coli (E.coli)* DNA polymerase 1 (Invitrogen) and 1 µl (2U) Ribonuclease H (RNase H, Invitrogen), was added to each sample. The mixture was incubated at 16°C for 2 minutes. 2 µl (10U) T4 DNA polymerase (Invitrogen) was added and the mixture incubated for 15 minutes at 16°C.

cDNA Precipitation

Half a microlitre of glycogen (10 mg/ml, Invitrogen), 75 µl 5M ammonium acetate and 375 µl 100% ethanol (Sigma) were added to the volume and mixed. The solution

was centrifuged at 4°C for five minutes at 10000 g. The supernatant was removed and the precipitate dried at 60°C, in a vacuum centrifuge (1000 g) until a dry pellet remained (10-50 minutes). The pellet was re-suspended in 8 µL RNase/ DNase free water.

RNA amplification

Amplified RNA (aRNA) was produced using the Ambion T7 MEGAscript high yield transcription kit accordance with the manufacturer's instructions (Ambion, UK). Briefly, 2 µl of 75 mM ATP, GTP, CTP and UTP were each added to the cDNA samples, with 10x reaction buffer and 2 ul RNA polymerase mix. The mixture was incubated for 16 hours at 37°C, after which 1 µl (2U) DNase was added to remove template complementary DNA.

Amplified RNA (aRNA) Clean Up

The amplified RNA (aRNA) was cleaned using a standard procedure. Briefly, the aRNA volume was increased to 100 µl by adding 79 µl RNase - DNase free water. 350 µl of RTL buffer was added, and the solution mixed. 250 µl ethanol (96-100%) was added and mixed. The samples was applied to an RNeasy mini column and centrifuged at >8000 x g for 15 seconds. The flow-through was discarded. The column was transferred to a new collecting tube and 500 µl RPE Buffer added to the column. The centrifugation was repeated, supernatant discarded and another 500 µl RPE added to the column. The column was centrifuged at >8000 x g for 2 minutes. The column was placed in a clean RNase - DNase free Epindorph tube, and the centrifugation repeated for at >8000 x g for 2 minutes. aRNA was eluted by adding 50 µl of RNase - DNase free water directly onto the RNeasy silica-gel membrane and

allowing the tube to stand for 1 minute at room temperature, followed by centrifugation at >8000 x g for 1 minute. The elution was repeated with a further 50 µl RNase/DNase free water, 1 minute standing time and centrifugation. The aRNA concentration was determined using a spectrophotometer (NanoDrop).

Second Round aRNA Amplification

As the total mRNA required for optimal microarray hybridisation (2 µg) was not achieved using a single amplification procedure, a second round of amplification was performed on all aRNA samples using the MessageAmp (Ambion) aRNA Amplification kit, in accordance with the manufactures instructions, as follows;

First Strand Synthesis

Briefly, the aRNA samples were reduced to 10 µl in volume using a vacuum centrifuge as previously described (so that each sample contained up to 2 µg aRNA). Two microlitres of second round primers were added and the mixture heated to 70°C for 10 minutes. The samples were briefly centrifuges and placed on ice. A mastermix of 2 µL 10x First Strand Buffer, 1 µL Ribonuclease inhibitor, 4 µl dNTP mix and 1 µl reverse transcriptase was added to each sample, and the samples incubated for 2 hours at 42°C. 1 µl of RNase H was added to each sample, followed by incubation at 37°C for 20 minutes.

Second Strand Synthesis

Five microlitres of T7 oligo (dT) primer were added to the first strand reaction and incubated at 70°C for 10 minutes. Second strand cDNA synthesis reagents were added

as follows; 10 µl 10x second strand buffer, 4 µl dNTP mix, 2 µl DNA polymerase and 58 µl RNase/DNase free water. The mixture was incubated at 16°C for 2 hours.

cDNA Purification

Double stranded cDNA was purified using by adding 250 µl cDNA binding buffer to each sample, and 50 µl to each filter cartridge. Each sample was centrifuged at 10000 x g for 1 minute, the flow through discarded and the centrifuge procedure repeated. 500 µl cDNA wash buffer was added to each filter cartridge, and each sample centrifuged again for 1 minute at 10000 x g. cDNA was eluted by adding 10 µl RNase – DNase free water pre-heated to 50°C to each filter cartridge, which was then left to stand for 2 minutes at room temperature before centrifuging at 10000 x g for 1 minute. The elution procedure was repeated with a further 10 µl of RNase – DNase free water.

Second Round RNA Amplification

Sixteen microlitres of each double stranded DNA sample was mixed with 4 µl each of 75 mM T7 ATP, GTP, CTP and UPT solutions, 4 µl 10x reaction buffer and 4 µl T7 RNA polymerase enzyme mix. Samples were incubated at 37°C for 14 hours, in a hybridization oven, followed by the addition of 2 µl DNase I to each reaction and incubation at 37°C for a further 30 minutes.

Second Round aRNA Clean Up

60µl of elution solution was added to each aRNA sample and 350µl of aRNA binding buffer added to each aRNA sample. 250 µl 100% ethanol was added to each aRNA sample and mixed by vortexing. Each sample was placed in an aRNA filter cartridge,

and centrifuged at 10000 x g for 1 minute. 650 µl of aRNA wash buffer was applied to each filter cartridge, which were then centrifuged at 10000 x g for 1 minute and the flow through discarded. Each filter cartridge was then spun at 10000 x g for 1 minute. aRNA was eluted by 50 µl RNase – DNase free water pre-heated to 50°C was added to each filter cartridge, which was then left to stand for 2 minutes at room temperature before centrifuging at 10000 x g for 1 minute. The elution procedure was repeated with a further 50 µl of RNase – DNase free water. The aRNA concentration was determined using a spectrophotometer (Nanodrop).

Chapter 4

Cartilage gene expression correlates with radiographic severity of canine elbow osteoarthritis

Dylan N. Clements[1,2],

Noel Fitzpatrick[3],

Stuart D. Carter[1],

Philip J.R. Day[2].

[1]Musculoskeletal Research Group, c/o Department of Veterinary Pathology, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX.

[2]Centre for Integrated Genomic Medical Research, The Stopford Building, University of Manchester, Oxford Rd, Manchester, M13 9PT.

[3]Fitzpatrick Referrals, Farnham, Surrey, GU2 9AD, UK

Paper published in *The Veterinary Journal* (2007) E-publication ahead of print

ABSTRACT

Introduction

The relationship between clinical measures of osteoarthritis (OA) severity, such as radiography, and changes in gene expression in affected joint tissues are not well characterised.

Materials and Methods

Messenger RNA (mRNA) was extracted from both articular cartilage and bone from radiographically imaged elbow joints of dogs treated surgically for fragmented coronoid disease. The expression of candidate genes involved in articular cartilage production and degradation and bone remodelling were quantified in normal and diseased cartilage and bone using qPCR. Absolute transcript numbers were calculated using template oligonucleotide calibrators. The subsequent patterns of expression were then correlated with the radiographic severity of OA.

Results

The expression of collagen genes (*COL1A2, COL3A1*), matrix metalloproteinase genes (*MMP2, MMP9, MMP13*) and matrix structural genes (*LUM*) were all significantly increased in OA cartilage. The expression of *TIMP2* and *CTSD* genes were decreased in OA cartilage. The expression of *COL1A2, MMP2, MMP9, MMP13* and *TIMP1* were all increased in OA bone, and the expression of *TIMP2* was decreased in OA bone. Significant correlations between gene expression level and radiographic OA grade were identified for *COL1A2, COL3A1* and *LUM* expression (positive) and *TIMP2* and *CTSD* expression (negative). Similar patterns of gene expression were observed in OA cartilage and OA bone, although the relative magnitude of gene expression and their changes in OA differed between the two tissues.

Conclusions

The positive correlation of cartilage gene expression with the radiographic severity of OA demonstrates that molecular measures of disease activity and associated tissue response can be directly related to clinical assessment of OA.

INTRODUCTION

Osteoarthritis (OA) is characterised by articular cartilage destruction, although other tissues such as bone (4), synovium (1), fat (2) and ligament (102) are also changed in OA joint, and may be involved in the pathogenesis of the disease. Up to 17.8% of dogs in purebred populations demonstrate radiographic evidence of elbow OA (34) and this is often secondary to other, primary processes such as fragmented medial coronoid process (FCP).

Fragmented coronoid process is a disease of elbow joints which results in changes in the bony architecture of the medial coronoid process of the elbow (70), invariably leading to the development of secondary OA, and is the primary cause of elbow lameness and elbow OA in dogs (322). At present, the pathogenesis of medial coronoid fragmentation is unknown, although the observation of microcrack damage and increased porosity in the subchondral and trabecular bone of the FCP suggests that the osteochondral fragment develops from chronic fatigue damage to the bone (70). These microscopic changes in bone structure are also the hallmark of OA bone (323,324), and so although FCP can occur in the absence of macroscopic evidence of OA (71), the bone pathology which characterises the disease may represent OA *per se*. The importance of bone in the pathogenesis of OA has been recognised, with some authors suggesting that OA may be primarily a disease of the skeleton (4). Bone changes are identified in OA, such as increases in subchondral bone density and activity (5), which may precede subsequent osteoarthritic changes in articular cartilage (325). Gene expression measurements in bone have been used to characterise pathways involved in the remodelling of bone in relation to remodelling associated with both OA (192,326) and loading (256).

The removal of the fragmented bone and associated cartilage in dogs with FCP is a routine surgical treatment for the condition (70) and thus provides an ethically acceptable method of obtaining bone and cartilage specimens for the study of gene expression in naturally occurring canine OA. A major advantage of using FCP fragments for molecular investigation is that the anatomical site of tissue collection is identical between patients, which is important as gene expression in both normal and OA joints varies widely between sites in the same joint (327). Clinical assessment of canine OA can be achieved by a number of means, but radiography is usually considered the most useful and quantifiable. In particular, there is a well-established radiographic scoring system for canine elbow OA (International Elbow Working Group Scoring Scheme) (76), which uses osteophyte size to "score" the grade of OA and thus allows an ordinal grading of the radiographic severity of elbow OA.

To date, few studies have examined whether there is any association of molecular events with measures of disease severity, such as histological scores (284,300), or clinical measures, such as radiographic scores (317,318). We have previously identified changes in expression of a number of matrix associated genes in end-stage canine hip OA cartilage (296), such as type I collagen alpha 2 chain [*COL1A2*], type II collagen alpha 1 chain [*COL2A1*], type III collagen alpha 1 chain [*COL3A1*], cathepsin D [*CTSD*], lumican [*LUM*], matrix metalloproteinase -13 [*MMP13*], tissue inhibitor of matrix metalloproteinase -1 [*TIMP1*], -2 [*TIMP2*] and tenascin C [*TNC*]. In the present study, we aimed to quantify the expression of these genes in canine elbow OA, and two other metalloproteinases known to have increased gene expression in human OA cartilage (matrix metalloproteinase -2 [*MMP2*] and -9 [*MMP9*]) (255). We selected two genes involved in primary bone structure (328)

(*COL1A2* and Type X collagen alpha 1 chain [*COL10A1*]), and seven genes which are differentially expressed in experimental models of bone remodelling (Annexin 2A [*ANXA2*], *CTSD,* fibronectin 1 [*FN1*], *MMP2, MMP9, MMP13, TIMP1, TNC*) (256,329) and quantified their expression in associated normal and bone fragments. Finally, we analysed whether there was any correlation between the levels of cartilage gene expression with OA severity, as assessed by radiographic score.

MATERIALS AND METHODS

Sample Collection and Radiographic Evaluation

Dogs presenting for investigation and treatment of elbow pain to a private referral orthopaedic clinic[3] between January 2005 and December 2005 were evaluated by a single clinician (NF). Standard elbow radiographs (flexed and neutral mediolateral) were taken and elbow osteophyte size was scored in an ordinal manner (IEWG grade 0= no osteophytes, grade 1 =osteophytes <2 mm in size, grade 2 = osteophytes 2 – 5 mm in size and grade 3 = osteophytes > 5 mm in size) using a standardised scoring scheme (76) by a single observer (NF). Inclusion criteria for the study were dogs where the presence of FCP was confirmed by both arthroscopic examination and direct surgical visualisation of the medial coronoid process (medial arthrotomy). Exclusion criteria were the absence of concurrent conditions of the elbow as determined by radiography, arthroscopy and arthrotomy. Twenty dogs met the inclusion criteria over the study period.

The fragmented medial coronoid process was surgically removed using an osteotome as part of a standard surgical procedure (70). Following the collection of the tissue,

the samples were washed in sterile Hartmann's solution (Isolec, Ivex Pharmaceuticals) and articular cartilage removed by cutting off the articular surface of the bony fragment with a No.10 scalpel blade. The trabecular bone from the bony cut surface created by the osteotome was gouged out at the centre of the cut surface in fine pieces using a pair of bone cutters and Lempert rongeurs. The trabecular bone of the medial coronoid process bone was separated and immediately stored in RNAlater (Qiagen), according to the manufacturer's instructions, until extraction. The medial coronoid process was collected from 12 dogs euthanatized for reasons other than, and with no gross evidence of, elbow joint disease. Samples were collected and stored in the same manner as diseased samples, except the joint was grossly dissected to confirm the absence of osteophytes or macroscopic evidence of articular cartilage damage on any parts of the articular surfaces. As osteophytes are evident macroscopically before radiographic changes are present in other canine joints (330), assigning a radiographic score of 0 was deemed appropriate where osteophytes were absent.

RNA Extraction and cDNA Synthesis

Total RNA was extracted from bone and cartilage samples using a liquid nitrogen cooled dismembrator, a phenol/guanidine hydrochloride reagent (Trizol, Invitrogen Ltd) with a chloroform extraction and ethanol precipitation, as previously described (276). An on-column DNA digestion step was included (RNase-Free DNase Set, Qiagen Ltd). Total RNA samples were stored at -80 ^{o}C until use.

Reverse transcription was performed using Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's instructions (331). Initially 200 μg (10

µl) total RNA was pre-incubated with 0.5 µg (1 µl) oligo-dT_{12-18} (Invitrogen) and 10 mM (1 µl) dNTP mix (Invitrogen) at 65°C for 5 minutes. 4 µl of 5x first strand buffer (containing 250 mM Tris-HCI (pH 8.3), 375 mM KC1, 15 mM $MgCl_2$), 2 µl of 0.1 M DTT and 40 units (1 µl) of RNAsin (Promega) were added to each sample and the samples incubated for 2 minutes at 42°C, followed by the addition of 200 units (1 µl) of Superscript II reverse transcriptase (Invitrogen) and incubation for 50 minutes. Reverse transcriptase activity was terminated by incubation at 70°C for 15 minutes.

Quantative Reverse Transcriptase Polmerase Chain Reaction Assay Design

The primer and probe sequences, efficiency values, dynamic ranges have been previously published (296) for the following canine mRNA expression assays; type I collagen alpha 2 chain [*COL1A2*], type II collagen alpha 1 chain [*COL2A1*], type III collagen alpha 1 chain [*COL3A1*], cathepsin D [*CTSD*], lumican [*LUM*], matrix metalloproteinase -13 [*MMP13*], tissue inhibitor of matrix metalloproteinase -1 [*TIMP1*], -2 [*TIMP2*], tenascin C [*TNC*]. Expression assays were also designed, as previously described (296), for matrix metalloproteinase -2 [*MMP2*], -9 [*MMP9*], annexin 2 A [*ANXA2*], type X collagen alpha 1 chain [*COL10A1*], fibronectin 1 [*FN1*], and the reference genes HIRA interacting protein 5 isoform 2 [*HIRP5*] and mitochondrial ribosomal protein S7 [*MRPS7*]. The primer and probe sequences, assay efficiencies and dynamic ranges are presented in Table 1. Primers were synthesized by MWG Biotech, and probes were synthesized by Roche Diagnostics using locked nucleic acid analogues with a 5'-end reporter dye fluorescein (FAM (6-carboxy fluorescein)) and a 3'-end dark quencher dye. Template oligonucleotides (332) for the amplicon generated by each assay were synthesized by Eurogentec (Southampton, UK). The two reference genes (*MRPS7* and *HIRP5*) were previously identified from

micro-array data, and confirmed as having stable expression in normal and OA bone and cartilage, using a statistical algorithm (238).

The quantative reverse transcriptase polmerase chain reaction (RT-qPCR) assays were all performed in triplicate using a TaqMan™ ABI PRISM 7900 SDS (Applied Biosystems) in 384-well plate format. Each assay well had a 10 µl reaction volume consisting of 5 µl 2X PCR master mix with Uracil N-Glycosylase (Universal PCR Mastermix, Applied Biosystems), 0.1 µl each of 20 µM forward and reverse primers, 0.1 µl of 10 µM probe (Exiqon, Roche Diagnostics) and 4.7 µl of sample cDNA (templates) or water (negative controls). The amplification was performed according to a standard protocol with 10 minutes at 50°C then 40 cycles of 95°C for 1 min and 60°C for 15 sec, as recommended by the manufacturer (Applied Biosystems). Real-time data were analysed using the Sequence Detection Systems software, version 2.2.1 (Applied Biosystems). The detection threshold was set manually for all assays at 0.05. Threshold cycle data was exported into Microsoft Excel 2003 (Microsoft) and evaluated using statistical software package (Minitab v14.1, Minitab Ltd). For each plate, a calibrator of known transcript number for each assay was used. The calibrator template number was determined as equal to:

Volume of template added x Template concentration (Molar) x N (Where $N = [6.022 \times 10^{23}]$*)*

Absolute quantification of gene transcript number in an unknown sample was determined as equal to:

$$\textit{Calibrator template number x E value of assay}^{(C_T \textit{ value of calibrator assay} - \textit{Mean } C_T \textit{ value of unknown assay})}$$

The template numbers were then averaged for each gene in each sample. To correct for the bias which might be induced by one reference gene being more abundantly expressed if an arithmetic mean were to have been used, an average of the absolute

quantities of the two reference genes transformed to a geometric mean (238), was applied using the following formula;

Geometric mean of the two reference genes

= Square root of [(Transcript number of MRPS7) X (Transcript number of HIRP5)]

The transcript numbers for each gene of interest, in each sample, were divided by the geometric mean of the two reference genes, and the normalised transcript number (referred to as molecules per reference gene) used in all statistical calculations and equations. Weights and ages, and normalised transcript numbers were not normally distributed, and thus were all compared by Mann-Whitney *U* tests. Comparisons of normalised transcript numbers were checked for multiple hypothesis testing, using a Benjamini and Hochberg false discovery rate test (281). Spearman's rank correlation coefficient (ρ) and their associated *P* value were calculated for the molecules per reference gene compared to the OA score using an online statistical calculator (303). Statistical significance was set at $P < 0.05$.

RESULTS

Population Comparisons

The patient groups had been age and weight matched; the median age of the control group (1.5 years, range 6 months to 4 years) was not significantly different ($P = 0.15$) to the OA group (1 year, range 7 months to 7 years). The median weight of the control group (30.8kg, range 12 kg- 43 kg) was not significantly different ($P = 0.43$) to the control group (30.0 kg, range 17.5 kg – 51 kg). The diseased dogs were categorised as having radiographic OA scores of 1 (n = 5), 2 (n = 13) and 3 (n = 2). When stratified into separate groups on the basis of radiographic OA score, the age and weight

variables were not significantly different between these 2 groups ($P = 0.110$ and $P = 0.155$ respectively).

Gene expression profiles of articular cartilage

Significant increases in the level of mRNA expression of *COL1A2, COL3A1, LUM, MMP2, MMP9* and *MMP13* were identified in the cartilage from cases with FCP, with concurrent significant decreases in the level of expression of *CTSD* and *TIMP2* (Table 2 and Figure 1). The largest fold changes in expression were identified for *MMP13* (4.2 fold increased expression), *TIMP2* (3.8 fold decrease in expression), *COL3A1* (2.8 fold increased expression) and *LUM* (2.7 fold increased expression). Significant positive correlation of gene expression with radiographic OA scores was identified for *COL1A2, COL3A1, MMP2, MMP9, MMP13* and *LUM* and significant negative correlation of gene expression with radiographic OA scores was identified for *CTSD* and *TIMP2* (Table 2 and Figure 1). The strongest correlations between radiographic OA score and gene expression change were identified for *COL3A1* ($\rho = 0.66$), *TIMP2* ($\rho = -0.64$) *LUM* ($\rho = 0.63$), and *COL1A2* ($\rho = 0.60$). The relationship between transcript number and OA grade for genes with a $\rho > 0.5$ and are illustrated in Figure 2.

Gene expression profiles of trabecular bone

Complete expression profiles could be generated in all of the control samples and 11 of the diseased (OA) samples (the quantity of RNA extracted from the remaining samples was not sufficient to provide complete data sets, and thus were ignored). Significant increases in the expression of *COL1A2, MMP2, MMP9* and *MMP13* were identified, with concurrent decreased expression of *TIMP2* (Table 2 and Figure 1).

The largest fold changes in expression were identified for *COL1A2* (3.7 fold increase in expression), and *MMP9* (2.3 fold increase in expression). The majority of diseased bone samples for which an expression profile could be generated were from cases with a grade 2 radiographic OA score so meaningful correlation analyses of bone gene expression and radiographic score could not be made.

DISCUSSION

The expression profiles of genes responsible for structural components of articular cartilage (*COL1A2, COL3A1,* and *LUM*) and the inhibition of metalloproteinase activity (*TIMP2*) were well correlated to the gross radiographic changes in OA joints in this study. Intuitively, one might expect the expression of structural matrix molecules and mediators of proteolytic degradation to correlate with radiographic severity of OA. However, the radiographic features of OA lag behind both the histological (330) and the molecular changes (changes in gene expression) (327) associated with the disease. The histological changes in OA cartilage are also delayed behind the changes in gene expression in experimental canine knee OA (327). Perhaps, most significantly, radiographic measures do not correlate with the most important clinical feature, lameness, in canine patients with elbow (333) or stifle (334) OA. Hence, although this study identified measures of gene expression which have the potential to function as markers of severity in canine elbow OA, the relationship with more meaningful measures of disease, would have been better investigated.

Data from human knee OA cartilage suggest that no changes in gene expression of structural matrix-associated genes (*COL1A2, COL2A1, COL3A1, LUM* and *TNC)* are identified in OA cartilage samples from patients with histological evidence of "early" degenerative changes when compared to normal cartilage samples, although marked changed are evident in OA cartilage from patients with "late" OA changes (249,312). The increased expression of *COL1* and *COL2* are early features of experimental canine knee OA (287), appearing at 6 weeks post injury, although the histological changes in the cartilage at this time-point are already greater than that considered "early degenerative" in other studies (249). This data supports the view that radiographic measures of canine elbow OA are somewhat unsubtle indicators of the degenerative process present in a joint, and thus joints with even minimal radiographic evidence of OA are likely to have marked gene expression changes in their articular cartilage as we detected. Although we identified significant correlations, the inherent radiographic bias (with 65% of samples being grade 2) suggests that these findings should be interpreted with caution until studies of larger numbers of samples, with the evaluation of additional variables, such as histological scores (284,300) have been performed.

Links between the radiographic phenotype and molecular events occurring in OA have been previously suggested. For example, the proteoglycan content of cultured chondrocytes from OA joints can be correlated to human knee radiographic OA score (318), the expression of aggrecan (*AGR)* has been correlated to clinical OA grade in the human knee (284) and measures of canine hip laxity (Norberg angle) can be correlated to IL1β activity in canine hip OA (335). However, other studies have reported that *AGR* expression in OA cartilage was not related to histological or

radiographic measures of OA severity (317,318). We did not evaluate the expression of *AGR* in this study as we had previously identified that *AGR* expression does not change in canine OA cartilage, in a different joint (the hip) (296) and other studies have documented the expression of *AGR* in experimental canine OA as being highly variable (287).

The general pattern of increased structural matrix (*COL1A2, COL3A1* and *LUM*) gene expression showed increases in OA cartilage, which were consistent with previous reports of expression profiles from other end-stage canine OA joints (hip) (296), and human knee OA cartilage (292,336). The magnitude of the increases in *COL1A2* and *COL3A1* gene expression and the absence of changes in *COL2A1* and *TNC* gene expression in canine elbow OA were in contrast to the changes previously reported in naturally occurring end-stage canine hip OA (296). Whether these disparities represent joint specific differences in gene expression or a lack of truly end-stage OA in this study is unclear.

MMP13 may be considered to be the primary collagenase responsible for the initial cleavage of type II collagen in OA (255), therefore it is unsurprising that we identified the greatest fold increases in expression of *MMP13* when compared to the fold increases of *MMP2* and *MMP9* in OA articular cartilage. *MMP13* gene expression was also increased in diseased bone, although to a lesser extent. Our findings in elbow OA were consistent with previous studies of naturally-occurring canine hip OA (296), experimental canine OA (287) and naturally occurring human OA (255). Increases in *MMP13* expression tend to be a feature of late experimental canine OA (287), and late stage human knee OA (337). Gene expression for gelatinase (*MMP2* and *MMP9*)

production was also increased in elbow OA cartilage, which is in agreement with studies of their expression in human OA studies (248,255,338) and animal models of OA (339). Interestingly, the absolute transcript number of *MMP9* and *MMP13* were similar, and always greater than *MMP2*, in both OA bone and cartilage.

The expression of *TIMP2* is repeatedly reported as decreased in canine OA, both in this study and those published previously (296,327,340) and in other joint diseases such as osteochondrosis (341). This pattern of expression of the *TIMP2* gene in canine OA is consistently different to that reported in human OA, where no change in *TIMP2* gene expression is noted (255), thus *TIMP2* expression appears to demonstrate a strong species, but not disease, specificity in its expression. Clearly, *TIMP2* warrants further investigation in canine OA, where its reduced expression could contribute to the development and progression of disease. Increased expression of *TIMP1* was identified in OA bone and suggests a response to the mechanical and molecular changes present in OA. We did not identify a corresponding significant increase in the expression of *TIMP1* in canine OA cartilage, in contrast to canine end-stage hip OA (296) and canine experimental OA (342), which may reflect the less end-stage nature of disease in this study, although decreased *TIMP1* expression has been reported in naturally occurring human OA (255).

COL1A2 expression in OA bone was in keeping with its function as the primary structural component of bone, and thus its increased expression was expected given the sclerotic changes which are reported in the fragmented bone of the medial coronoid process (69) in cases of FCP. The increased activity of MMPs in OA bone is also in agreement with previous study of *MMP2* activity in human OA bone (343).

The selection of other genes (*ANXA2, CTSD, COL10A1, FN1* and *TNC*) which have been reported to be involved in experimental bone remodelling pathways (256) did not identify similar pathways being activated dog, with the exception of the increased expression of matrix metalloproteinase genes (*MMP2, MMP9, MMP13*) and reduced expression of one of their inhibitors (*TIMP2).* The change in bone density observed in the MCP in dogs with elbow OA may be governed by pathways which are not activated or identified in animal models of bone remodelling (256,329), or it may have been that the pertinent changes in bone gene expression were too small to be identified.

Complete expression profiles of bone were not always possible because the extraction of mRNA from bone is problematic, due to the structural nature of the tissue (lower cellularity of trabecular bone when compared to articular cartilage), or because of the small quantities (often <10 mg) of tissue available. Although RNAlater is specifically not recommended for use on hard tissues, previous work has shown that usable high quality mRNA can be recovered from suitable thin bone specimens (256), and the tissues were fragmented prior to storage to maximise the penetration of the storage solution. As samples were obtained from a surgical practice distant from the laboratory, alternative storage techniques, such a liquid nitrogen freezing, were not available.

CONCLUSIONS

Gene expression changes in canine elbow OA cartilage correlated with radiographic assessment of elbow OA. The most significant correlations were identified positively with structural genes (*COL1A2, COL3A1* and *LUM*) and negatively with *TIMP2.* Cartilage and bone samples demonstrated similar changes in protease and protease inhibitor expression. The accurate quantification of gene transcription in clinical tissues may allow the identification of biomarkers which accurately reflect measures of gross disease status, such as the radiographic score.

Table 1

Primer sequences, probes sequences, the dynamic range, correlation and PCR efficiency values for each assay.

Gene	*Forward Primer Sequence*	*Reverse Primer Sequence*	*LNA Probe Sequence*	*Lower detection limit (CT value)*	*Upper detection limit (CT Value)*	r^2 *Value*	*PCR Efficiency (%)*
ANXA2	agaaagtatggcaagtccctgt	ctttctggtagtcgcccttg	catccagc	20.9	32.0	0.992	90.2
COL10A1	acctggacaacagggaccta	ccccttttctcctggaaatc	agccccag	25.9	37.4	0.992	93.1
FN1	gaccagaagaggcacaaggt	gctggtttaggccttggtc	gggaggag	19.1	32.8	0.988	100.1
HIRP5	aattcagaacatgctgcaatttta	tgattcatcatccataacctgttc	aggtggag	21.7	32.0	0.980	94.9
MMP2	acctgcaaggcagtggtc	tccaaatttcacgcttttca	agctggag	15.3	32.2	0.992	95.1
MMP9	cacgcatgacatcttccagt	cgagaattcacacgccagta	cttctgcc	14.5	30.5	0.992	101.8
MRPS7	agtgcagggagaagaagcac	cagcagctcgtgtgacaact	ggatgctg	22.4	32.3	1.000	98.0

Table 2

The gene expression values in normal and OA articular cartilage and trabecular bone.

Gene	*Transcript Number (Control Cartilage)*	*Transcript Number (OA Cartilage)*	*Mann-Whitney U Test P Value*	*Correlation coefficient (Spearmans ρ value)*	*p value for correlation*	*Transcript number (Control Bone)*	*Transcript Number (OA Bone)*	*Mann-Whitney U Test p Value*
ANXA2						2.5 (+/-0.2)	1.9 (+/-0.3)	0.0455
COL1A2	38.6 (+/-18.8)	126.5 (+/-32.1)	0.0006	0.6008	0.0006	61.1 (+/-15.9)	288.8 (+/-46.7)	0.0002
COL2A1	11.8 (+/-3.5)	33.4 (+/-24.5)	0.3811	0.2285	0.1936			
COL3A1	2.6 (+/-0.5)	9.7 (+/-2.2)	0.0002	0.6688	0.0002			
COL10A1						0.04 (+/-0.03)	0.01 (+/-0.00)	0.7344
CTSD	0.9 (+/-0.1)	0.6 (+/-0.2)	0.0108	-0.5067	0.0214	1.5 (+/-0.2)	1.0 (+/-0.1)	0.0694
FN1						22.9 (+/-4.4)	17.2 (+/-4.4)	0.1316
LUM	11.4 (+/-2.9)	42.3 (+/-10.4)	0.0007	0.6333	0.0002			
MMP2	0.3 (+/-0.1)	0.6 (+/-0.2)	0.0167	0.483	0.0042	0.3 (+/-0)	0.9 (+/-0.2)	0.0051
MMP9	0.7 (+/-0.3)	1.8 (+/-0.3)	0.0012	0.3942	0.0164	1.6 (+/-0.4)	5.1 (+/-0.8)	0.0015
MMP13	0.5 (+/-0.1)	2.6 (+/-0.9)	0.001	0.4955	0.0034	1.2 (+/-0.3)	3.2 (+/-0.4)	0.0015
TIMP1	1.4 (+/-0.2)	3 (+/-1.2)	0.0645	0.2666	0.2666	1.2 (+/-0.2)	2.6 (+/-0.4)	0.0247
TIMP2	4.7 (+/-0.7)	1.2 (+/-0.4)	0.0001	-0.6421	0.0026	1.9 (+/-0.2)	0.9 (+/-0.1)	0.001
TNC	2.8 (+/-0.5)	2.8 (+/-0.9)	0.3403	-0.1137	0.7948	2 (+/-0.5)	3.2 (+/-0.5)	0.0605

Figure 1

Mean fold change in gene expression (and standard error) of each gene evaluated in OA cartilage or bone when compared to the control tissue.

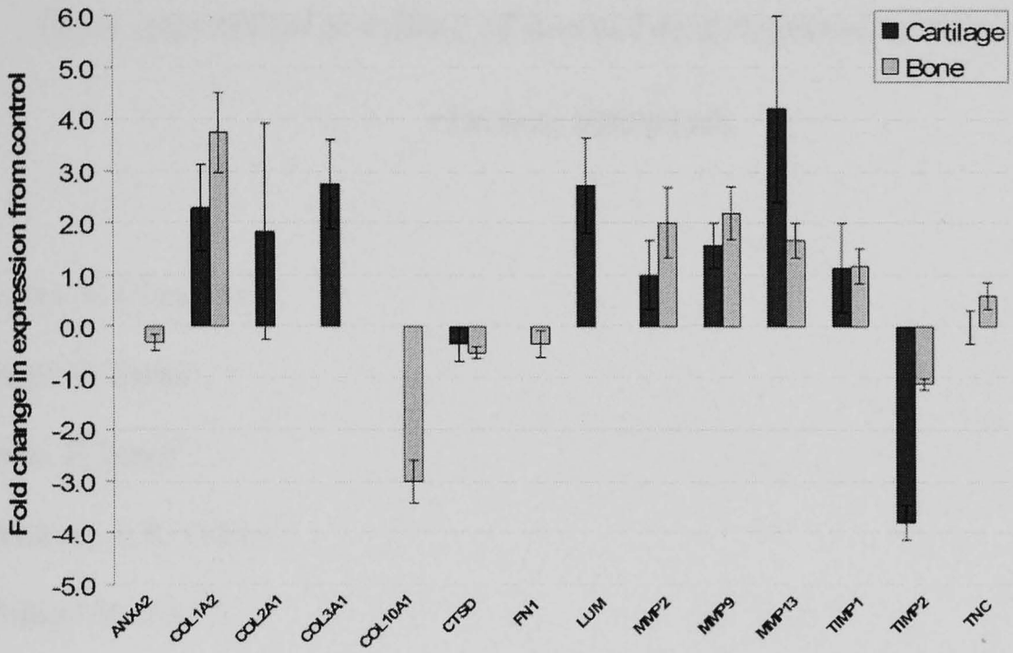
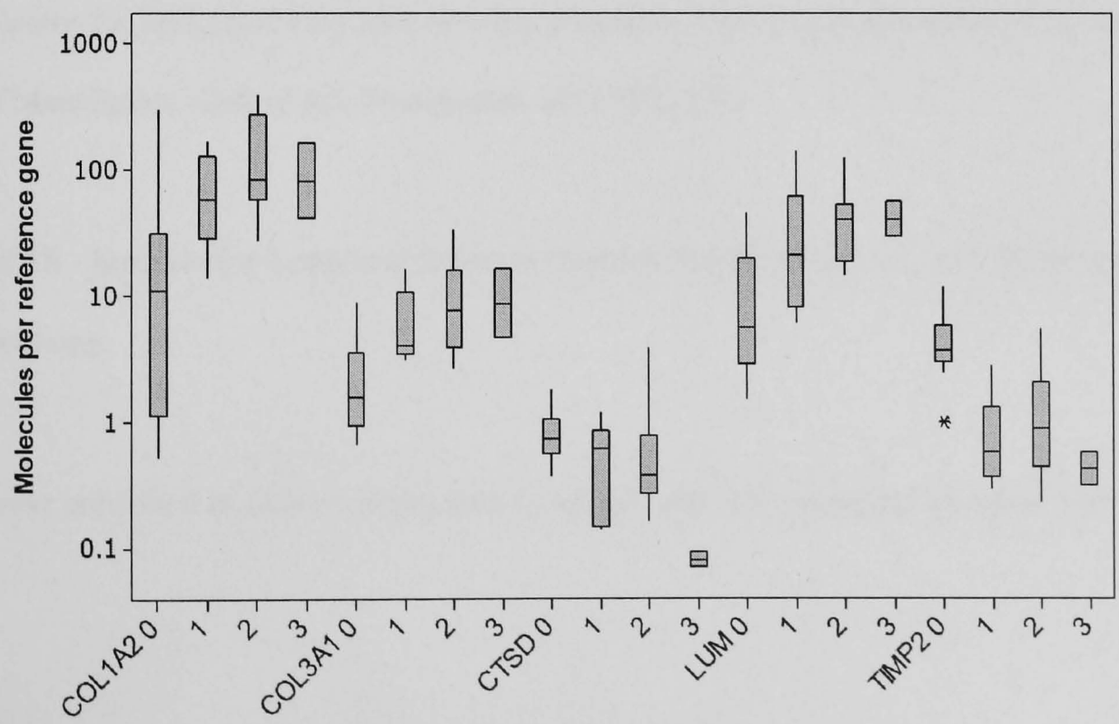


Figure 2

The number of mRNA transcripts of each gene in normal and OA articular cartilage in relation to the OA score. Median (horizontal bar), interquartile range (box) and 95% confidence intervals (whiskers) are shown, with outliers (*). The number assigned denotes the OA score of the group of values represented.



Chapter 5

Gene expression profiling of normal and ruptured canine cranial cruciate ligaments

Dylan N. Clements[1,2],

Stuart D Carter[1],

John, F. Innes[1],

William, E.R. Ollier[2],

Philip J.R. Day[2,3].

[1]Musculoskeletal Research Group, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX, UK.

[2]Centre for Integrated Genomic Medical Research, The Stopford Building, University of Manchester, Oxford Rd, Manchester, M13 9PT, UK.

[3]ISAS - Institute for Analytical Sciences, Bunsen-Kirchhoff-Str. 11, 44139 Dortmund, Germany

Paper published in *Osteoarthritis and Cartilage* (2007) E-published ahead of print

ABSTRACT

Introduction

Cranial cruciate ligament (CCL) rupture is associated with marked breed risks in canine populations, and ligament laxity is an important component of the disease. We aimed to identify genes which may be involved in the development of cranial cruciate ligament (CCL) laxity and rupture in naturally-occurring canine disease.

Materials and Methods

Three groups of dog were studied: (i) dogs with CCL rupture; (ii) dogs with intact CCLs from a breed predisposed to CCL rupture; (iii) dogs with intact CCLs from a breed at very low risk of rupture. The transcriptomes of the CCLs from each group were compared using a whole genome microarray and the quantitative reverse transcriptase (real-time) polymerase chain reaction (RT-qPCR).

Results

A general pattern of increased protease and extracellular structural matrix gene expression was identified in the ruptured CCLs when compared to intact CCLs. No significant differences were identified between the gene expression profiles of normal CCLs of a breed predisposed to CCL rupture when compared to a breed relatively resistant to CCL rupture, although a degree of risk-specific clustering was observed for expression profiles of genes which were differentially expressed in CCL rupture. A strong association was identified between the genes whose expression was changed in ruptured CCLs when compared to normal CCLs, although statistical significance for individual genes was not identified when corrected for multiple hypothesis testing.

Conclusions

The expression profiles of ruptured canine CCLs were similar to those previously reported for ruptured human ACLs. The aetiopathogenesis of the spontaneous canine model may be relevant to human knee OA. A transcriptomic basis to breed specific risk for the development of canine CCL rupture was not identified.

INTRODUCTION

Joint laxity is hypothesised to be an important contributor to the pathogenesis of human knee OA, with laxity in the valgus-varus (103) and anterior-posterior (103,104) planes increasing with the severity of OA. The anterior cruciate ligament (ACL) is the primary stabiliser of the knee joint, and rupture of this ligament results in joint instability and the development of osteoarthritis (OA) (344-346). ACL rupture is identified more commonly in patients with knee OA (102), although this is not necessarily a sequel to a previous traumatic event (102).

Pathological changes to the ACL resulting in knee laxity may predispose patients to knee OA. This hypothesis is supported by spontaneous animal models of OA which highlight the association of ligament laxity, specifically of the ACL, and the development of OA (107). A similar spontaneous condition is recognised in dogs where disease of the canine cranial cruciate ligament (CCL, the anatomical equivalent of the human ACL) results in a progressive pathological failure of the ligament, the development of joint instability and secondary OA of the affected joint (11).

Epidemiological studies have highlighted that dogs demonstrate a breed-associated risk to CCL rupture (88,108), with "at-risk" breeds such as the Labrador Retriever (LR) demonstrating much higher levels of risk than "protected" breeds, such as the Greyhound (GH). Dogs from breeds predisposed to CCL rupture have reduced ligament stiffness and reduced loading to ultimate failure when compared to dogs from breeds with low risk of CCL rupture (91,105,106). This implies that the genetic susceptibility to the development of CCL rupture manifests itself through changes in the mechanical properties of the CCL. Increased levels of pro-matrix

metalloproteinase -2 (pro-*MMP2*) have been identified in normal CCLs of dogs with a high risk of CCL rupture (LR) compared to dogs with a low risk of CCL rupture (GH) (91). Similar changes in gross CCL biomechanical properties (increasing laxity) have been related to molecular differences (increased *MMP2* and pro-*MMP2* protein) in the CCL in an animal model of spontaneous knee OA (107), further supporting the link between the development of knee OA, knee laxity and molecular changes at a cellular level.

Gene expression within the normal and pathological ACL/CCL has not been extensively studied in humans or canines. In man, the ruptured ACL expresses higher quantities of mRNA coding for Type I collagen (*COL1*), Type III collagen (*COL3*), biglycan (*BGN*) and tissue inhibitor of metallopeptidase -1 (*TIMP1*), than the normal (non-ruptured) ACL (347). In the dog, the messenger RNA (mRNA) expression of matrix metalloproteinase -2 (*MMP2*) and -9 (*MMP9*), tartrate-resistant acid phosphatase (*TRAP*) and cathepsin S (*CTSS*) have been reported to be increased in the ruptured CCL when compared to normal CCL (348), although studies evaluating the level of protein present suggest that pro-*MMP2* is increased in the ruptured canine CCL, but not active *MMP2, MMP9, TIMP1* or *TIMP2* (349).

Differences between the transcriptome of diseased and normal tissue can be identified by gene expression profiling. Canine specific microarray platforms are now available to perform genome wide expression profiling of canine cells from dissected tissues (222). Microarray analysis has allowed the identification of differential gene expression pathways in connective tissues, such as cartilage and tendon, which further

our understanding of the molecular pathways involved in OA (249) and tendon repair (350).

We hypothesized that altered gene expression profiles would be observed when comparing the normal CCL from dogs of breeds predisposed to CCL rupture (LR) with normal ligaments from dog breeds at low risk from CCL rupture (GH). Secondly, we hypothesised that, in breeds at risk of CCL rupture, differential expression of genes would be identified between the transcriptomes of normal CCL and the ruptured CCL.

MATERIALS AND METHODS

RNA extraction

Normal canine ACLs were harvested by sharp dissection and from the knees (stifles) of dogs (7 LR [mean age 5.4 years (standard deviation (SD)+/-3.3 years, range 1-10 years), 3 male neutered, 2 female neutered, 1 entire male and 1 entire female mean weight 28.7kg (SD +/-2.4kg, range 25-32kg)], 5 GH [mean age 3.8 years (SD +/-4.1 years, range 1.5-10 years), 3 entire female, 2 entire male, mean weight 33.0kg (SD +/-3.8kg, range 30-38kg)]) without any evidence of knee pathology, and which were euthanatized for reasons unrelated to orthopaedic disease. The central third of the ligament was preserved. Ruptured canine ACLs were obtained from 5 LRs (mean age 7.7 years [SD+/-1.6 years, range 5.5-10 years], 3 neutered male, 1 neutered female, 1 male, mean weight 33.4kg [SD +/-3.0kg, range 30-37kg]) during routine surgical treatment for the ACL rupture (medial parapatellar arthrotomy), and stored in RNAlater as recommended by the manufacturer (Qiagen Ltd, Crawley, RH10 9NQ,

UK), at room temperature for 24 hours, then at -20°C until use. All dogs with ACL rupture had radiographic evidence of stifle OA (osteophytosis).

RNA extraction

Tissue samples were removed from RNAlater and total RNA was extracted using phenol / guanidine HCl reagents (Trizol, Invitrogen Ltd, Dorset, UK) and isolated as previously described (258,276). An on-column DNA digestion step was included (RNase-Free DNase Set, Qiagen Ltd). Final elution of the total RNA was performed using 30 µl of RNase free water, and repeated to maximise the recovery of RNA.

RNA quality assessment

The concentration of total RNA of each sample was quantified in a spectrophotometer (NanoDrop technologies Ltd, Utah, USA). RNA integrity was analysed by evaluating the capillary electrophoresis trace (Agilent 2100 Bioanalyser, Agilent Technologies, California, USA) of the sample using the RNA integrity number [RIN] algorithm (224), Degradation Factor [DF] (230) and ribosomal peak ratio. The samples determined to have no, or minimal loss of integrity (RIN > 6.4, and / or DF < 10, and / or a ribosomal ratio > 0.4) were deemed suitable for use in experiments in accordance with a previously developed quality algorithm (276). For full details of the RNA amplification and labelling methods see Chapter 4, supplementary material.

RNA Amplification

Messenger RNA (mRNA) was amplified for each sample, starting with 200 ng total RNA using a commercially available kit (Ambion T7 MEGAscript high yield transcription kit, Ambion [Europe] Ltd, Cambridge, UK) as previously described

(297). A second round mRNA amplification was performed using MessageAmp aRNA Amplification kit (Ambion [Europe] Ltd), following the manufacturer's instructions. The amplified RNA (aRNA) was quantified using a spectrophotometer.

aRNA Labelling

2 μg of aRNA was labelled by reverse transcription with Cyanine-3dCTP (Cy3) or Cyanine-5dCTP (Cy5), using a fluorescent dye labelling kit (Agilent Technologies UK Ltd, South Queensferry, UK) following the manufacturer's instructions. Fluorescent dye incorporation was determined using a spectrophotometer, ensuring that > 750 ng complementary RNA (cRNA) was labelled, and that the label incorporation was > 8 pmol per μg RNA. Samples were stored at -80°C until use.

Microarray hybridization and slide reading

750 ng of both Cy3 and Cy5 cRNA was fragmented and hybridized to a canine-specific, custom designed, whole genome 44219 spot 60mer oligonucleotide microarray chip (298) at 65°C for 17 hours using the manufacturer's protocol (Agilent Technologies UK Ltd). Slides were washed according to the manufacturer's instructions, read using an Agilent DNA Microarray Slide Reader, and fluorescence data extracted by employing the Agilent Feature Extraction 8.5 software (Agilent Technologies UK Ltd). DNA spots were automatically located and subtracted from the intensity of the local backgrounds. Where intensities of the spots were below set thresholds, data was discarded from further analysis. Spots were flagged if they exhibited poor hybridization signals or when they were saturated.

One hybridization (GH) was determined to be a quality control hybridization technical failure by the software, thus only four arrays were used for the GH data analysis.

Data Normalisation and Statistics

Data were imported into Genedata Expressionist Analyst (Genedata AG, Basel, Switzerland), and the Cy3 and Cy5 fluorescence intensities normalised using lowest weighted linear regression (LOWESS) (299). Ultimately 41,623 spots (94%) coding for transcripts were considered acceptable for application in data analysis. Expression data were then exported into Excel 2003 and comparison between groups was performed using paired Students t-tests.

Comparisons of the number of genes up- or down-regulated in both the normal GH CCL and ruptured LR CCL when compared to the normal LR CCL were made using Chi squared analysis (Minitab v14.1, Minitab Ltd, Coventry, UK). Correction for multiple hypothesis testing was performed using the Benjamini and Hochberg false discovery rate (FDR) (281). Correct *P* values were calculated by dividing the true *P* value by the individual correction factor and multiplying by 0.05.

Real-Time Polymerase Chain Reaction (RT-qPCR)

Synthesis of complementary DNA

The original (un-amplified) mRNA samples used for the microarray experiments, and further suitable high grade RNA sample collections from further patients (totals, n = 21 ruptured CCL, n = 13 normal LR CCLs and n = 7 normal GH CCLs) were obtained and used in RT-qPCR experiments. Reverse transcription was performed

using 10 µl RNA (200 µg total RNA) with oligo-dT_{12-18} and Superscript II reverse transcriptase (Invitrogen, Dorset, UK). Following reverse transcription the template was diluted with 500 µl RNase /DNase free water. Complementary DNA (cDNA) was stored at -80°C until subsequent analysis by RT-qPCR.

Assay design

Assay sequences were obtained from the canine genome database (277), with cross reference to the National Centre for Biotechnology Information (278). BLAST searches were performed for all primer sequences to verify gene specificity.

Genes were selected for assay on the basis of their perceived relevance to the CCL extracellular matrix (from literature review), and from the results of the microarray screen. Assays were designed for quantification of expression of 23 genes of interest and 5 reference genes. The primer and probe sequences for each assay have been previously published (296) or are listed in Table 1. The assays were used to quantify; three collagen genes (Type I collagen, alpha two chain [*COL1A2*], Type III collagen, alpha 1 chain [*COL3A1*], Type V collagen, alpha 1 chain [*COL5A1*]), seven extracellular matrix genes (aggrecan [*AGC1*], biglycan [*BGN*], chondroitin sulphate proteoglycan 2 [Versican/*CSPG2*], decorin [*DCN*], lumican [*LUM*], tenascin C [*TNC*], vimentin [*VIM*]), proteases and their inhibitors (a disintegrin and metalloproteinase with thrombospondin-like motif -4 [*ADAMTS4*], -5 [*ADAMTS5*], cathepsin B [*CTSB*] and D [*CTSD*], matrix metalloproteinases -2, -9, -13 [*MMP2, MMP9, MMP13*], caspase-8 [*CASP-8*], tissue inhibitors of metalloproteinase -1 [*TIMP1*] and -2 [*TIMP2*]), a growth factor, (insulin like growth factor-1 [*IGF1*]),

prostaglandin -G, -H synthetase-2 (*PTGS2*), inducible nitric oxide synthetase 2A (*NOS2A*) and genomic DNA (*GEN*).

Reference genes were selected using a previously published reference gene stability algorithm (238). Five reference genes (beta 2-microglobulin [*B2M*], glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], TATA box binding protein [*TBP*], ribosomal protein L13a [*RPL13A*] and succinate dehydrogenase complex, subunit A [*SDHA*]) were assayed in all samples. Again, the primer and probe sequences for each assay have been previously published (296) or are listed in Table 1.

Primers were synthesized by MWG Biotech (London, UK). Locked nucleic acid probes with a 5' reporter dye FAM (6-carboxy fluorescein) and 3' quencher dye TAMRA (6-carboxytetramethyl-rhodamine) were synthesized by Roche Diagnostics Ltd (Lewes, West Sussex, UK).

The quantitative (real-time) reverse transcriptase polymerase chain polymerase reaction (RT-qPCR) assays were all performed in triplicate using a TaqMan™ ABI PRISM 7900 SDS (Applied Biosystems, California, USA) in 384-well plate format. Each assay well had a 10 µl reaction volume consisting of 5 µl 2X PCR master mix with uracil N-glycosylase (Universal PCR Mastermix, Applied Biosystems, California, USA), 0.1 µl each of 20 µM forward and reverse primers, 0.1 µl of 10 µM probe (ProbeLibrary, Roche Diagnostics, Lewes, UK) and 4.7 µl of sample cDNA (templates) or water (negative controls).

The amplification was performed according to a standard protocol with 2 minutes at 50°C and 10 minutes at 95°C, followed by 40 cycles of 95°C for 1 min and 60°C for 15 sec, as recommended by the manufacturer (Applied Biosystems, California, USA). Real-time data were analyzed by using the Sequence Detection Systems software, version 2.2.1 (Applied Biosystems, California, USA). The detection threshold was set manually for all assays at 0.05.

RT-qPCR Data Analysis

RT-qPCR transcript data was produced in triplicate for each sample and analysed by generation of mean C_T values. Geometric means (238) were calculated for the combined four reference genes (*B2M, SDHA, RPL13A, TBP*), and used to calculate the delta-delta C_T values and the relative amount of each target gene (280). The fifth reference gene (*GAPDH*) was not included as a reference gene because it had near-differential expression between the normal and ruptured ligament samples, even when included as part of the normalisation calculation.

RT-qPCR data for each group were compared with the calculations of means, standard deviations, fold changes from normal and paired two-tailed t-tests (bodyweight and age) performed in a spreadsheet program (Microsoft Excel 2003). Significance was established at $P < 0.05$, and data was checked for errors due to multiple hypothesis testing using the Benjamini and Hochberg false discovery rate (FDR) (281).

Cluster Analysis

The normalised microarray data for sixty-three genes differentially expressed between diseased (LR) CCL when compared to normal (LR) CCL and with complete annotation were loaded into a gene clustering software (Cluster, Eisen Labs (301)). Data was log transformed and genes centred to the mean. Hierarchical clustering of differentially expressed genes was then performed for arrays and genes using Spearmans Rank Correlation and complete linkage link clustering. Clustering of genes and arrays were visualised with publicly available software (TreeView, Eisen Labs (351)).

RESULTS

Comparison of patient signalment

A significant difference (p=0.025) was noted in the patient weight of tissues used in the microarray experiment to compared the ruptured ACL LRs and the normal ACL LRs. A second significant difference (p=0.024) was noted in neuter status of the patients used in the RT-qPCR experiment to compared the ruptured ACL LRs (more likely to be neutered) and the normal ACL LRs. No other differences in age, weight sex or neuter status were determined.

Microarray

In the normal GH CCL, compared to the normal LR CCL, 925 transcripts were up-regulated ($P < 0.05$). Conversely 1050 transcripts were down-regulated in the normal GH CCL, compared to the normal LR CCL. Of the 925 transcripts up-regulated in the normal GH CCL, 455 were also significantly up-regulated in the ruptured LR CCL when compared to the normal LR CCL, and 450 of the 1070 transcripts were down-

regulated in both the normal GH CCL and LR CCL when compared to the normal LR CCL. The number of transcripts whose expression was increased or decreased in both normal GH CCL and ruptured LR CCL was significantly greater (P <0.001) than by chance alone. The FDR determined that none of the transcripts were significantly up- or down- regulated when corrected for multiple hypothesis testing.

4038 transcripts were up-regulated in ruptured LR CCL when compared to normal LR CCL. 5419 transcripts were down-regulated in ruptured LR CCL when compared to the normal LR CCL. The FDR determined that 99 transcripts were significantly up-regulated and 17 transcripts were significantly down-regulated when corrected for multiple hypothesis testing. 87 transcripts (of which 24 transcripts were repeats) had a defined annotation; 29 transcripts had no defined annotation. The annotated transcripts whose expression was changed by 3 fold or more are listed in Table 2.

Real-time PCR

The results of the real-time RT-qPCR are presented in Table 3. Two genes (*COL5A1* and *RPL13A*) were determined to be up-regulated ($P < 0.05$) in the normal LR CCL when compared to the normal GH CCL. The FDR determined that neither gene was significantly up-regulated.

Sixteen genes were significantly up- (n = 14) or down- (n = 2) regulated in the ruptured LR CCL, when compared to the normal LR CCL. The FDR determined that fourteen of these genes were significantly up-regulated, after correcting for multiple hypothesis testing (*AGC, CASP8, COL1A2, COL3A1, COL5A1, CTSB, CTSD, IGF1,*

LUM, MMP2, MMP9, PTGS2, TIMP1, TNC) and two were significantly down-regulated (*DCN, TIMP2*).

When the real-time RT-qPCR results for the sample subsets used for microarray analysis were examined separately, twelve genes were significantly up-regulated in the ruptured LR CCL when compared to the normal LR CCL, of which eight genes were still significantly up-regulated after FDR correction (*AGC, CASP8, COL1A2, COL3A1, CTSB, MMP9, TIMP1* and *TNC*). One gene was determined as being up-regulated in the normal LR CCL when compared to the normal GH CCL (*IGF1*), although this was not significant when corrected for multiple hypothesis testing.

When FDR correction was applied to microarray data, but only to the 24 candidate genes evaluated by RT-qPCR, then eleven genes were determined to be significantly up-regulated (*CASP8, COL1A2, COL3A1, COL5A1, CTSB, GAPDH, IGF1, LUM, MMP13, TIMP1, TNC*). Nine of these genes, (with the exceptions being *MMP13* and *GAPDH*) were confirmed to be significantly up-regulated in the larger cohort of samples evaluated by RT-qPCR. When FDR correction was applied to the complete microarray data set (41623 genes), only two genes evaluated by RT-qPCR were differentially expressed (*CASP8* and *COL3A1*). Both these genes were also differentially expressed in the larger cohort of samples evaluated by RT-qPCR. Overall, for genes significantly up- or down-regulated by either microarray (n = 2), RT-qPCR of a larger cohort of samples (n = 9) or both techniques (n = 7), the direction fold change in expression agreed for all except one gene (*TIMP2,* down-regulated in RT-qPCR samples). A graphical representation of the fold changes in

expression of genes up-regulated in ruptured LR CCL, as determined by microarray, is illustrated in Figure 1.

Cluster Analysis

The corresponding hierarchical cluster analysis plot is presented in Figure 2. Clustering of the arrays illustrates clear separation of the diseased and normal samples. This is unsurprising as the list of genes clustered was filtered on the basis of differential expression between normal and diseased samples. Samples of three of the four dogs from a breed at low risk of CCL rupture (GH) demonstrated patterns of gene expression which clustered together, and which were more closely related to ruptured CCL than those from dogs at high-risk of CCL rupture. Clustering of the genes demonstrated alignment of genes known to have similar function, such as *TUBA, TUBB, ACTA* and *ACTB* within closely related branches, and supports both the method of analysis and the validity of the data.

DISCUSSION

Differential gene expression was identified in ruptured CCL when compared to normal LR CCL using both expression profiles generated by microarray and RT-qPCR. The overall pattern of gene expression reported in the ruptured CCL suggests that both catabolism (MMP and cathepsin production) and repair (collagen and extracellular matrix production) (352) are increased in the ruptured canine CCL when compared to the normal CCL. The changes in gene expression are consistent with both histological features of collagen disruption and epiligamentous repair (3) and molecular changes (348) reported in the ruptured human ACL. To date, the results of

primary repair of ruptured CCLs have been poor both in man (353) and animal models (352). Thus the anabolic response of the ruptured CCL, which has lost its ability to resist mechanical load, would appear to be futile.

A transcriptomic basis for breed risk to CCL rupture was not identified. Over 40% of genes up- or down- regulated in the GH-CCLs (before correction for multiple hypothesis testing) showed a similar differential expression as observed in the transcriptome of ruptured LR CCLs, suggesting that expression profiles reported for normal GH CCLs were more akin to those identified in the ruptured LR CCLs. Furthermore, hierarchical cluster analysis of the most differentially expressed genes in ruptured CCL demonstrated clustering and separation of the expression profiles of three of the four low-risk (GH) normal CCL samples from the high-risk (LR) normal CCL samples. Thus, a transcriptomic basis to the breed specific risk may exist, but our methods were not sensitive enough to characterise it. Alternatively the small sample size evaluated by microarray may have dictated that individuals in the high-risk group were simply not at risk of CCL rupture, although the breed itself is, because CCL rupture does not affect all individuals in the breed. Conversely, one may interpret the results to indicate that a transcriptomic risk to the development of CCL rupture truly does not exist, despite differences in relative risk of CCL rupture between dog breeds (88,108), and biomechanical (91,105) and biochemical (91,354) differences in the intact CCL between dog breeds protected or at risk of CCL rupture. Clearly, there are advantages to using laboratory animal models of disease which demonstrate a consistent phenotype, thus providing more homogeneous information as to the molecular basis to the disease. However, such models may not always reflect

polygenetic nature to a disease, or the associated environmental influences, which the canine population shares with its human counterparts.

Although no pathways were consistently represented in the differentially expressed genes as determined by microarray, a number of interesting and functionally related rupture associated genes were up-regulated. *ACTA, ACTB, TUBB* and *TUBA* are all genes encoding intracellular structural molecules, whose up-regulation suggests an increase in cytoplasmic activity in ruptured LR CCLs. *SPARC*, an extracellular matrix protein which is involved in ligament development, remodelling and repair (355), was also increased in ruptured LR CCLs, suggesting that the gene may have a key role in the anabolic response to CCL rupture. The majority of the genes identified as being differentially expressed in ruptured CCL have no previous known association with OA or ligament pathology, which makes their precise role in the ruptured CCL difficult to define.

The normal resorption of ACL matrix collagen has been hypothesised to occur by fibroblast phagocytosis and intracellular digestion with lysosomal cathepsins, whereas inflammatory remodelling of collagen is thought to be mediated by MMPs (356). Our results suggest that both processes are active, as there were increases of both cathepsin (B and D) and matrix metalloproteinase (2 and 9) expression in the ruptured CCL, although the relative importance of expression changes in each of the molecules requires further study. Ligament *CTSD* expression increases with mechanical stress (357), and the experimental induction of immune-mediated synovitis increases *CTSD* in synovial fluid, with concomitant reduction in the mechanical properties of the CCL (358). At the protein level, proMMP-2 is raised in ruptured CCLs in dogs (349).

Increased *TIMP1* expression was identified in ruptured CCLs, which is consistent with evaluation of gene expression in ruptured human ACLs (347), although this increase does not reflect enzyme activity in the ruptured canine CCLs (349). The reduced expression of *TIMP2* in CCL rupture is interesting, because this is contrary to the change one may expect on the basis of its biological activity yet is consistent with similar reductions in the expression of this gene reported in osteoarthritic articular cartilage (296,327) and reductions in the level of protein of this gene reported in the synovial fluid of dogs with CCL rupture (340).

Up-regulation of extracellular matrix gene expression (*COL1A2, COL3A1, COL5A1, AGR, LUM* and *TNC*) suggests increased extracellular matrix production as part of an attempted reparative process. This concurs with reports documenting an increased expression of *COL1* and *COL3* in the ruptured human ACL when compared to normal (non-ruptured) ACLs (347) although we did not record an increased expression of biglycan in ruptured canine CCLs, as has been reported for ruptured human ACLs (347). Reduced expression of a number of these genes (*COL1, COL3, DCN* and *LUM*) is reported in the CCL of rabbits during pregnancy, and are associated with a concurrent increase in CCL laxity (359). We did not record a detectable difference in the expression of these genes in the normal CCL of two breeds known to demonstrate different laxity measurement (91). Interestingly a similar pattern of gene expression changes (increased *COL1A2, COL3A1, COL5A1, LUM* and *TNC* gene expression) is also reported in end-stage canine hip OA cartilage (296), which suggests that these changes reflect a more primordial mesenchymal tissue response in OA. Tenascin C expression is directly related to mechanical load in ligamentous fibroblasts (360), thus it increased expression in the ruptured CCL may represent a response to increased

loading prior to rupture, or an attempt at ligament repair, and similarly the increase in *IGF1* expression in the ruptured CCL is consistent with that reported role of this gene in models of tendon repair (361).

Histological changes are reported in other species following CCL rupture or transection, such as ligament remodelling (362) and ligament resorption (352), and vary with time post injury. Thus the precise cellular composition and activity within the ruptured CCL are in part a reflection of the length of time post injury (362), and as such the selection of tissues on the basis of time post-injury, or histological grading of cellular composition could have reduced the heterogeneity of the data produced. The insidious onset of CCL rupture in dogs (11) dictates that the precise timing of CCL injury is very difficult to determine (363). Other variables which were not controlled for in the study populations were sex, weight and neuter status. Although the increased risk of ACL rupture for females reported in humans (364) is reported in some (97) but not all (88) canine epidemiological studies, neuter status is associated with a risk of CCL rupture, with neutering increasing the risk of CCL rupture (88,97). Indeed there was a significant increase in the number of neutered dogs in our LR CCL rupture population, and these dogs were signficiantly heavier than the LR controls. Clearly the finer nuances may have been lost for data evaluating ruptured CCLs, and this may explain in part the limited number of differentially expressed genes as determined by microarray, when compared to RT-qPCR of a large sample set, and the large variation in fold change in expression of a number of the genes evaluated.

The highly significant increase in expression of alpha actin in the ruptured CCL, as determined by microarray, implies that the cellular differentiation of cells in the epiligamentous synovial layer to myofibroblasts was occurring, as has been reported in the proliferative phase of the ruptured human ACL (362). Furthermore, the significant increases in other matrix components suggested a proliferative response, whilst an epiligamentous inflammation recorded on histological studies of ruptured canine CCL (365) is indicated by the increase in *PTGS2* expression.

CONCLUSIONS

In summary, we could not identify a transcriptomic basis to the breed-specific risk for the development of canine CCL rupture, although a large number of new genes were identified as being differentially expressed in ruptured CCL. The expression profiles of ruptured CCLs were similar to those previously reported for ruptured human ACLs (347).

Figure 1

Fold change in gene expression (plus standard deviation) of candidate genes in ruptured CCL when compared to normal CCL from high-risk dogs (LR) as determined by microarray (MA), RT-qPCR of MA samples (MA RT-PCR), and RT-qPCR of all samples (RT-PCR). The candidate genes selected for comparison are those differentially expressed by microarray evaluation.

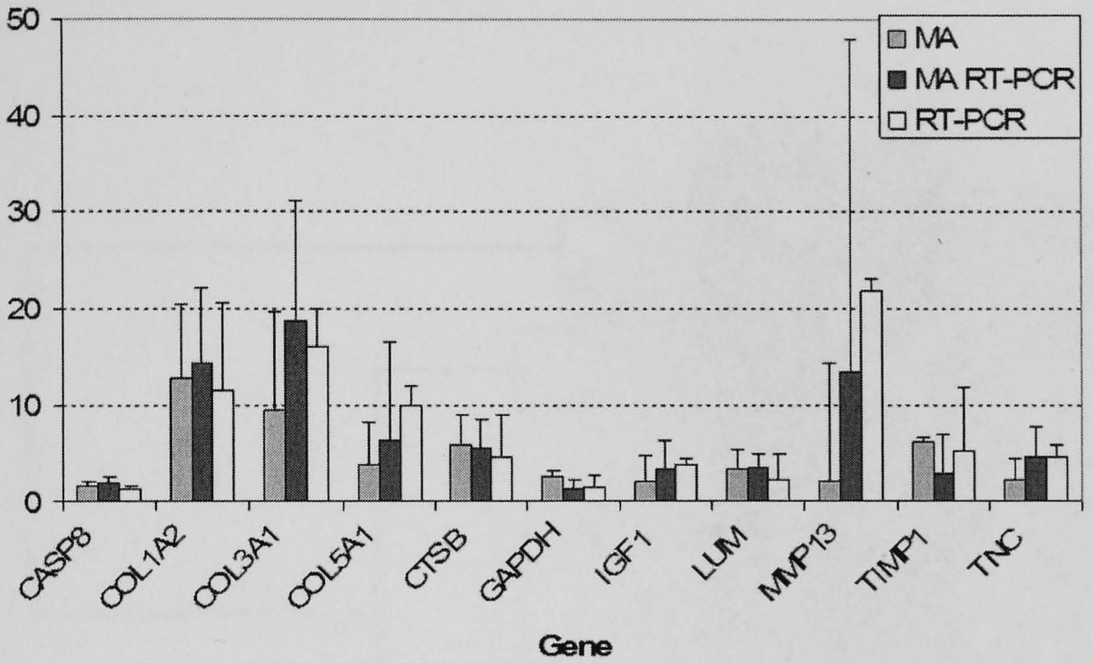


Figure 2

Heat map and hierarchical clustering plot of differentially expressed genes and arrays generated from canine ruptured (suffixed LD) and normal (suffixed LN) cranial cruciate ligament (CCL) from high-risk dogs (LR), and normal CCL from low risk dogs (suffixed GN). Colour coding; Green= Decreased expression, red= increased expression, black= no change in expression, grey=missing value).

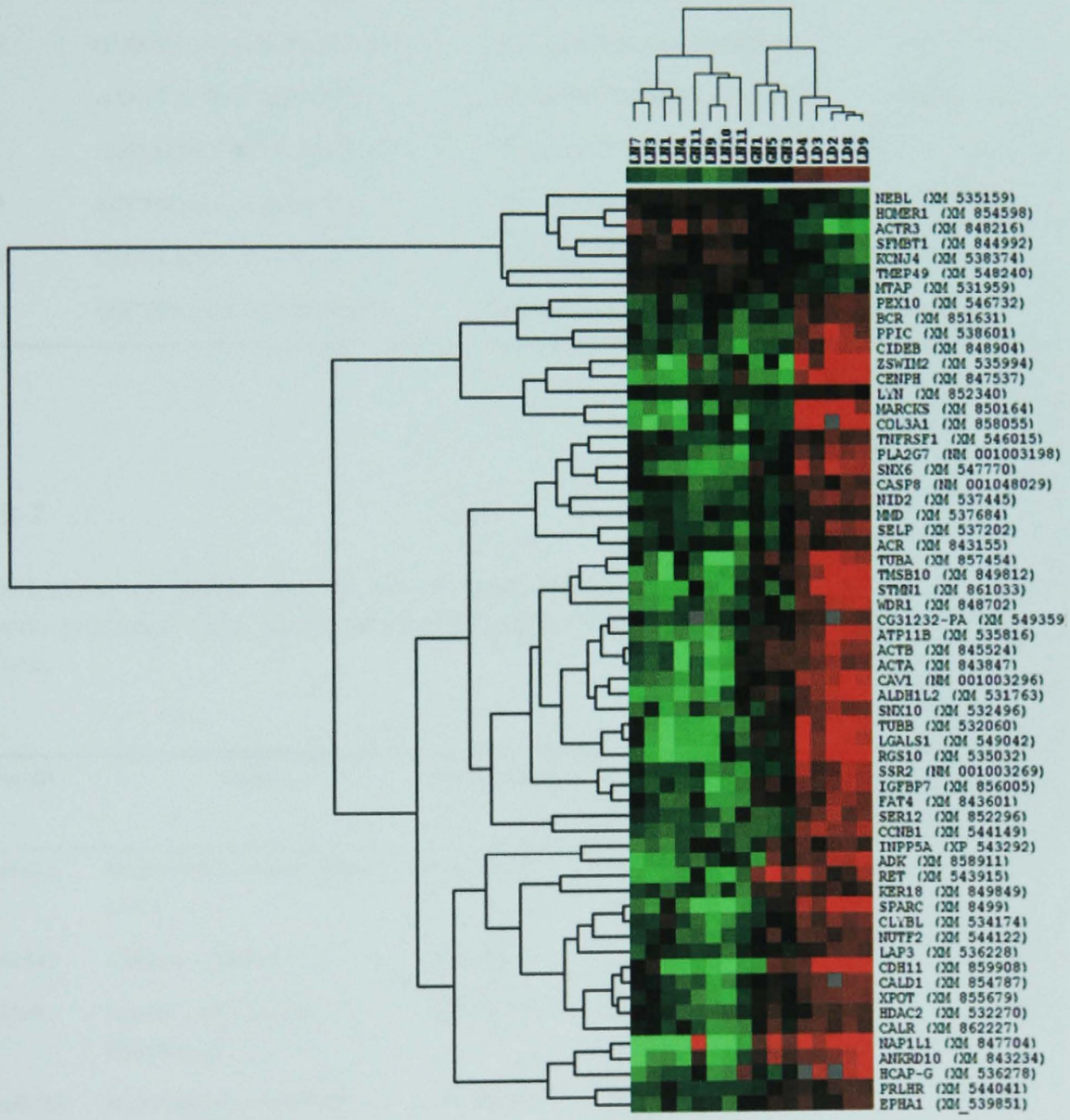


Table 1

Primer and probe sequences for quantitative PCR assays.

Gene	*Forward*	*Reverse*	*Probe*
ADAMTS4	GACCAGTGCAAACTCACCTG	CAGGGAGTCCCATCTACCAC	GGCCCTGG
B2M	CCTTGCTCCTCATCCTCCT	TGGGTGTCGTGAGTACACTTG	CAGCATCC
CASP8	GAGCTTCAGATACCAGGCAGA	TGAAATCTGAAAAAGCATGACC	CTCTGCCT
COX2	AAATTGCTGGCAGGGTTG	TCGAAGCTTTTGCTACTTGTTG	GGTGGCAG
IGF1	GGGGGTTCTACTTCAACAAGC	TCATCCACGATGCCTGTCT	CTCCAGCA
MMP2	ACCTGCAAGGCAGTGGTC	TCCAAATTTCACGCTTTTCA	AGCTGGAG
MMP9	CACGCATGACATCTTCCAGT	CGAGAATTCACACGCCAGTA	CTTCTGCC
NOS2A	GGCTCAAATCACAACGGAAT	AGAGCTCGACCAGGAGAGTG	CCAGCCGC

Table 2

Identification of genes up- or down-regulated (3 fold or more) when compared between ruptured CCL and normal CCL in Labrador Retrievers in the microarray analysis.

Gene ID	*Name*	*Ref Seq Number*	*Function*	*Fold Change*	*Corrected P Value*
NAP1L1	Nucleosome Assembly Protein 1-Like 1	XM_847704	nucleic acid processing	19.0 +/- 7.4	0.0373
COL3A1	Collagen 3, Alpha 1	XM_858055	structure	9.5 +/- 4.0	0.0328
RGS10	Regulator Of G Protein Signalling 10	XM_535032	signalling	8.2 +/- 2.5	0.0220
MARCKS	Myristoylated Alanine-Rich Protein Kinase C Substrate	XM_850164	structure	7.8 +/- 2.2	0.0189
CDH11	Cadherin 11, Type 2, OB-Cadherin	XM_859908	structure	6.9 +/- 2.4	0.0476
SPARC	Osteonectin	XM_8499	signalling	6.3 +/- 2.1	0.0387
STMN1	Stathmin 1	XM_861033	structure	6.2 +/- 1.7	0.0194
ZSWIM2	Zinc Finger, SWIM Domain Containing 2	XM_535994	metabolism	5.9 +/- 1.9	0.0328
TMSB10	Thymosin Beta-10	XM_849812	structure	5.7 +/- 1.9	0.0388

Gene ID	*Name*	*RefSeq Number*	*Function*	*Fold Change*	*Corrected P Value*
LGALS1	Lectin, Galactose Binding, Soluble 1	XM_549042	signalling	5.4 +/- 1.3	0.0195
CENPH	Centromere Protein H	XM_847537	nucleic acid processing	5.4 +/- 1.8	0.0480
ALDH1L2	Aldehyde Dehydrogenase 1 Family, Member L2	XM_531763	metabolism	5.3 +/- 1.6	0.0345
ANKRD10	Ankyrin Repeat Domain 10	XM_843234	nucleic acid processing	5.2 +/- 1.3	0.0207
RET	RET Tyrosine Kinase/Camp Protein Kinase A Subunit RI	XM_543915	signalling	4.9 +/- 1.7	0.0478
CAV1	Caveolin 1	NM_001003296	structure	4.8 +/- 0.9	0.0137
TUBB	Tubulin Beta	XM_532060	structure	4.8 +/- 1.4	0.0255
TUBA	Alpha-Tubulin	XM_857454	structure	4.7 +/- 1.0	0.0195
ATP11B	ATPase, Class I, Type 11B	XM_535816	metabolism	4.7 +/- 1.5	0.0381
WDR1	WD Repeat Domain 1	XM_848702	structure	4.4 +/- 1.3	0.0263
SNX6	Sorting Nexin 6	XM_547770	transport	4.4 +/- 0.9	0.0158
ACTB	Beta Actin	XM_845524	structure	3.8 +/- 0.8	0.0188
IGFBP7	Insulin-Like Growth Factor Binding Protein 7	XM_856005	signalling	3.7 +/- 0.8	0.0196
ADK	Adenosine Kinase, Transcript Variant 3	XM_858911	signalling	3.6 +/- 1.1	0.0472
CCNB1	Cyclin B1	XM_544149	structure	3.6 +/- 0.4	0.0009
SSR2	Signal Sequence Receptor, Beta	NM_001003269	signalling	3.5 +/- 0.9	0.0342
CIDEB	Cell Death-Inducing DNA Fragmentation Factor, Alpha Subunit-Like Effector B	XM_848904	signalling	3.2 +/- 0.9	0.0477
FAT4	FAT Tumour Suppressor Homolog 4	XM_843601	signalling	3.0 +/- 0.7	0.0341
ACTR3	Actin-Related Protein 3	XM_848216	structure	0.3 +/- 0.1	0.0379

Table 3

Gene expression folds change and significance (uncorrected P Values) in ruptured Labrador Retriever CCL (D), normal Labrador Retriever CCL (N) and normal greyhound CCL (GH) as determined by microarray (MA), RT-qPCR of MA samples (MA RT-qPCR), and RT-qPCR of all samples (Significant results, surviving FDR correction (to 24 genes) are in bold).

	MA				*MA RT*				*ALL RT*			
Gene	**Ratio LD vs LN**	**P Value**	**Ratio GH vs LN**	**P Value**	**Ratio LD vs LN**	**P Value**	**Ratio GH vs LN**	**P Value**	**Ratio LD vs LN**	**P Value**	**Ratio GH vs LN**	**P Value**
ADAMTS4	1.0	0.982	1.1	0.408	**0.9**	0.881	51.7	0.365	2.7	0.084	3.4	0.558
ADAMTS5	0.8	0.605	1.6	0.286	**0.3**	0.248	4.4	0.278	0.6	0.217	1.9	0.427
AGC	3.9	0.043	0.5	0.126	**13.9**	**0.007**	0.9	0.759	14.3	**0.000**	1.1	0.814
BGN	1.7	0.086	0.9	0.795	**0.9**	0.916	1.3	0.702	1.6	0.133	1.2	0.652
CASP8	1.8	**0.000**	0.9	0.479	**2.1**	**0.003**	0.9	0.630	1.5	**0.028**	0.9	0.574
COL1A2	12.8	**0.001**	0.7	0.306	**14.3**	**0.004**	0.9	0.849	11.5	**0.000**	1.0	0.905
COL3A1	9.5	**0.000**	0.6	0.086	**18.7**	**0.004**	0.9	0.839	16.0	**0.000**	1.1	0.854
COL5A1	3.9	**0.004**	0.9	0.727	**6.4**	0.025	0.5	0.117	9.9	**0.003**	0.5	0.019
CSPG2	0.8	0.078	0.8	0.818	**1.0**	0.893	0.7	0.498	1.0	0.910	0.7	0.319
CTSB	6.0	**0.012**	0.7	0.293	**5.6**	**0.010**	0.9	0.499	4.7	**0.000**	1.0	0.920
CTSD	1.0	0.475	1.0	0.756	**1.9**	0.201	2.1	0.115	2.9	**0.001**	1.3	0.389
DCN	0.7	0.100	0.9	0.742	**0.5**	0.053	0.7	0.175	0.5	**0.000**	0.8	0.462
GAPDH	2.6	**0.011**	0.6	0.124	**1.4**	0.327	1.1	0.606	1.6	0.043	1.2	0.204
IGF1	2.2	**0.001**	1.3	0.029	**3.4**	0.086	4.3	0.042	3.8	**0.003**	2.0	0.071
LUM	3.5	**0.012**	0.4	0.035	**3.5**	0.025	0.7	0.226	2.4	**0.002**	0.8	0.238
MMP13	2.2	**0.023**	1.2	0.420	**13.5**	0.051	3.3	0.449	21.9	0.038	2.7	0.402
MMP2	36.1	0.042	1.5	0.345	**34.1**	0.032	1.2	0.612	24.5	**0.000**	1.1	0.793
MMP9	1.5	0.068	1.1	0.758	**336.4**	**0.013**	1.2	0.826	491.9	**0.018**	1.5	0.526
NOS2	0.7	0.515	1.8	0.450	**0.0**	0.324	20.1	0.395	0.2	0.225	9.6	0.445
PTGS2	1.0	0.990	1.3	0.228	**4.9**	0.033	1.4	0.622	6.0	**0.028**	0.7	0.474
TIMP1	6.2	**0.020**	0.9	0.791	**3.0**	**0.002**	1.5	0.492	5.3	**0.001**	1.0	0.952
TIMP2	1.0	0.889	0.6	0.143	**0.5**	0.107	0.6	0.180	0.6	**0.006**	0.7	0.132
TNC	2.4	**0.008**	1.1	0.136	**4.6**	**0.008**	0.9	0.897	4.7	**0.000**	0.9	0.650
VIM	1.1	0.607	0.7	0.125	**1.1**	0.601	0.9	0.567	1.1	0.418	1.1	0.593

Chapter 6

The identification of new reference genes from microarray data for the normalisation of canine osteoarthritic joint tissue transcripts

Lindsey J. Maccoux[1,2*],

Dylan N. Clements[1,2*],

Fiona Salway[1],

Philip J. Day[1,3].

[1]Centre for Integrated Genomic Medical Research, University of Manchester, The Stopford Building, Oxford Rd, Manchester, M13 9PT. UK

[2]The Musculoskeletal Research Group, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX, UK

[3]ISAS - Institute for Analytical Sciences, Bunsen-Kirchhoff-Str. 11, 44139 Dortmund, Germany

Paper published in *BMC Molecular Biology* (2007) **8**, 62

*These authors contributed equally to this work.

ABSTRACT

Introduction

Comparison of gene expression in different samples measured by the real-time reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) requires the transformation of data by the comparison of the relative expression of the target gene to that of an internal "reference" (or housekeeping) gene. Reference genes are candidate genes which are selected on the basis of their constitutive expression across samples, and are ideally unaffected by the disease process being investigated.

Materials and Methods

Microarray data was filtered to identify new reference genes generated from total RNA isolated from normal and osteoarthritic (OA) canine articular tissues (bone, ligament, cartilage, synovium and fat). RT-qPCR assays were designed and applied to each different articular tissue. Reference gene expression stability and ranking was compared using three different mathematical algorithms.

Results

Twelve new potential reference genes were identified from microarray data. One gene (mitochondrial ribosomal protein S7) was stably expressed in all five of the articular tissues evaluated. One gene (HIRA interacting protein 5 isoform 2) was stably expressed in four of the tissues evaluated. A commonly used reference gene (Glyceraldehyde-3-phosphate dehydrogenase) was not stably expressed in any of the tissues evaluated. Most consistent agreement between rank ordering of reference genes was observed between Bestkeeper and geNorm algorithms.

Conclusions

Microarray data normalised by a conventional manner can be filtered using a simple stepwise procedure to identify new reference genes, some of which will demonstrate good measures of stability. Different methods of reference gene stability assessment will generally agree on the most and least stably expressed genes.

INTRODUCTION

The real-time reverse transcriptase polymerase chain reaction (RT-qPCR) provides the most accurate and specific measure of gene expression, with an unsurpassed dynamic range and a high level of reproducibility. A number of variables contribute to the variability of gene expression measurements, such as the number and type of cells in the tissue evaluated, the method and efficiency of mRNA extraction, mRNA handling techniques (223), mRNA integrity (224,225), method of reverse transcription (226) and analytical detection chemistry method (223). These inter-sample differences are addressed through the process of normalisation (235), whereby the expression of an individual gene within a sample is related to that of a calibrating gene known as a reference, control or "housekeeping" gene. Ideally, a reference gene is expressed at a consistent and repeatable quantity across all samples (normal and diseased) being compared, so that relative differences in gene expression can be measured with confidence. Commonly used reference genes such as *B2M, GAPDH* and *ACTB*, are not constantly expressed across all tissue types and disease states (236,238). Thus it is widely accepted that the selection of reference genes should be established through the validation of expression stability in the tissue or cells of interest, before use.

A number of statistical algorithms exist for the optimisation of reference gene selection, such as geNorm (238), Global Pattern Recognition (239), Bestkeeper (240), equivalence tests (242) and NormFinder (241). In each case, mathematical evaluation of expression data allows the ordering of candidate reference genes, on the basis of their relative expression stabilities. At present, no gold standard method exists for the selection of reference genes, and although methods have been compared with similar

results in some reports (245,366,367) but not in others (241), the optimal method for selections remains unknown.

New reference genes identified from microarray data, within a particular tissue type, have been demonstrated to be more "stable" than conventionally used reference genes, when compared using stability algorithms (241,244-246). Microarray data can be stratified on the basis of fold changes in expression (245), the variance of expression (241,246) or integrative correlations (244). Candidate genes can then be selected from stratified data, and frequently demonstrate expression stabilities greater than conventionally used reference genes (241,244,245). However, microarray data has yet to identify a new reference gene which shows consistent stability across multiple tissue or cell types, and / or disease situations. Therefore, a ubiquitous reference gene suitable for normalisation of gene expression of all experiments probably does not exist, but the identification of new reference genes to improve in reference gene stability is important to reduce error in RT-qPCR experiments and needs to be considered on a tissue by tissue basis.

In this study, we identified candidate reference genes from microarray expression profiling data generated from the evaluation of two different canine articular tissues (cartilage and cranial cruciate ligament). The relative stability of expression of each reference gene in normal and osteoarthritic canine articular tissues was determined from RT-qPCR reactions using statistical algorithmic packages. The stability of the new reference genes were compared between tissues, and related to a commonly used reference gene (Glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*]).

MATERIALS AND METHODS

Microarray data

Expression profiling data from 10 hip articular cartilage samples (5 control, 5 from osteoarthritic [OA] joints) and 16 cranial cruciate ligament (CCL) samples (4 normal CCL from a breed at low-risk of rupture, 7 normal CCL for a breed at high-risk of rupture, and 5 ruptured CCL from OA joints) were generated from a custom designed 44000 transcript canine whole genome 60mer oligonucleotide microarray (298). Raw data was normalised by two methods; locally weighted scatter plot smoothing (LOWESS), or using the geometric mean of 3 conventional reference genes arbitrarily selected (glyceraldehyde-3-phosphate dehydrogenase [*GAPDH*], ribosomal protein L13a [*RPL13A*], succinate dehydrogenase flavoprotein subunit A [*SDHA*]). Expression quantification was exported into an Excel Datasheet (Microsoft Excel 2003), and the data compared in three separate experiments as follows;

1. Normal hip articular cartilage was compared to OA cartilage,
2. Normal CCL (high-risk of rupture) was compared to normal CCL (low-risk of rupture), and
3) Normal CCL (high risk of rupture) was compared to ruptured CCL .

Selection of reference gene candidates

The stepwise procedure for identifying candidate reference genes is illustrated in Figure 1. Data for each reference gene candidate was compared in each experiment by calculating the fold change in mean expression level (between the two comparison groups), student's t-tests and percentage standard deviation (co-efficient of variation).

To identify the most stably expressed genes across each of the experiments, the prospective reference genes were then selected using the following criteria;

1. Student's t-test *P* value > 0.5 (in all experiments).
2. Ratio of expression between the two groups compared in each experiment <1.5 (in all experiments).
3. Standard deviation of the mean expression in each experimental group being <30% (in all experiments).

The data sets were reduced to 420 transcripts (LOWESS normalised) and 13 transcripts (reference gene normalisation). To further refine and filter the new reference gene list, data was ordered on the average signal intensity and the most abundantly expressed transcripts evaluated first.

The probe sequences used from the microarray experiments were entered into the National Centre for Biotechnology Information Basic Local Alignment Search Tool (278) to confirm the gene identity. Gene function was determined (278) and the associated gene information checked to ensure no known involvement in OA.

Complete filtering reduced the data set to 12 genes, of which 10 were selected from the LOWESS normalised data, (CG14980-PB [*C7orf28B*], Gu binding protein [*PIAS1*], HIRA interacting protein 5 [*HIRP5*], 5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/ IMP cyclohydrolase [*ATIC*], Mitogen-activated protein kinase 6 [*MAPK6*], Mitochondrial 28S ribosomal protein S25 [*MRPS25*], ORM1-like 2 [*ORMDL2*], Phosphatidylserine synthase 1 [*PTDSS1*], Mitochondrial

ribosomal protein S7 [*MRPS7*] and Transketolase [*TKT*]), 2 were selected from the reference gene normalised data (Hematopoietic stem/ progenitor cells 176 [*TRAPPC2L*] and Cytoplasmic protein NCK2 [*NCK2*]), and one gene was selected on the basis of its common usage in RT-qPCR experiments (*GAPDH*). The sequence details and putative functions (determined by reference to the human transcript at (278) are listed in Table 1.

Sample collection and storage

A separate set of samples were collected for the analysis of the new putative reference gene panel. Infrapatella fat (n = 5), ruptured cranial cruciate ligament (n = 5), femoral head articular cartilage (n = 5), ulnar subchondral bone (n = 5) and synovial membrane (n = 5) were obtained from dogs with clinical OA secondary to naturally occurring joint disease. In each case, the samples were obtained as part of the standard surgical treatment for the disease in question (total hip replacement, cranial cruciate ligament rupture surgery or fragmented coronoid process removal). Control samples (healthy) were obtained from infrapatella fat pad (n = 5), cranial cruciate ligament (n = 5), synovial membrane (n = 5), hip articular cartilage (n = 5) and ulnar bone (n = 5) of dogs euthanized for reasons other than, and with no evidence of, joint disease. Following the collection of the tissue, the samples were weighed and immediately stored in RNAlater (Qiagen Inc, Crawley, UK), according to the manufacturer's instructions, until extraction.

RNA extraction

For all of the tissue samples total RNA was extracted using a phenol/guanidine hydrochloride reagent (Trizol, Invitrogen Ltd, UK) with a chloroform extraction and

ethanol precipitation, as previously described (276). An on column DNA digestion step was included (RNase-Free DNase Set, Qiagen Ltd, Crawley, UK). Final elution of the total RNA was performed using 30 µl of RNase free water, and repeated to maximize the amount of RNA eluted. Total RNA samples were stored at -80 °C until use. The concentration of total RNA in each sample was quantified by using a NanoDrop® ND – 1000 UV/Visible Spectrophotometer (NanoDrop Technologies Ltd, Utah, USA).

cDNA synthesis

Reverse transcription was performed using Superscript II reverse transcriptase (Invitrogen, Dorset, UK) according to the manufacturer's instructions (331). Initially 200 µg (10 µl) total RNA was pre-incubated with 0.5 µg (1 µl) oligo-dT12-18 (Invitrogen, Paisley, UK) and 10 mM (1 µl) dNTP mix (Invitrogen, Paisley, UK) at 65°C for 5 minutes. After chilling on ice, 4 µl of 5x first strand buffer (containing 250 mM Tris-HCI (pH 8.3), 375 mM KC1, 15 mM MgCl2), 2 µl of 0.1M DTT and 40 units (1 µl) of RNase (Promega, Southampton, UK) were added to each sample and the samples incubated for 2 minutes at 42 °C, followed by the addition of 200 units (1 µl) of Superscript II reverse transcriptase (Invitrogen, Dorset, UK) and incubated for 50 minutes. Reverse transcriptase activity was terminated by incubation at 70 °C for 15 minutes, and samples stored at -80 °C until use.

RT-qPCR assay design

Transcript sequences were obtained from the National Centre for Biotechnology Information (278) and were cross referenced to the Ensembl canine genome database (277). Primer and probe sequences were then designed for each of the reference genes

by using the Universal Probe Library Assay Design Centre (279). BLAST searches were performed for all primer sequences to confirm gene specificity, and electrophoresis of the PCR reaction mixture confirmed a single product of the appropriate length in all cases. Primers were synthesized by Metabion International AG (Martinsried, Germany), and probes were synthesized by Roche Diagnostics (Lewes, U.K) using locked nucleic acid with 5'-end reporter dye fluorescein (FAM [6-carboxy fluorescein]) and 3'-end dark quencher dye.

Real-time -qPCR assays were performed in triplicate using the LightCycler 480 (Roche Diagnostics, Lewes, U.K) in 384 well format, with three no template controls used for each assay. The reaction volume in each well consisted of 5 µl LightCycler 480 Probes Master 2x concentration (Roche Diagnostics) (containing FastStart Taq DNA Polymerase, reaction buffer, dNTP mix (with dUTP) and 6.4 mM $MgCl_2$), 0.7 µl of LightCycler 480 Probes Master H2O (Roche Diagnostics), 0.1 µl of 20 µM forward primer, 0.1 µl of 20 µM reverse primer, 0.1 µl of 10 µM fluorescein-labelled probe and either 4 µl of sample cDNA, diluted template, or 4 µl of LightCycler 480 Probes Master H2O. The standard amplification conditions consisted of 1 cycle at 95°C for 5 minutes, followed by 50 cycles of 95°C for 15 seconds and 60°C for 30 seconds. Real-time qPCR data was then analysed by using LightCycler 480 Basic Software (Roche Diagnostics). Standard curves were generated for each reference gene by employing cDNA or template oligonucleotides (332), the parameters of which are listed in Table 1. All samples were checked for absence of genomic DNA contamination using a canine genome specific RT-qPCR assay, previously described (296). The assays were deemed to be reproducible, as determined by the average

standard deviation of the triplicate repeats of each assay being less than 30% (Table 1).

Reference gene stability analysis

Real-time RT-qPCR data was exported into an Excel datasheet (Microsoft Excel 2003) and analysed using three separate reference gene stability analysis software packages; geNorm (238), Bestkeeper (368) and NormFinder (241). Each of these methods generates a measure of reference gene stability, which can be used to rank the reference genes in order of stability. GeNorm generates a stability measure (the M value) for each gene which is arbitrarily suggested to be lower than 0.4 (with a lower value indicating increased gene stability across samples), and a pairwise stability measure to determine the benefit of adding extra reference genes for the normalisation process, with again a lower value indicating greater stability of the normalised genes, and a lower value indicating greater stability with an arbitrary cut off value of lower than 0.15 indicating acceptable stability of the reference gene combination (238). NormFinder generates a stability measure of which a lower value indicates increased stability in gene expression. By using a model-based approach, NormFinder groups samples to allow for a direct estimation of expression variation, compared to the pairwise comparison approach that ranks genes according to the similarity of their expression profiles. Therefore, taking a sample set which consists of two sample subgroups where all of the candidates but one show little difference between the groups, the one candidate which shows no difference will have the smallest stability value across all candidates and be the most stably expressed gene. Bestkeeper generates a pairwise correlation co-efficient between each gene and the Bestkeeper index (the geometric mean of the threshold cycle values of all the reference genes grouped together). Stability measures for combined (normal and diseased) samples

were recorded, as ultimately it is these measures which would be used to determine which genes were suitable for normalising expression data from genes of interest in a particular disease (OA in this case).

BestKeeper can only be used to analyse a maximum of 10 housekeeping genes so the three genes least stably expressed (as determined by NormFinder) were excluded from BestKeeper analysis. The stability values for each gene, as determined by each method of analysis, are illustrated in Figure 2, 3, and 4. Statistical tests were performed using a statistical software package (Minitab V14.1, Minitab Ltd., Coventry, U.K). Spearman rank correlation coefficients were then calculated using the ranking order of genes to compare the relationship of the relative ordering of genes by different methods of analysis (Table 2). Finally, the stability parameters of the new reference genes were compared to those generated for commonly used reference genes in a similar study of canine OA tissues (369) (Table 3).

RESULTS

New reference genes

Identities and putative functions of each of the new potential reference genes are listed in Table 1. Although the genes selected did not localise to common pathways or functions, two of the genes coded for mitochondrial ribosomal proteins. The metrics of the candidate reference gene stability are presented in Table 2. These were the used as reference gene candidates in RT-qPCR on specific articular tissues.

Articular cartilage

All methods of stability analysis agreed in identifying the (new reference) genes *MRPS7* and *MRPS25* as stably expressed in cartilage samples. Likewise, *C7orf28B* and *NCK2* were determined to be the least stably expressed genes by both geNorm (Figure 3) and NormFinder (Figure 2). *GAPDH* was identified as the 4th most stably expressed gene by both geNorm and Bestkeeper, and the 8th most stably expressed gene by NormFinder.

Infrapatella fat pad

All three methods of reference gene analysis agreed on the most stably expressed reference genes in the fat pad, which were *C7orf28B, MRPS7* and *MAPK6*. GeNorm (Figure 3) and NormFinder (Figure 2) agreed that the least stably expressed gene was *NCK2*. *GAPDH* was identified as the 9th most stably expressed gene by NormFinder, the 7th most stably expressed gene by geNorm, and the 5th most stably expressed gene by Bestkeeper.

Cranial cruciate ligament

Methods did not agree on the most stably expressed genes in the CCL, although all methods agreed on the five most stably expressed genes (albeit, not their order); *ATIC, MRPS7, C7orf28B, ORMDL2* and *HIRP5*. *MRPS25* was the least stably expressed gene as determined by both NormFinder (Figure 2) and geNorm (Figure 3). *GAPDH* was identified as the 7th most stably expressed gene by NormFinder, the 9th most stably expressed gene by geNorm, and the 8th most stably expressed gene by Bestkeeper.

Synovial membrane

Although Bestkeeper and NormFinder agreed on the six most stably expressed genes in synovial membrane (*MRPS25, ATIC, HIRP5, TKT, MRPS7, PTDSS1*), and *NCK2* was determined to be the least stably expressed gene by NormFinder (Figure 2) and geNorm (Figure 3), no further patterns of agreement in rank ordering of the expression profiles were identified. *ATIC* was identified as the most stably expressed gene by NormFinder (Figure 2) and Bestkeeper (Figure 4), and the 6th most stably expressed gene by geNorm.

Bone

Rank ordering between NormFinder and geNorm agreed on the seven most stably expressed (*C7orf28B, MRPS25, PIAS1, PTDSS1, ATIC, MRPS7* and *HIRP5*) bone genes but not their order, and the least stably expressed gene (*NCK2*). Bestkeeper (Figure 4) and NormFinder (Figure 2) agreed on the most stably expressed gene (*C7orf28B*).

Comparison of reference gene performance in all tissues

Using the reference gene stability value (M) of 0.40 as the determinant of stable expression (238), *MRPS7* was stably expressed in all five tissues, and *HIRP5* was found to be stably expressed in four tissues (Figure 3). *GAPDH* was found to be unstable in all of the tissues evaluated, which is consistent with the findings of a previous study of reference genes in these tissues (370). Comparison of gene stability (M) and pairwise stability (V) values with a previous study of commonly used reference genes using similar tissues further illustrates how optimal reference gene stabilities, can be achieved using the new reference genes rather than the commonly used reference genes (Table 3).

No single reference gene was consistently identified as being the most stably expressed by NormFinder, geNorm or BestKeeper across all tissues. There was not consistent agreement in the rank ordering, or the selection of the optimal candidates by the different analysis methods, although agreement was generally reached on the most and least stable gene. For example, BestKeeper and NormFinder always identified the same gene as being most stably expressed. When looking at rank order across all three reference gene stability programs, fat pad showed the highest correlation between methods, followed by cruciate ligament, cartilage, bone and synovium as the least consistent (Table 2).

When the data for all tissues was compared together (Figure 2, 3, 4), a much clearer pattern of reference gene stability was observed. The stability metrics of the reference genes in different tissues show similar patterns across all three methods. *MRPS7* demonstrates the most consistent metric (low geNorm M value, low NormFinder value and high Bestkeeper correlation), with *HIRP5* and *ATIC* demonstrating a similarly consistent stability across all tissues. This is supported by the finding that *MRPS7* was consistently identified as being stably expressed in all tissues by geNorm, as well as being ranked as one of the two most stable reference genes in four of the five tissues by geNorm (cartilage, fat, bone and synovium), and in three of the five tissues using NormFinder and BestKeeper (cartilage, ligament and fat).

Comparison of genes identified by different methods

Identification of new reference genes using RT-qPCR methodology for gene normalisation was not successful at identifying new reference genes with increased

stability when compared to a commonly used reference gene such as *GAPDH*. *NCK2* was determined to be the least stably expressed gene in synovium and fat pad, and one of the four least stably expressed genes in cruciate ligament and cartilage. *TRAPPC2L* was not identified as being stably expressed in any tissue using the geNorm algorithm, and was not ranked higher than the 8th most stably expressed gene in any tissue using the NormFinder algorithm.

DISCUSSION

A number of different strategies have been employed to filter microarray data to identify new reference genes, such as selection on the co-efficient variation and level of expression (241), fold changes of expression (245,246), or integrative correlations (244). We used a combination of filtering on statistical significance, fold change and coefficient of variation (percentage standard deviation) to narrow the potential number of reference genes. Furthermore, these criteria were applied to three different experiments, using two different data sets, to identify genes which were more likely to have generic stability across multiple tissues for diseases. Genes were finally filtered on the basis of defined annotation and level of expression. In retrospect, genes should also have been selected on the basis of single transcript expression (i.e. the absence of splice variants). Although the two most stably expressed genes (*MRPS7* and *HIRP5*) currently have no splice variants reported, the absence of splice variants did not necessarily confer reference gene stability across multiple tissues (as demonstrated by *GAPDH* and *C7orf28B*, genes which do not have splice variants annotated but which were not stably expressed) but should be taken into account when selecting new reference genes, as another potential indicator of instability. Our filtering method was

straightforward, quickly performed and easily completed by any person without a full understanding of microarray data set handling, and as such could be applied to publicly available microarray data sets for a given experiment or disease.

Variability in the expression of commonly used reference genes has been recognised by the analysis of both cell culture experiments (371) and clinical tissue specimens (372). The selection of reference genes upon their stability as determined by the mathematical assessment of their expression values in a test cohort of samples is a widely accepted technique (238,243-245,366,367,373). We identified one gene which showed stable expression across normal and diseased articular tissues (*MRPS7*), and a number of genes which demonstrated a relatively consistent stability across the majority of tissue specimens (*HIRP5*). One should bear in mind that the tissues evaluated were from the same embryological origin (mesenchymal tissue), and hence there may have been a tendency towards identifying a reference gene which was stable in all tissues, although this is not supported by previous reports of reference gene stability in different tissues (243). Likewise, the diseases compared in the microarray data sets were the same as those affecting the tissue samples evaluated by real-time RT-qPCR, which may further tend towards identifying reference genes whose stability was constant. Therefore, although we identified one gene as being stably expressed in all tissues, we would not advocate its use as a reference gene in other tissues or diseases without assessment of its stability in the samples under analysis (238,243,246). The utopia of a universal reference gene suitable for all studies probably does not exist on basis of the published evidence to date.

Mitochondrial ribosomal protein S7 is involved in mitochondrial protein synthesis. The precise function of this gene is unknown in eukaryotes, but the protein is thought to be involved in organising the 3' domain of the 16 S rRNA in the mitochondria of prokaryotes, and thus be involved in the initiation of translation in mammalian mitochondria (374). Microarray data analysis indicated the *MRPS25* was also stably expressed, although it was only stably expressed in two of the four tissues analysed by RT-qPCR (cartilage and fat pad). In a separate study, mitochondrial ribosomal protein L19 was one of six genes identified from microarray data obtained from different tissues and cells, as a good reference gene for real-time RT-qPCR experiments, when compared to conventional reference genes for mammary tumour expression profiling (246). Mitochondrial ribosomal gene expression appears to show greater stability across different tissues and thus may be better potential candidate reference genes for other real-time RT-qPCR experiments.

Comparing the results of this study to a similar previous study of commonly used reference genes in multiple articular tissues demonstrates the increased stability of the "new" reference genes (Table 3) (370). The selection of candidate reference genes from microarray data identified new genes which were more stably expressed and is consistent with the general outcome of previous studies using this methodology (241,244-246). The normalisation of microarray data by geometric mean of three reference genes (238) did not identify genes (*NCK2* or *TRAPPC2L*) with appropriate stability to be suitable for use as reference genes. The instability of these genes may be reflected, in part, by the greater variation identified in the triplicate repeats of each assay when compared to more genes determined as being more stably expressed such as *HIRP5* or *MRPS7*. The less stable expression of the three conventional reference

genes (*GAPDH, RPL13A* and *SDHA*) probably resulted in the selection of similarly "unstably" expressed reference genes from microarray data, and thus accounted for this being a futile method of trying to select reference genes, which agrees with the evaluation of these types of methodologies for the accurate normalisation of microarray data (375). These genes were selected on the basis of a preliminary study of reference gene stability in canine OA tissues (369), however subsequent work evaluating greater sample numbers has determined that one of these genes (*SDHA*) demonstrates differential expression in OA cartilage (296) and thus its use may have further predisposed to the selection of genes which were not stably expressed. Furthermore, the conventionally used reference gene we evaluated (*GAPDH*) did not show acceptably stable expression in any of the tissues we analysed.

Previous studies have demonstrated that the generation of rank orders can be very similar between different methods of reference gene selection (245), but this is not always the case (241). The best correlation in rank ordering was observed between geNorm and BestKeeper, across all the tissues which is unsurprising as both are generated by pairwise comparisons (although geNorm uses un-transformed data, whereas BestKeeper uses threshold cycle values, which are a *log* relation to the true transcript number), although BestKeeper and NormFinder always identified the same gene as being most stably expressed. The rank order of reference gene stability was identified most consistently for fat pad, followed by cruciate ligament, cartilage, bone and least consistently for synovium.

The advantage of using a model based stability assessment is that rank ordering can be changed if co-regulated genes are included in the stability assessment procedure, as

pairwise assessments will determine an increase in stability between these methods (241). As we identified a number of new reference genes which have very little functional information associated with their annotation, we checked for co-regulation between the most stably expressed genes by removing one of the highest ranked genes (as determined by pairwise comparisons) alternately, and re-assessing the rank ordering of reference genes stabilities. No major changes in rank ordering or reference gene stability were observed when this was performed. However, it should be noted that other factors besides gene expression pathway similarities can contribute to co-regulation. For example transcription factors may target multiple genes resulting in complex relationships across the apparently un-related genes (376).

CONCLUSIONS

The use of microarray data for the selection of reference genes allowed the identification of multiple genes demonstrating greater stability than a conventional reference gene in multiple tissues. Mitochondrial ribosomal protein S7 is suitable for use as a reference gene in all the experimental conditions we analysed. Different methods of assessment of gene stability do not always show correlation between the rank order of gene expression stability, but they do generally agree on which genes are suitable for use to normalise gene expression experiments.

Authors Contributions

DNC and LM carried out the microarray data analysis. LM and FS carried out the assay design. DNC, LM and FS performed the molecular genetic studies and DNC performed the statistical analysis. DNC and PJRD conceived the study, its design and coordination, and drafted the manuscript with LM.

Figure 1

Microarray data normalised by two different methods was filtered to identify new reference genes using statistical significance, fold changes in expression between experimental group and the co-efficient of variation.

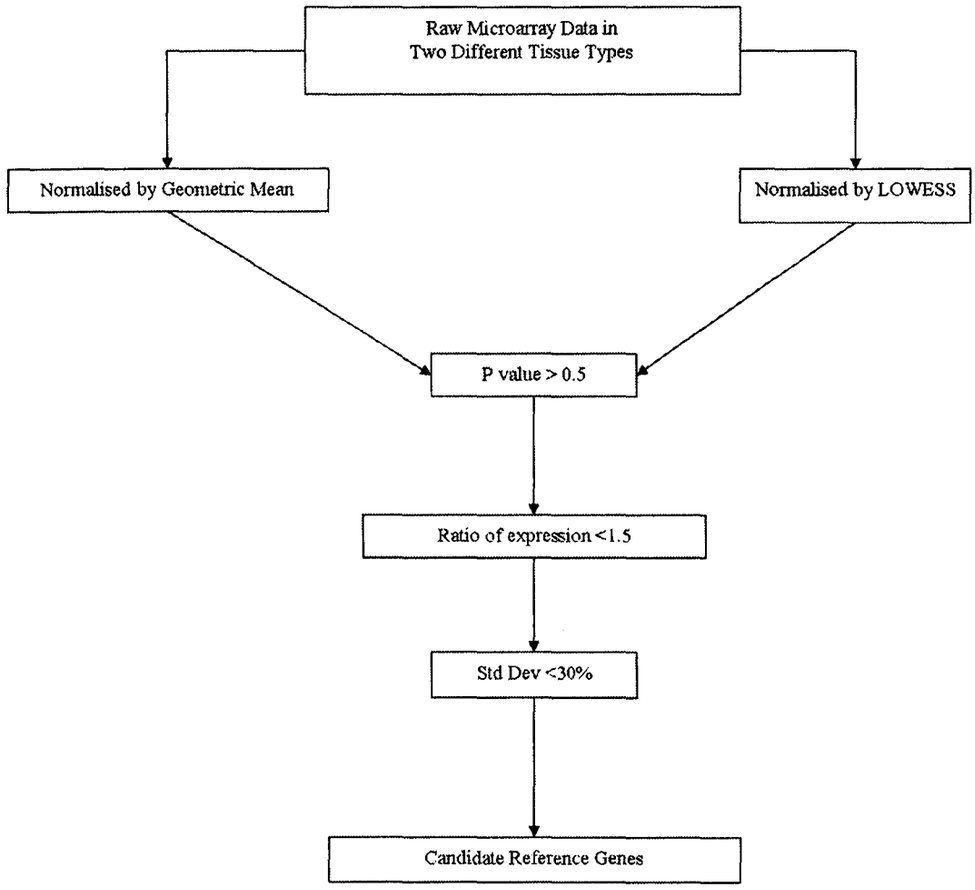


Figure 2

Reference gene stability measures as determined by the NormFinder Algorithm (with a lower value indicating increased reference gene stability).

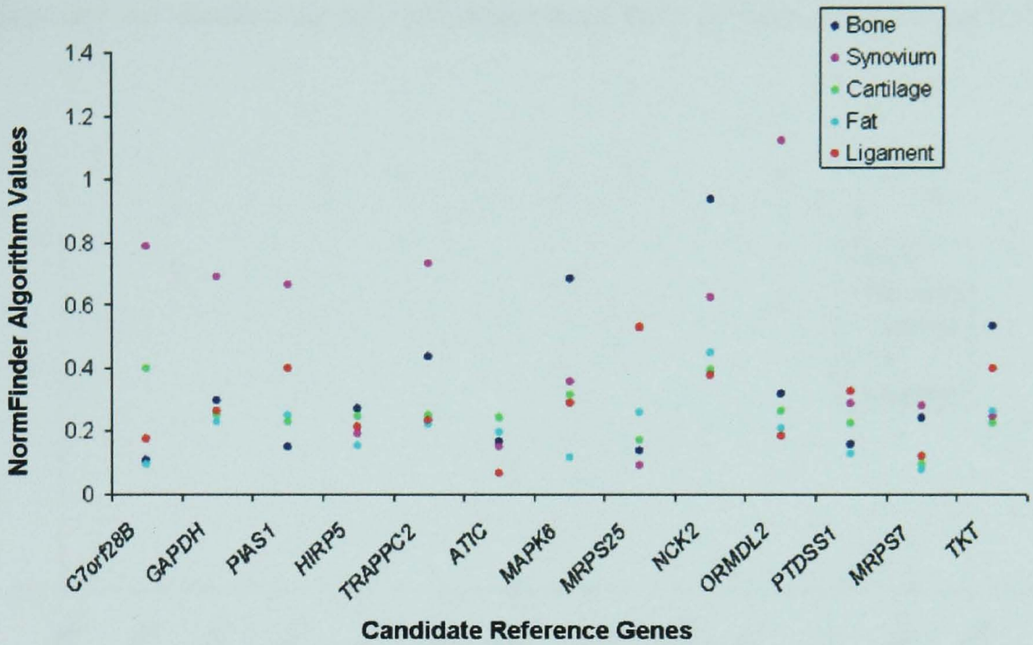


Figure 3

Reference gene stability measures as determined by the geNorm algorithm (with a stability measure [M value] <0.4 indicating appropriate reference gene stability).

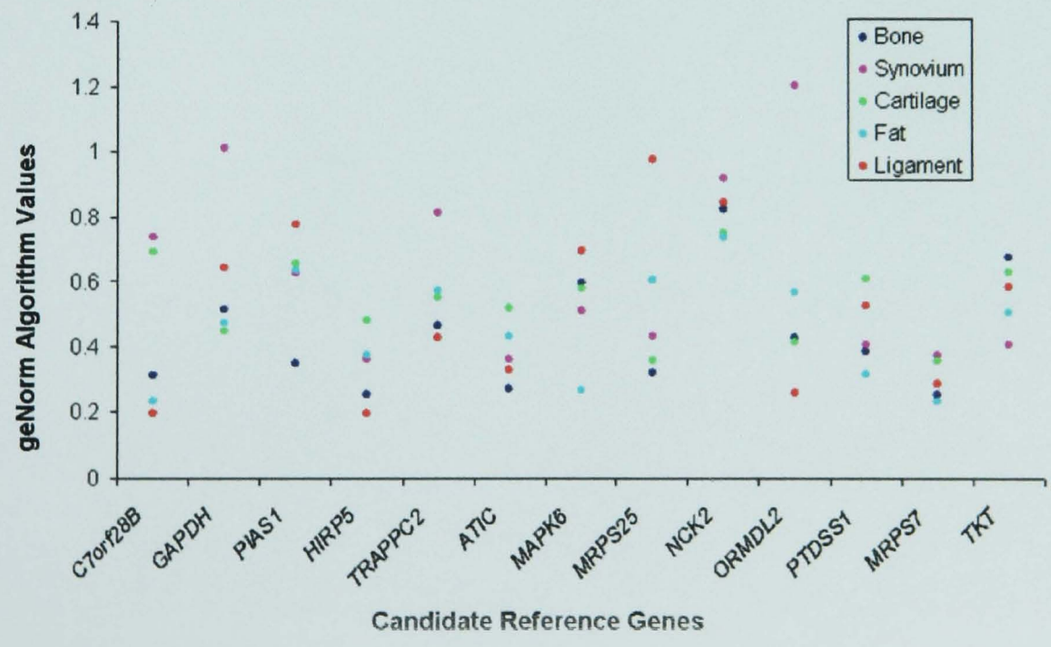


Figure 4

Reference gene stability measures as determined by the Bestkeeper algorithm (with a higher value indicating increased reference gene stability). Please note that as only the top 10 genes (as ranked by the NormFinder algorithm) are selected for analysis, thus there are not necessarily data points for each gene corresponding to each tissue.

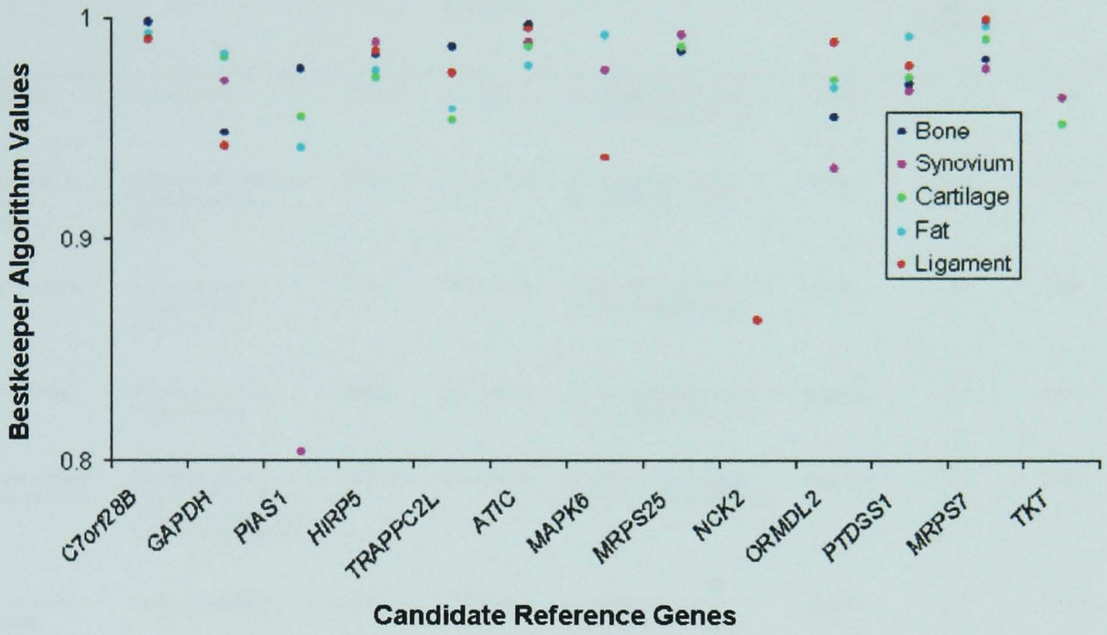


Table 1

A list of the gene annotations, functions, primer and probe sequences, and qPCR metrics for the 12 new reference genes, and glyceraldehyde 3-phosphate dehydrogenase.

Gene Name	*Gene Function*	*Gene Symbol*	*Accession Number [GenBank]*	*Forward (F) and Reverse (R) Primers*	*Probe Sequence*	*Average Standard Deviation of Triplicate*	R^2	*PCR Efficiency*
CG14980-PB	Protein coding	*C7orf28B*	XM_536878	F - gcaggaagggattctccag R - ggtccagtaagaaatcttccataa	gccaggaa	19.8	0.986	104.3
Glyceraldehyde-3-phosphate dehydrogenase	Enzyme in the glycolysis / gluconeogenesis pathway	*GAPDH*	NM_001003142	F - ctggggctcacttgaaagg R - caaacatgggggcatcag	ctgctcct	20.3	0.991	101.1
Gu binding protein	Nuclear receptor in transcriptional co-regulation	*PIAS1*	XM_535524	F- ggagacaatcagcattataacacct R - tgatcatctgacactgctgct	ggctgctg	16.9	0.990	99.6
HIRA interacting protein 5	Histone-interaction-DNA packaging	*HIRP5*	XM_850340	F - aattcagaacatgctgcaatttta R - tgattcatcatccataacctgttc	aggtggag	8.6	0.998	96.9
Hematopoietic stem/ progenitor cells 176	Transport protein particle involved in endoplasmic reticulum to Golgi vesicle transport	*TRAPPC2L*	XM_844929	F - gatgatccaggtgtgctgag R - caatacggttatgtcaacagcact	ctggagga	25.2	0.993	97.2
5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/ IMP cyclohydrolase	Purine biosynthesis	*ATIC*	XM_858011	F - cgctgcctctttcaaacat R - tttggcctcatcttcactgag	cagcaggt	13.4	0.991	97.7
Mitogen-activated protein kinase 6	Phosphorylates microtubule-associated protein 2 (MAP2)	*MAPK6*	XM_858091	F - tcttcttgggatagccagtttg R - cctcacctcacaacaaaactgat	ggtggtgg	14.9	0.992	97.6
Mitochondrial 28S ribosomal protein S25	Mitochondrial ribosomal subunit protein synthesis	*MRPS25*	XM_533729	F - tgaaggtcatgacggtgaac R - tggatctgaggtatgttgaaaaac	gccaggaa	14.6	1.000	95.5
Cytoplasmic protein NCK2	Regulates cell proliferation	*NCK2*	XM_538440	F - cagacgctctacccgttca R - gtctcgcccttctcgaagtt	aggaggag	28.7	0.975	96.7
ORM1-like 2	Protein folding in the endoplasmic reticulum	*ORMDL2*	XM_843143	F - atggactacgggctccaat R - ctggccaggaggtagagtaca	ctcctccc	28.2	0.996	103.1
Phosphatidylserine synthase I	Membrane bound protein that catalyses the replacement of phospholipids by L-serine	*PTDSS1*	XM_849686	F - actcagaatgcgacgatgg R - tcagaacctttgaacctttcg	ctggtctc	15.3	0.996	100.9
Mitochondrial ribosomal protein S7	Mitochondrial protein synthesis	*MRPS7*	XM_846915	F - agtgcagggagaagaagcac R - cagcagctcgtgtgacaact	ggatgctg	12.1	0.998	100.8
Transketolase	Enzyme in pentose phosphate pathway	*TKT*	XM_533792	F - caacttctgtggctcccact R - ccagatcttccagagccatc	tggggaag	11.8	0.993	103.4

Table 2

Correlation coefficients for the rank ordering of gene stability by different reference gene analysis methods.

Tissue	*Method*	*NormFinder*	*GeNorm*
Cartilage	GeNorm	0.462	
	BestKeeper	0.515	0.721
Cruciate Ligament	GeNorm	0.835	
	BestKeeper	0.915	0.794
Synovium	GeNorm	0.833	
	BestKeeper	0.745	0.579
Fat Pad	GeNorm	0.907	
	BestKeeper	0.867	0.939
Bone	GeNorm	0.710	
	BestKeeper	0.382	0.475

Table 3

Comparison of M and V values generated in this study when compared to a previous study (Ayers and others 2007) evaluating similar tissues.

	Current Study			*Ayers (2007) Study (369)*		
Tissue	*Reference Genes*	*M (Gene Stability) Value*	*V (Pairwise Stability) Value*	*Reference Genes*	*M (Gene Stability) Value*	*V (Pairwise Stability) Value*
Articular Cartilage	***MRPS7***	0.37	0.122	***RPL13A***	0.57	0.31
	MRPS25			***SDHA***		
Synovium	***MRPS7***	0.2	0.091	*N/A*	N/A	N/A
	ATIC					
Cruciate Ligament	***HIRP5***	0.2	0.093	***B2M***	0.59	0.27
	C7orf28B			***TBP***		
Fad Pad	***C7orf28B***	0.23	0.088	***B2M***	1.02	0.35
	MRPS7			***SDHA***		
Bone	***MRPS7***	0.36	0.084	*N/A*	N/A	N/A
	HIRP5					

Chapter 7

An assessment of the quality and quantity of genomic DNA recovered from canine blood samples by three different extraction methods

Dylan N. Clements[1,2]*,

Shona Wood[1,2],

Stuart D. Carter[1],

William E.R. Ollier[1,2].

[1]Musculoskeletal Research Group, c/o Department of Veterinary Pathology, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX.

[2]Centre for Integrated Genomic Medical Research, The Stopford Building, University of Manchester, Oxford Rd, Manchester, M13 9PT.

Paper published in *Research in Veterinary Science* (2007) E-publication ahead of print

ABSTRACT

Introduction

The ideal method for genomic DNA (gDNA) extraction from whole blood should recover high quantities of pure, integral gDNA from the original sample source with minimal co-extraction of inhibitors of downstream processes. Three different methods of gDNA extraction from canine whole blood were compared.

Materials and Methods

Genomic DNA was extracted from ethylenediamine tetra acetic acid (EDTA) treated and clotted blood samples by three different methods (a silica column method, a phenol-chloroform method and a modified salt precipitation method). The quantity of gDNA recovered was compared by spectrophotometric measurement and the quantative polymerase chain reaction (qPCR). The quality of the gDNA recovered was compared by quantification of PCR inhibition, spectrophotometric measurement and agarose gel electrophoresis.

Results

Phenol-chloroform and modified salt precipitation based extractions demonstrated similar relative recovery of gDNA with EDTA preserved blood, but were less efficient at recovering gDNA from clotted blood. Spectrophotometer measurement of phenol-chloroform based extractions tended to overestimate the quantity of gDNA recovered from extractions, and was associated with the greater co-extraction of PCR inhibitors.

Conclusions

The silica column method recovered gDNA with equal efficiency, purity and integrity irrespective of the sample type or method of quantification. Spectrophotometric measurement of DNA quality may not reflect the true utility of the gDNA present in a sample.

INTRODUCTION

The publication of the canine genome (218) has provided great opportunities for canine genomic research. For studies investigating the genomic basis of canine traits (377-379), the isolation of high quality genomic DNA (gDNA) from dog tissues or fluids is a prerequisite. At present the major restricting factor on canine genetic studies is the availability of clinical material with linked high quality phenotypic information. This genotyping data can be provided by many potential sources of canine gDNA, such as blood (380), hair (381), toe nails (382), tissue (383), urine (384) or buccal swabs (385). Of these sources, blood is one of the most readily available and contains large quantities of gDNA which are not likely to be contaminated with bacterial DNA extracted at the same time.

Haematology and biochemistry profiles are routinely performed in veterinary practice, and excess blood remaining, after the tests have been completed, can provide a simple and ethical potential source of gDNA for use in genetic studies (386), providing owners have been given informed consent for the collection, extraction and use of the sample. Collection of residual blood is particularly appealing as the blood sample is obtained for the sole purpose of clinical investigation yet provides a source of gDNA without the necessity for repeated venepuncture or additional volume collection, neither of which are allowed in the UK without Home Office approval. Secondly, phenotypic information about the patient, such as pertinent clinical findings and the results of subsequent diagnostic tests can easily be linked to an individual sample. The collection of such samples for archival storage is now commonly performed, both in

the United Kingdom (387) and the United States of America (388), and promise to provide the large sample cohorts required for progress in canine genomic research.

Traditionally, gDNA was extracted from whole blood using organic solvents, such as phenol and chloroform, followed by precipitation with ethanol. New techniques utilise alternative methodologies to increase the speed of extraction and reduce the exposure to hazardous chemicals. For example, DNA may be adsorbed onto silica gel membranes (389) or magnetic beads (390), or protein and DNA can be separated by sequential precipitation using a modified salt precipitation protocol (391). Additional factors also need to be taken into account when selecting a method of gDNA extraction such as the sample size, sample type, quantity of gDNA recovered (389), the purity of gDNA recovered (228), the integrity of gDNA recovered (392), the co-extraction of inhibitors of PCR (393), as well as cost and speed of extraction. In this report we evaluate three different methods of canine gDNA extraction on the quantity and quality of gDNA recovered using spectrophotometer absorbance values, ethidium bromide stained agar gel electrophoresis scans and quantitative polymerase chain reaction (qPCR) data.

MATERIAL AND METHODS

Clinical Material

Blood was collected from ten dogs presenting to the University of Liverpool Small Animal Teaching Hospital for diagnostic investigation which included routine haematological and biochemical tests. The inclusion criteria were cases for which over 3 ml of both clotted and ethylenediamine tetra acetic acid (EDTA) preserved

blood remained following that used for the diagnostic investigation. Written informed consent was given by all owners, facilitating the use of residual blood (which would normally be disposed of) for research purposes. Samples collected into EDTA were each divided into three aliquots of one millilitre, and clotted samples were divided into three roughly equal proportions, placed into plain collecting tubes and weighed. One millilitre and one gram was selected as the unit size of sample to analyse in this experiment, as in the author's experience this was estimated to be the average submission of residual blood samples to the UK DNA Archive for companion animals (387) for gDNA extraction and storage.

A precise nucleated cell count of each sample was measured twice using a haemocytometer and the average cell count used in calculations (for a full description see the supplementary information at the end if this manuscript). The gDNA copy number present in each sample was determined using the cell count, the sample weight, and a correction factor for blood density (1.06 g/ml, Geigy Scientific Tables). All samples were stored at room temperature for 24 hours, followed by freezing at -20°C for at least 7 days, to mimic the typical conditions of transport and storage prior to extraction at the UK DNA Archive for Companion Animals.

Genomic DNA Extraction

Prior to extraction, clotted samples were macerated with the end of a 3 ml Plastic Pasteur pipette.

(i) Phenol-Chloroform Method

Genomic DNA was extracted by phenol-chloroform extraction and ethanol precipitation (for a full description see the supplementary information at the end of this manuscript). Genomic DNA was re-suspended in 300 µl of Tris-EDTA (10 mM Tris, pH 8.0; 1 mM EDTA, pH 8.0) and overnight rotation. The results for EDTA samples extracted by the phenol-chloroform method are denoted PE, and the results of the clotted samples extracted by the phenol-chloroform method are denoted PC.

(ii) Silica Gel Membrane Column Method

Silica gel membrane columns were used (QIAamp DNA Blood Midi Kit, Qiagen, Crawley, UK) following the manufacturer's instructions (389). The technique utilises a protease digestion, cell lysis and then binding of DNA to a silica gel, which allows serial washes to remove non-DNA cell contents followed by elution with a low salt buffer. Elution was performed with 300 µl elution buffer, run through the silica column twice (Qiagen, 10 mM Tris-Cl, 0.5 mM EDTA, pH 9.0). The results relating to EDTA samples extracted by silica column method are denoted QE, and for clotted samples extracted by the silica method are denoted QC.

(iii) Modified Salt Precipitation Method

In the third method, DNA was extracted with a modified salt precipitation technique (Puregene, Gentra, UK) following the manufacturer's instructions (391). The principle of the technique is that proteins are separated from DNA using a salt solution, and the DNA is isolated and purified further using an alcohol solution. Clots were digested with proteinase K (in addition to the red blood cell lysis solution), as recommended by the manufacturer. Genomic DNA pellet hydration was achieved with 300 µl Hydration Buffer (Gentra, 10 mM Tris-Cl, 1 mM EDTA, pH 7.0-8.0).

The results relating to EDTA samples extracted by the modified salt precipitation method are denoted GE, and for clotted samples extracted by silica columns are denoted GC.

DNA Quantification

All samples were quantified for DNA concentration using a spectrophotometer (NanoDrop ND-1000, Labtech International Ltd, East Sussex, UK), blanked using the appropriate hydration solution or elution buffer. The quantity of DNA recovered for each sample was normalised to allow comparison between methods by calculating the ng per 10^6 nucleated cells originally counted in each blood sample.

Accurate quantification of gDNA copy number was determined using a canine genome specific qPCR assay as previously reported (296) which amplifies a section of gDNA (location; chromosome 12: 4071650:4071709), with no known exon, intron or promoter position annotated. Samples were diluted to 20 ng/µl with TE for use in the assay, on the basis of the spectrophotometer quantity results. The qPCR assays were all performed in triplicate using a TaqMan™ ABI PRISM 7900 SDS (Applied Biosystems, California, USA) in 384-well plate format, as previously described (296). The precise number of gDNA transcripts in each sample was calculated with reference to a standard curve generated on the same plate, using serial dilutions of a synthetic template oligonucleotide (332) of the amplicon (Eurogentec, Southampton, UK) across the dynamic range of the quantities measured (5 pM, 500 fM, 50 fM and 5 fM). The transcript number calculated by the SDS software was then divided by the original nucleated cell count (assuming 2 copies per cell, and taking into account the

dilution factors) for each sample, and expressed as a percentage gDNA recovery (Figure 1).

Genomic DNA purity and integrity measurements

The purity of gDNA extracted was determined using the spectrophotometer, by measuring the 260 nm to 280 nm absorbance ratio (A_{260}:A_{280}). A gross assessment of gDNA integrity was generated by electrophoresis of 5 µl of each sample normalised to 20 ng/µl (were possible) on a 2% agarose gel containing 0.5 µg/ml ethidium bromide, with 2 µl of 1 kb ladder (Gel Pilot 1 kb Ladder, Qiagen, Crawley, UK) for 1 hour at 140V, followed by visualisation and photography under ultraviolet light (Gel Doc 1000, Bio Rad Laboratories Limited, Hemel Hempstead, UK). Visual assessment of the electrophoresis traces was performed independently by a single blinded observer (DC) experienced in analysing gDNA quality. Each sample was designated as high or low quality DNA. No descriptors were provided, although assessment was based on loss of a main band and smearing in the gel.

PCR Inhibition

The level of PCR inhibition on the normalised gDNA samples was quantified by using an internal amplification control (227). A synthetic oligonucleotide was synthesized to spike gDNA samples, thus allowing the quantification of gDNA inhibition. A genomic sequence from the varicella zoster virus (VZV) genome was utilised for the spike (Table 1), as canines are not infected with this virus. All gDNA samples were checked with un-spiked tests wells to confirm the absence of assay cross-reactivity. The qPCR inhibition assays were performed as previously described except each assay well had a 20 µl reaction volume consisting of 10 µl 2X PCR

master mix with Uracil N-Glycosylase, 0.2 µl each of 20 µM forward and reverse primers, 0.2 µl of 10 µM probe, 4.4 µl of 10 pM synthetic template oligonucleotide or water (negative controls) and 5 µl of sample (genomic DNA, 20 ng/µl) or TE (positive controls). The percentage inhibition (relative to the positive control) was calculated as follows =

$$100- [100 \times E \text{ value of VZV assay}^{(\text{Mean } C_T \text{ value of unknown assay} - C_T \text{ value of positive control assay})}]$$

Statistical Analysis

Statistical analyses were performed using statistical software (Minitab Version 14.1, Minitab Ltd, Coventry, UK). Comparison of the evaluation was performed using One-Way Analysis of Variance, with post hoc analysis using the Tukey test. Significance was set at $P < 0.05$.

RESULTS

Spectrophotometer quantification of gDNA

A graphical illustration of the quantity of gDNA recovered using different methods of extraction is presented in Figure 1. The greatest quantity of gDNA extracted from canine blood samples was obtained using the phenol-chloroform extraction method (PE [mean 13.4 ng/10^6 cells, standard error [SE] +/-2.4] and PC [14.2 ng/10^6 cells +/-2.5]), and the lowest quantity was for the modified salt precipitation method (GE [8.0 ng/10^6 cells +/-1.7] and GC [1.2 ng/10^6 cells +/-0.4]). The silica column method extracted similar quantities of gDNA from both EDTA preserved (QE [8.2 ng/10^6 cells +/-1.3]) and clotted blood (QC [8.1 ng/10^6 cells +/-0.8]). The quantity of gDNA extracted from clotted blood using the modified salt precipitation method (GC) was

significantly less than that extracted from clotted blood using the phenol-chloroform method (PC) or from EDTA preserved blood using the silica column method (QE) ($P < 0.001$).

PCR quantification of gDNA

The greatest quantity of gDNA extracted from canine blood, as determined by qPCR, was obtained using the modified salt precipitation method on EDTA blood (GE mean 72.1%, SE +/-17.9), and least was obtained using the same method on clotted blood samples (GC, 3.0% +/-2.6). The quantities of gDNA recovered by silica spin columns on EDTA blood (QE, 68.7% +/-11.3) and clotted blood (QC, 50.5% +/-5.4) were similar to that obtained by phenol-chloroform extraction of EDTA blood (PE, 59.5% +/-29.2) but greater than of phenol-chloroform extracted clots (PC, 13.0% +/-2.5). The quantity of gDNA recovered was significantly less ($P = 0.004$) using the modified salt precipitation method on clotted blood (GC) than using the same method on EDTA preserved blood (GE) or using a silica column method on EDTA preserved blood (QE).

DNA Purity

A graphical illustration of the purity of gDNA recovered using different methods of extraction is presented in Figure 2. The purity of gDNA extracted from canine blood samples using the modified salt precipitation method on clotted blood (GC mean 2.95 SE+/-0.54; QC 49.3% +/-6.0) samples had a significantly greater A_{260}:A_{280} ratio than those extracted by the other methods (GE, 1.96 +/-0.05; PE, 1.63 +/-0.04; PC, 1.65 +/-0.04; QE, 1.86 +/-0.02; QC, 1.90 +/-0.03 [$P = 0.001$]).

The low quantity of gDNA recovered in some GC and GE samples dictated that the electropherogram could not be assessed for quality. No differences in the quality of gDNA were evident by visual analysis of electrophoresis patterns, as all samples were assigned to be of "high" quality.

PCR Inhibition

A graphical illustration of the inhibition metrics of recovered gDNA using different methods of extraction is presented in Figure 3. All gDNA samples resulted in a degree of assay inhibition, when compared to the positive control. The level of PCR inhibition was approximately 50% for the silica column (QE mean 50.7% SE+/-4.5; QC, 49.3% +/-6.0) and modified salt precipitation methods (GE, 50.0% +/-6.8; GC 46.2% +/-13.9) using clotted or EDTA blood, and the phenol-chlor0form method used on EDTA blood (PE, 67.2 +/-2.6) when compared to the positive control. However, the level of inhibition significantly increased (i.e. the percentage of positive control reduced) for gDNA samples from clots extracted using the phenol-chloroform method (PC, 82.3%+/-3.5, $P = 0.005$) when compared to all of the other methods.

DISCUSSION

The quantification of gDNA is most accurately determined using qPCR. Unfortunately this technique is relatively slow and expensive in comparison to other methods such as spectrophotometric (393) or fluorometric analysis (386), and therefore less suited to high throughput methodology (394). However, qPCR directly reflects the utility of the extracted sample for end-point usage, unlike the other methods. The data shows that the method of extraction did not significantly affect the

quantity of gDNA recovered from EDTA blood as determined by qPCR or the spectrophotometric measures, although the rates were slightly higher than those quoted in the product literature of both the silica columns (5.9 μg per 10^6 cells (395)) and the modified salt precipitation method (2.3-7 μg per 10^6 cells (396)).

Marked differences were noted in gDNA recovery from clotted blood between different methods. The extremely low level of recovery of gDNA from clotted blood extracted by the modified salt precipitation method (without marked concurrent PCR inhibition) suggested that gDNA was more susceptible to accidental loss during this procedure when compared to the other techniques utilising organic solvent extraction or DNA binding methodologies. The low level of gDNA recovery from clotted blood using the phenol-chloroform method suggested contamination of the extracted samples with PCR inhibitors such as phenol (227), which was supported by the results of the internal amplification control experiment. Similarly, a degree of protein contamination (A_{260}:A_{280} ratio less than 1.8 (228)) may also have contributed to qPCR inhibition, and the subsequent disparity between the spectrophotometric results and the qPCR results. Thus, although the spectrophotometer quantification of gDNA is more convenient it is not an accurate assessment of the functional quantity of gDNA extracted from clotted blood using the phenol-chloroform method.

Endogenous endonuclease activity results in the loss of DNA integrity (the breakage of DNA strands) when normal cellular processes are stopped (397). The integrity of gDNA may be determined from ethidium bromide-stained agar gel electrophoresis (386) or by length determining quantitative PCR (392). We did not identify differences in gDNA integrity in samples extracted by different methods in this

experiment, as all detected samples were defined as high quality, and thus integrity could not be used to discriminate the different methods of gDNA extraction.

A degree of PCR inhibition was observed in all our samples, although the effect was the greatest for clotted blood DNA extracted using a phenol-chloroform method. In practical terms this translated to the PCR, requiring one extra cycle to achieve the same quantity of product as the control when silica column or modified salt precipitation extracted gDNA was present in the PCR well with the internal amplification control. Conversely, if gDNA from clotted blood samples extracted by the phenol-chloroform method was used as template, approximately two and a half extra cycles were required. Thus, one may infer from this result that the quantity of gDNA extracted using the phenol-chloroform method on clotted blood may, in reality be similar to that extracted using the silica column method, but inhibition of the qPCR results in the lower quantity reported as measured by the qPCR quantification. Furthermore, the true level of PCR inhibition in the phenol-chloroform extracted samples may be an underestimate, as the gDNA samples were normalised using the spectrophotometer result.

Whole genome amplification (WGA) provides an alternative method for maximising the quantity of gDNA obtained from an individual sample. Less invasive methods of sample collection from animals, such as buccal swabs, can yield suitable quantities of gDNA for downstream tests when amplified using this technique (385). However, the concordance of genotyping tests using WGA samples, when compared to source material, is lower for buccal swabs than for blood samples (385). Therefore, differences in quantity of extracted gDNA may not be critically important, but the

value of residual blood as a gDNA source is fundamental for developing large scale sample cohorts of well phenotyped samples. Furthermore, other important factors not considered in this experiment, such as processing time and cost, must also be taken into account when selecting the most appropriate method of gDNA extraction.

CONCLUSIONS

The silica column method recovered gDNA in a more consistent manner irrespective of the sample preservation when compared to other methods. Phenol-chloroform and modified salt precipitation extractions demonstrated similar relative recovery of gDNA with EDTA preserved blood, but were less efficient at recovering qPCR quantifiable gDNA from clotted blood. Ultimately, the quantification of these differences on specific downstream processes will determine how important these effects are.

Table 1

Primer, probe and amplicon sequence, and the performance metrics off the VZV inhibition assay.

Forward Primer Sequence	*Reverse Primer Sequence*	*Amplicon Sequence*	*Probe*	R^2	*% Efficiency*	*Range low*	*Range high*
caaagcagaacatc gagcac	gggtgtcacagggt gactaag	caaagcagaacatc gagcaccgtttcctg ctggacgttttggattt cttagtcaccctgtga caccc	tcctgctg	1	100.81	23.01	32.97

Figure 1

Mean quantity of gDNA extracted from canine whole blood using three different methods of gDNA Extraction, as determined by a spectrophotometer (Spec; µg DNA measured per 10^6 cells) and direct quantification by qPCR (percentage DNA recovery with standard error bars) (Puregene Gentra extracted EDTA blood [GE] or clotted blood [GC]; Phenol-chloroform extracted EDTA blood [PE] or clotted blood [PC]; and QIAamp Midi Kit extracted EDTA blood, [QE] or clotted blood [QC]).

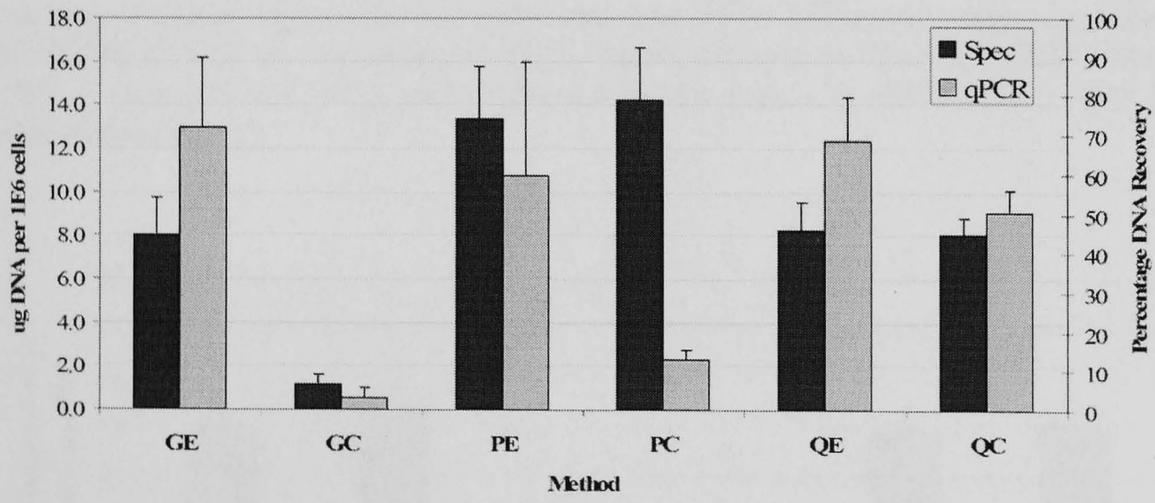


Figure 2

The mean 260 nm to 280 nm absorbance ratio of gDNA extracted from canine whole blood using three different methods of DNA Extraction, as determined by a spectrophotometer. Data presented with standard error bars (Puregene Gentra extracted EDTA blood [GE] or clotted blood [GC]; Phenol-chloroform extracted EDTA blood [PE] or clotted blood [PC]; and QIAamp Midi Kit extracted EDTA blood, [QE] or clotted blood [QC]).

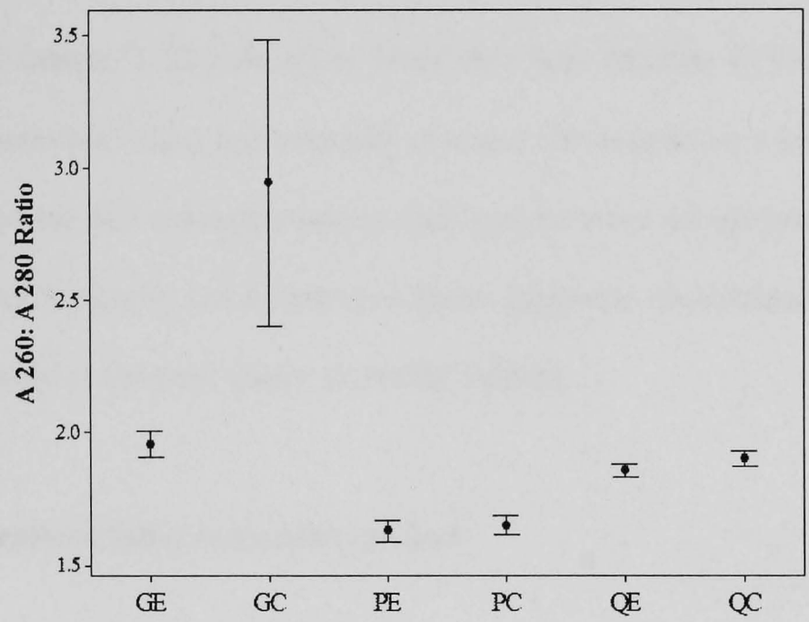
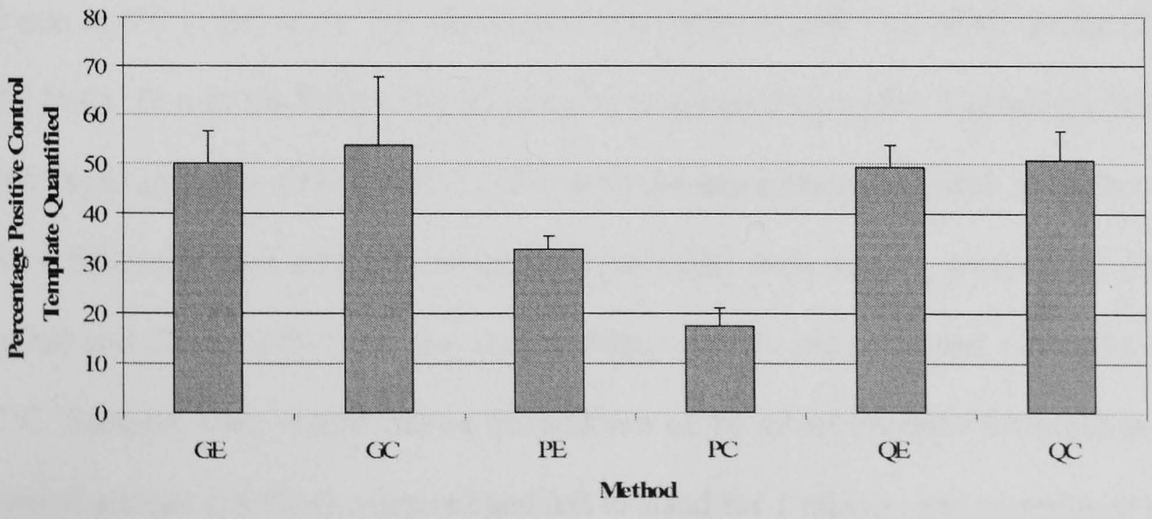


Figure 3

Mean percentage inhibition of a positive internal amplification control, using a synthetic template oligonucleotide (plus standard error) (Puregene Gentra extracted EDTA blood [GE] or clotted blood [GC]; Phenol-chloroform extracted EDTA blood [PE] or clotted blood [PC]; and QIAamp Midi Kit extracted EDTA blood, [QE] or clotted blood [QC]).



Supplementary information

White blood cell counting method

A precise nucleated cell count of each sample was obtained by diluting 10ul of each EDTA blood sample 1:20 with white blood cell lysis solution (3.0% glacial acetic acid, 0.1% methylene blue) and manually counting the cells using a haemocytometer. Estimations of the cell counts present in each sample were determined using the cell count, the sample weight, and a correction factor applied to the calculations to account for blood density (1.06 g/ml, Geigy Scientific Tables).

Phenol-chloroform DNA extraction method

Clotted samples were macerated with the end of a 3 ml plastic Pasteur pipette. Red blood cells were lysed by incubation with 20 ml of a lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 0.1 mM Na_2EDTA, pH 7.4), then incubated for 30 min on ice, followed by centrifugation at 1200 rpm for 10min at 4°C. The supernatant was removed and a further 10 ml of lysis buffer added to re-suspend the pellet, and sample centrifuged for 10 min at 4°C (1200 rpm). The supernatant was removed and 5 ml of SE-buffer (75 mM NaCl, 25 mM Na_2EDTA, pH 8) added to re-suspend the pellet. The sample was centrifuged again for 10 min at 4°C (1200 rpm) the supernatant discarded. A further 5 ml of SE-buffer was added to re-suspend the pellet with 40 µl proteinase K (10 mg/ml) and 250 µl 20% SDS, the sample shaken gently, and incubated overnight at 37°C. Samples were cooled before the addition of 10 ml of Phenol / Chloroform / Isoamyl alcohol (25:24:1), vortexed and left to stand for 1 minute, and centrifuged at

3000 rpm for 45 minutes. Genomic DNA (gDNA) was precipitated from the aqueous layer by addition to 10 ml of absolute ethanol and stored at -20°C overnight. Precipitated gDNA was separated by centrifugation at 3000 rpm for 15 minutes and washed with 2 ml 70% Ethanol and centrifuged at 3000 rpm for 5 minutes, before removal of the supernatant. The gDNA pellet was air dried for 15-20 minutes before re-suspension in 300 µl Tris-EDTA (10 mM Tris, pH 8.0; 1 mM EDTA, pH 8.0) and overnight rotation.

Chapter 8

A candidate gene association study of canine joint disease: A common pathogenesis?

Dylan N. Clements[1,2],

Andrea D. Short[2],

Annette Barnes[1],

Lorna Kennedy[2],

John Ferguson[3]

Steven J. Butterworth[4],

Noel Fitzpatrick[5],

Matthew Pead[6],

David Bennett[7],

John F. Innes[1],

Stuart D. Carter[1],

William E.R. Ollier[2].

[1]The Musculoskeletal Research Group, Faculty of Veterinary Science, The University of Liverpool, L69 7ZJ

[2]Centre for Integrated Genomic Medical Research, The University of Manchester, M13 9PT

[3]East Neuk Veterinary Clinic, St Monans, Fife, KY10 2DW

[4]Weighbridge Referral Centre, Swansea, SA6 8QF

[5]Fitzpartick Referrals, Farnham, Surrey, GU10 2DZ

[6]Queen Mother Hospital for Small Animals, The Royal Veterinary College, AL9 7TA

[7]The University of Glasgow Veterinary School, Glasgow, G61 1QH

Paper not yet submitted for publication

ABSTRACT

Introduction

Canine osteoarthritis (OA) commonly occurs in association with articular diseases, such as hip dysplasia (HD), elbow dysplasia (ED) or cranial cruciate ligament rupture (CCLR). We hypothesised that the genomic risk for the development of OA would be identified by evaluating the allele frequencies of candidate gene single nucleotide polymorphisms (SNPs) in dogs with OA associated with different articular diseases.

Materials and Methods

DNA was extracted from blood samples from Labrador Retrievers (LR) and Golden Retrievers (GR) surgically treated for ED (LR n = 81), HD (LR n = 32), CCLR (LR n = 51, GR n = 45) and confirmed to have radiographic evidence of OA, and a general population of dogs (LR n = 341, GR n = 94). One hundred and thirteen SNPs in twenty candidate genes were genotyped using the Sequenom MASSarray platform. Odds ratios, minor allele (MAF) and haplotype frequencies were calculated and compared by Chi square analysis and corrected by Monte Carlo simulation tests. Data were also stratified on the basis of known sex (ED and HD) and neuter (CCLR) status.

Results

Significant associations were identified for SNPs in interleukin 12B (*IL12B*) and tumour necrosis factor alpha (*TNFα*) with ED in LRs, interleukin 4 (*IL4*) and interleukin 6 (*IL6*) with HD in LRs, *IL4* and *IL12B* with CCLR in LRs, and interleukin 10 (*IL10*) and ankyrin repeat domain 10 (*ANKRD10*) and for CCLR in GRs. Following population stratification, significant associations were identified for SNPs in osteonectin (*SPARC*) and leptin receptor (*LEPR*) with ED in LRs, *IL4* with HD in LRs, and Zinc finger SWIM-type containing 2 (*ZSWIM2*) with CCLR in LRs.

Conclusions

A common genomic risk for the development of CCLR was not identified between the two breeds of dog evaluated in this study. Common genomic risks were identified across the different articular diseases evaluated in LR populations. The pathological basis to different articular diseases may be similar within a single dog breed, and demonstrate similar gene associations to their human equivalent diseases.

INTRODUCTION

Osteoarthritis (OA) is a common, debilitating condition of mammalian joints, characterised by the destruction of articular cartilage, resulting in pain and dysfunction of the affected joint. OA is estimated to affect up to 20% of dogs over one year of age (7) in the general dog population. The joints most commonly affected by OA in the dog are the hip, the elbow and the stifle. Historically, OA of these joints was understood to be secondary to primary diseases, such as hip dysplasia (HD), elbow dysplasia (ED) and cranial cruciate ligament rupture (CCLR).

Breed risks for the development of HD, ED and CCLR are marked. The Labrador Retriever demonstrates a 3.4 fold increase in risk for the development of HD (48), a 20.5 fold increase in risk for developing the primary component of ED (48) (fragmented coronoid disease) and a 5.5 fold increase risk for developing CCLR (108). Sex predisposition to the development of each of theses diseases also exist, with males dogs being a greater risk for developing hip OA (59) and ED (FCP) (83), and neuter status conferring an increased risk for developing CCLR (88,108). The estimates of heritability for HD vary between 0.18 and 0.74 (49-53), estimates for ED vary between 0.10 and 0.77 (73,78-81), and estimates for CCLR vary between 0.27 and 0.31 (33,92). A genetic correlation between the HD and ED have also been identified in a population of Rottweilers (80), suggesting that these traits may be influenced by the same genetic and/or environmental factors in certain breeds.

Recent evidence suggests that genetic factors may additionally affect the development of OA in dogs affected by an articular disease. Differences in the breed tolerance of

passive hip laxity for the development of hip OA imply that genetic differences affect the development of canine hip OA (12). The severity of OA in dogs presenting with ED, HD and CCLR can vary widely, which is a function of disease duration, animal activity, nutrition status and almost certainly genetic profile. Thus, whilst the significance of primary versus secondary canine OA is undetermined, canine OA *per se* is likely to have a significant genetic background.

In contrast to human OA (46), the genomic basis of OA in dogs has received limited investigation to date. A previous case-control candidate gene study of single nucleotide polymorphisms (SNPs) in four candidate genes (Fibronectin 1, type 9 collagen alpha 1 chain, type 9 collagen alpha two chain and cartilage oligometric protein) failed to identify any significant associations with between SNPs and the development of CCLR in a population of Newfoundland dogs (398). A study of microsatellite markers adjacent to 14 candidate collagen genes for ED also failed to identify significant associations with the development of the disease. A microsatellite marker (FH2320) on canine chromosome 3 (CFA3) has been linked with the development of osteophytosis of the cranial and caudal acetabulum in Portuguese Water dogs (60). In a separate study of a Labrador Retriever and Greyhound cross pedigree, putative quantitative trait loci (QTL) contributing to macroscopic evidence of hip OA were identified on CFA05, 18, 19, 23 and 30 (CFA30) (61). Further quantitative assessment of the same pedigree also revealed that hip OA was inherited additively (59).

The majority of candidate genes studies in human OA have evaluated genes which were associated with the molecular pathogenesis of the disease, such as cytokines and

structural components of the extracellular matrix (46). The most successful study investigating candidate gene associations with human OA focused on genes which are differentially expressed in human OA synovium and cartilage (130). Subsequently, a number of these gene polymorphism associations have been reproduced in separate cohorts of women (Oestrogen receptor alpha, bone morphogenic protein 2), men (vitamin D receptor), and both women and men (a disintegrin and metalloprotease domain 12, cartilage intermediate layer protein and osteoprotegrin) (399). Furthermore the prediction of OA risk in males and females could be attained by combining several of the genes which were consistently shown to be involved in OA susceptibility (216).

We hypothesized that genomic risk to the development of canine OA, or joint disease *per se*, would be similar across the three most common articular diseases affecting dogs (ED, HD and CCLR) and between different breeds of dog for the same disease (CCLR). We further hypothesized that these genomic risks could be elucidated by evaluating the allele frequencies of SNPs in candidate genes in populations of dogs with ED, HD and CCLR of a single breed, and between two breeds with a common disease (CCLR).

MATERIALS AND METHODS

Candidate gene selection

Candidate genes were selected on the basis of previous association of polymorphisms reported with OA in man or from differential gene expression in articular tissues from

canine or human OA joints. A full list of the genes evaluated, the SNP positions, gene function and justification for inclusion as a candidate gene are reported in Table 1.

The genes selected were; Ankyrin repeat domain 10 (*ANKRD10*), ATPase, Class VI, Type 11B (*ATP11B*), Interleukin 1 alpha (*IL1α*), -2 (*IL2*), -4 (*IL4*), -6 (*IL6*), -10 (*IL10*), -12B (*IL12B*), Leptin Receptor (*LEPR*), Matrix Metalloproteinase 3 (*MMP3*), -9 (*MMP9*), -13 (*MMP13*), Secreted protein, acidic, cysteine-rich (osteonectin / *SPARC*), Tissue inhibitor of metalloproteinase 1 (*TIMP1*), -2 (*TIMP2*), -3 (*TIMP3*), -4 (*TIMP4*), Tenascin C (*TNC*), Tumour Necrosis Factor (*TNFa*), Zinc finger, SWIM-type containing 2 (*ZSWIM2*).

Cohort Collection

Genomic DNA was extracted from residual clotted and ethylenediamine tetra acetic acid preserved blood samples using a standard phenol - chloroform extraction method. Samples were suspended in Tris-EDTA and normalised to 5 ng/μl. All diseased samples were obtained from the UK DNA Archive for Companion Animals (387) and all control samples were obtained from a general population of dogs undergoing vaccination. The breeds and orthopaedic diseases evaluated were selected by choosing cohorts for which at least 30 samples had been collected from individuals within a single breed, for a given condition. All samples from cases were collected by veterinary orthopaedic specialist surgeons from dogs surgically treated for ED, HD or CCLR, and with no clinical evidence of a concurrent orthopaedic condition (ED, HD or CCLR) at the time of treatment. All cases had radiographic evidence of osteoarthritis of the affected joint at the time of surgery. Samples were collected from Labrador Retrievers (LR) and Golden Retrievers (GR) being surgically treated for ED

[FCP] (LR, n = 81 [11 female, 7 female neutered, 61 male, 1 male neutered, 1 male neutering status unknown]), HD (LR, n = 32 [9 female, 7 female neutered, 14 male, 1 male neutered, 1 male unknown neutering status]), CCLR (LR n = 51 [10 female, 19 female neutered, 15 male, 7 male neutered], GR 45 [5 female, 26 female neutered, 10 male, 4 male neutered]), and a general population of dogs undergoing vaccination (LR n = 341 [89 female, 29 female neutered, 105 male, 28 male neutered, 93 sex and neuter status unknown], GR n = 94 [42 female, 11 female neutered, 34 male, 6 male neutered, 1 sex and neutered status unknown]). An internal genotyping control was included on each plate.

SNP Identification

Selected regions of each candidate gene were amplified by PCR. The PCR product was assessed for the presence of a polymorphic product using denaturing high performance liquid chromatography (DHPLC) (400), and amplicons with melt curve analyses indicating a SNP were sequenced. SNPs identified were annotated to a genomic position by alignment of the sequence with the canine genome (277). SNPs were designated a genomic location on the basis of their position (Pre-gene = <10000bp upstream of exon 1, Promotor [<1000bp of the start of exon 1], post-gene = <10000bp downstream of exon 1). Forty six SNPs were selected from a previous study (In; *IL1α, IL2, IL4, IL6, IL10, IL12* and *TNFα*) (401), twenty four SNPs were identified as described by DNC (15 SNPs in; *ATP11B, ANKRD10, SPARC, TNC, ZSWIM2*) and AB (9 SNPs in; *LEPR* and *MMP13*). For full details of the methods see appendix 3. A further forty three SNPs were selected from the canine genome sequence (278) and a canine SNP database (402). A further 27 SNPs could not be analysed as genotyping tests failed, but are listed in appendix 3.

Genotyping

Genotyping was performed using the Sequenom MASSarray platform (Sequenom, Hamburg, Germany) as previously described (403). Briefly, primers and probes were designed using the Sequenom Assay Design Software Version 3, and synthesized by Metabion AG (Martinsried, Germany). Primers and probes were pooled as recommended by the manufacturer's instructions (404). Multiplex PCR reactions, product clean-up and probe extension reaction were performed in 384-well plates with 20 ng of DNA per well, using iPLEX Gold reagents (404). Samples were diluted and de-salted with 6 mg of resin before dispensation onto a SpectroCHIP (Sequenom) using the Sequenom Nanodispenser, before genotype identification by matrix-assisted laser desorption / ionisation - time of flight mass spectrometry (MALDI-TOF MS). For full details of the methods see appendix 3.

Data Analysis

Genotype and phenotype data were imported into BCgene software (405), which was used to calculate genotyping rates, minor allele frequencies (MAF) and Hardy-Weinberg equilibrium (HWE) for each control population. SNPs were not analysed further if the call rates were below 80% or if the control population was not in HWE. Case-control comparison of MAFs was performed by Chi^2 (χ^2) comparison and odds ratio (OR) and 95% confidence intervals (CI) calculated using the BC gene software. Data analysis was repeated after stratification of each population on the basis of previously reported disease risk factors (neuter status for CCLR [evaluation of neutered animals only], and sex for ED and HD [evaluation of male animals only]). Significant differences ($P < 0.05$) were checked for multiple permutations by Monte Carlo simulation (406), using a freely available software program (T1 statistic,

CLUMP) (407). Any SNPs with MAF demonstrating sex or neuter associations within the control population were removed from further analysis (3 SNPs eliminated).

Haplotype frequencies were estimated for each cohort (control populations, disease population and control population stratified on sex or neuter status). SNPs were considered for haplotype analysis if the minor allele frequencies were greater than 5% and the SNPs were in HWE in the population analysed. Thus haplotypes were calculated for 13 genes in GRs (*ANKRD10, ATP11B, IL1α, IL4, IL6, IL10, IL12, LEPR, SPARC, TIMP3, TIMP4, TNC, TNFα* and *ZSWIM2*) and for 10 genes in LRs (*ANKRD10, MMP9, IL4, IL6, IL10, IL12, SPARC, TIMP3, TNC* and *TNFα*). Maximum likelihood haplotype frequencies were computed using an expectation-maximisation algorithm, using HelixTree version 4.1.0 software (GoldenHelix, Inc., Bozeman, USA). Haplotype frequency estimates were multiplied by the number of chromosomes in diseases and control groups to generate contingency tables. Frequency estimates were compared between controls and cases by χ^2 analysis checked for multiple permutations using CLUMP (T1 statistic) and ORs and CIs calculated for the haplotypes of each gene using a web-based statistical calculator (408). Contingency tables containing values less than 5 were analysed using web-based Fishers exact test calculator (409). The significant associations are presented in Tables 3 and 4.

For each cohort of disease samples using SNPs with a MAF ranging from 5% to 50% this study was powered to detected risk alleles with ORs ranges of 1.64 - 2.43 (LR ED) to 2.10- 3.46 (GR CCLR) and protective alleles with OR ranges of 0.61 - 0.13

(LR ED) to 0.48 - undetectable (GR CCLR) if the allele is protective, at 80% power ($P < 0.05$).

RESULTS

A total of 113 SNPs were analysed; in Labrador Retrievers, 44% (n = 54) of SNPs were informative (MAF >1%), 14% (n = 16) of SNPs demonstrated low heterozygosity (MAF<1%), 15% (n = 17) were monoallelic and 27% (n = 30) were not in HWE; in Golden Retrievers 61% (n = 69) SNPs were informative, 27% (n = 31) were monoallelic and 12% (n = 13) were not in HWE. The average genotyping rate for the LR samples was 94.5% (range 81.3% - 99.6%), and 96.6% for the GR samples (range 84.9% - 100%). The concordance of the internal genotyping control between plates was 100%. A further 27 SNPs could not be analysed as genotyping tests failed.

Case-control comparison of MAF identified fifteen SNPs with allele frequencies which were significantly associated with either risk (n = 3) or protection (n = 12) of orthopaedic disease after correction for multiple permutation testing (Table 2). Minor allele frequencies of two SNPs were associated with the risk of developing more than one disease (4_13S97; increased risk of CCLR and HD in LRs, and 12B_01M115; increased risk of CCLR and ED in LRs).

The MAF of six SNPs in two genes (*ANKRD10* and *IL10*) were associated with CCLR protection in GRs. The MAF of single SNPs in each of two genes (*IL4* and *IL12B*) were associated CCLR risk in LRs, and the MAF of one SNP *IL4* was

associated with the protection from CCLR in LRs. The MAF of two SNPs in *IL4* were associated with the risk of HD in LRs, and the MAF of one SNP in *IL6* was associated with HD protection in LRs. The MAF of four SNPs in *IL12B* and *TNFα* were associated with ED protection in LRs, and one SNP in *IL12B* was associated with risk of ED in LRs.

When the data were stratified on the basis of previously described breed and neuter risks (i.e. only "at-risk" cases and controls were considered), then two associations of SNP MAFs with disease status were maintained consistent with that identified in the non-stratified populations. The MAF of two SNPs in *IL4* were again associated with the risk of HD in male LRs. The MAF of individual SNPs in three other genes which were not in HWE in the general LR population but which were in HWE in the stratified LR population were associated with protection from ED in male LRs (*SPARC* and *LEPR*) or the risk of CCLR (*ZSWIM2*) in neutered LRs.

One *ANKRD10* haplotype (GAGG) was associated with a reduced risk of CCLR in GRs (OR O.49), one haplotype of *IL10* (GGACGACA) associated with a reduced risk of CCLR in GRs (OR 0.45) and one haplotype of *IL10* (AGGCCATG) was associated with an increased risk of CCLR (OR 1.74) in GRs. One haplotype of *IL4* (CCAGAG) was associated with the risk of both CCLR (OR 1.69) and HD (OR 2.17) in LRs. A further haplotype of *IL4* (TGCACT) was associated with a reduced risk of CCLR in LRs (OR 0.57). A single haplotype of *IL6* (GGGAG) was associated with the protection from risk of HD in LRs (Haplotype frequency controls 6.8%, haplotype frequency in diseased <0.1%). One haplotype of *IL12B* (CCAATGGC) was associated with both the risk of CCLR (OR 2.03) and ED (OR 1.58) in LRs. A further

haplotype of *IL12B* (ATCTCAGT) was associated with the reduced risk of ED in LRs (OR 0.47). No haplotype associations were identified in any other genes, or between the neuter or sex status of the control populations. For full details of the results see appendix 3.

DISCUSSION

The identification of gene sequence polymorphisms associated with disease phenotypes is fundamental to understand the genomic basis of disease, but more importantly to facilitate the development of molecular diagnostic tests which identify individuals carrying disease related mutations within the general population. The great advantage of such tests in domestic animals is that they can be applied to identify carrier intervals so that they may be removed from breeding populations, without relying on the identification of the phenotype, which may not be evident at the breeding age. Many mutations causing monogenetic ocular disorders have been identified in dogs (410), and these discoveries have led to the development of genomic DNA-based screening tests which are widely used, and resulted in the production of novel gene replacement-based therapies of such disorders (411). Whilst these strategies are extremely successful for simple monogenetic disorders, the characterisation of mutations causing polygenetic disorders, such as canine hip or elbow dysplasia is extremely difficult. Consequently progress in developing tests which may identify disease associated mutations has been extremely slow. Thus the identification of individuals carrying the disease associated mutations currently remains radiographic assessment of the phenotype, and predictably progress in reducing the prevalence of disease in the general population has been relatively slow

(54). Furthermore the prospects of developing gene based therapeutic treatments for such complex disorders are likely to be even more difficult, because of the numbers of likely causative mutations involved and the marked secondary changes which have developed in the diseased joint by the time of presentation. For canine articular diseases at least, prevention will probably be more attainable than cure.

A number of SNPs in candidate genes were found to associate with the development of canine articular disease in this study. These may represent risks for susceptibility to either the primary articular disease or to the subsequent development of OA. Furthermore, individual SNPs and gene haplotypes of *IL4* and *IL12B* demonstrated associations with disease risk and these were replicated in different disease cohorts of LRs. Thus our hypothesis, that genomic risk to the development of canine OA or joint disease *per se*, would be similar across common articular diseases was supported, although we could not identify a common risk between two different breeds for the same disease.

The significant association of two SNPs in the *IL4* gene (4_13S97 and 4_8R458) with the risk of HD in LRs was identified in both the overall and stratified dog population, and demonstrated the greatest OR values. Further association of one of these SNPs (4_13S97) with the risk of of CCLR in LRs supports the hypothesis that the genomic risk to canine OA, or articular disease, was common across different diseases. A common haplotype of *IL4* (CCAGAG) was also associated with the risk of CCLR and HD in LRs. Interleukin 4 function has also been implicated in human hip OA, where functional variants of the interleukin 4 receptor gene have been associated with the development of female hip OA (164). The importance of *IL4* in joint homeostasis at

the cellular level is well recognised, with *IL4* demonstrating marked protective effects on chondrocytes *in vitro* such as the prevention of *IL1* mediated *MMP3* expression, and *IL4* inhibits *IL1β* and *TNFα* synthesis in OA synovium (412).

The *IL4* gene is positioned on CFA11, which has previous been reported to harbour a putative quantitative trait loci for canine hip laxity in a Labrador Retriever cross Greyhound pedigree (57), using a marker approximately 8Mbp downstream from the *IL4* gene. Thus the association of *IL4* SNPs with HD in LRs may have been anticipated. Without further study it is impossible to determine whether these associations represent functional variants which contributed to disease, or whether they are simply in linkage disequilibrium with a functional genomic mutation, as LD is extremely high in dogs (117). Intriguingly, the same *IL4* SNPs which are associated with the risk of HD (4_13S97 and 4_8R458) have also both been previously associated with an increased risk of diabetes in Cairn terriers (401). This suggests that if these SNPs do represent functional mutations, they may contribute to the development of very different diseases in different breeds of dog.

Interleukin 12 is a pro-inflammatory cytokine which demonstrates increased expression in human OA synovium (413), although it has not been investigated as a potential candidate gene in canine or human OA. Significant associations of both an individual SNP (12_M115) and a haplotype of *IL12B* (CCAATGGC) with the risk of CCLR and ED in LRs further supports the hypothesis that a common risk factor to canine OA or articular disease could be identified within an individual breed. Interleukin 12 is located on a canine chromosome harbouring QTL which are linked to hip laxity (CFA04) in a LR-GH pedigree, although at a long distance (21Mbp)

from the linked markers, which suggests that there may be no significance to this finding (57).

Interleukin 10 is an immunomodulatory cytokine, produced primarily by monocytes. The association of SNPs and haplotypes of *IL10* with CCLR in GRs was of interest as we have previously demonstrated an increased expression of *IL10* in the synovial tissue of dogs with CCLR (414). Mutations in *IL10* have been associated with human knee OA (166), but not human hand OA (167). Thus, the association of *IL4* mutations with HD in LRs and *IL10* mutations with CCLR in GRs, supports the assertion that genomic risk factors for canine disease often mirror those identified for the equivalent human disease (human hip and knee OA respectively) as has been reported with other canine diseases such as diabetes mellitus (401).

The common association between SNPs and haplotypes with CCLR and ED or HD, but not between HD and ED in LRs, or between CCLR in LRs and GRs was contrary to what we might have expected. Furthermore, one cannot exclude the possibility that these associations are significant in only a single condition (e.g. CCLR) which had not developed at the time of investigation, but may have developed over time in the co-associated disease cohort (e.g. ED or HD). Without longitudinal studies, this possibility cannot be excluded.

Two of the SNPs which were significantly associated with OA were also positioned on canine chromosomes previously linked to canine joint disease. One of the genes was associated with the development of ED (*LEPR*) and is located on a canine chromosome (CFA05) harbouring a QTL which is linked to the development of OA

(57). The long distance of the *LEPR* SNPs (32Mbp) from the linked markers, suggesting that there may be no significance to this observation. Similarly, *ANKRD10* which was associated with the development of CCLR in the GR dog population is only 2Mbp from a marker of hip laxity (in a Labrador Retriever cross Greyhound dog population) on CFA22. However, no SNPs in *ANKRD10* were associated with HD in the LR population we evaluated, which would have been expected if a common genomic risk exists between the breeds for these different diseases.

The two breeds compared have markedly different risks for developing CCLR. Labrador Retrievers demonstrate a significantly increased risk of developing CCLR (OR 5.1 [95% CI = 3.5-7.4]) compared to a general population of dogs, whereas GRs are not at risk (OR 0.48 [0.22-1.05]) (108), and other studies suggest the prevalence of CCLR in GRs is lower than the average for other breeds (88). We may have been more likely to identify a common genomic risk to CCLR if we had evaluated a second breed with an increased risk of developing CCLR. Unfortunately no cohort was available with sufficient sample numbers to meet the inclusion criteria. Furthermore, the absence of a genomic risk to the development of CCLR in GRs when the data were stratified on the basis of the neutering status reflected in part the small number of samples present in the data set.

In most cases the odds ratios for SNPs associated with diseases evaluated in this study were below the limits calculated for 80% statistical power. A proportion of our control cases were likely to have developed HD, ED or CCLR in the future. This would artificially lower the differences and significance of the changes in the minor allele and haplotype frequencies recorded. Thus the true odds ratios for the minor

allele and haplotype frequencies would probably be greater if control populations were available which had been accurately phenotyped for all three diseases. The true prevalence's of each of these diseases in the control populations are unknown so estimations of the degree of misclassification "noise" in the control population could not be made, as robust epidemiological studies of orthopaedic diseases have not been performed in dogs to date.

The collection of well phenotyped controls would have required general anaesthesia and radiographic examination of a large number of older dogs, and was precluded on practical and ethical grounds. The phenotypic quality of control populations for this type of study is further compounded by the fact that many of these diseases are progressive and clinical and radiographic signs may not develop until late age, making accurate phenotype identification difficult. Conversely it should be noted that the data quality of the disease cohorts is extremely high, as all cases were at the extreme end of the phenotype, requiring surgery for their underlying condition, and each was diagnosed by a veterinary orthopaedic specialist. Difficulties in obtaining appropriate numbers of samples to accurately determine differences in allele and genotype frequencies for case-control study of canine disease with dogs populations are a recognised issue (403). The limited cohort sizes were probably the primary reason for the differences between the associations identified in general and stratified populations. This was because the allele frequencies, for most genes, were broadly similar between the general and stratified populations, and only a limited number of SNPs (n = 3) and no haplotypes demonstrated sex or neuter associations in the control population.

A significant proportion of the SNPs evaluated were out of HWE (27% for LR, 12% for GR) although the proportions of SNPs out of HWE were slightly lower in this study than has been previously reported (415), which reflects the large number of individuals genotyped. Interestingly, the MAFs of candidate genes SNPs in the control populations evaluated in this study were virtually identical to those reported in a previous study of candidate susceptibility genes in canine diabetes (401), even though the populations were entirely unrelated which highlights both the repeatability of the genotyping method and the random nature of the control population. For example, the MAF of the G allele of the *IL10* SNP 10S308 was 5% in the LR control population, and 81% in the GR control population, and the previously reported frequencies of these alleles in these breeds are 6% and 81% respectively (403).

The majority of SNPs we evaluated were in the intronic (n = 33) or promotor (n = 28) regions, in contrast to the predominantly exonic SNPs evaluated most successful human candidate gene association studies of OA (130,216). Ideally exonic SNPs would have been evaluated for all genes, but few or none were identified in the genes screened. Although open-access canine SNP databases are available (402), detail is presently lacking for individual genes. This is because the canine genome has only recently been published (218), and the number of researchers working in this field is relatively small. Consequently, SNP identification is most thoroughly performed using *in vitro* methodology, which is both expensive and time consuming. We anticipated that the high level of linkage disequilibrium present in the dog (117) would reduce the number of SNPs required to identify significant associations with a diseased gene, and thus complete evaluation of each gene for SNPs would not be required. Clearly, the limited number of SNPs evaluated for some genes (and the

subsequent inability to perform haplotype analysis) means that they cannot be fully eliminated as candidate genes by association with the diseases under investigation on the basis of the results of this study. Furthermore meaningful interpretation of the disease associated mutations in non-coding regions is impossible without further *in vitro* evaluation of their function, and the complete characterisation of all polymorphisms in each gene.

CONCLUSIONS

A common genetic risk for the development of common articular disorders or OA was identified between different diseases within a breed, but not within the same disease in two different breeds of dog. A number of the genetic associations identified for canine diseases have also been reported for their equivalent human disease. All our significant associations require repeat testing with different canine cohorts to validate their true significance.

Table 1

Candidate genes selected for evaluation in canine OA and their SNP positions. (CFA = canine chromosome, U =upstream, P = Promotor, E = Exon, I = Intron, D = Downstream), and their justification for evaluation as candidate genes.

Name	*CFA*	*U*	*P*	*E*	*I*	*D*	*Function*	*Justification for Evaluation Inclusion*	*Reference*
Ankyrin repeat domain 10 (*ANKRD10*)	22			2	3		Structural component of muscle	Increased expression in canine ruptured CCL	(319)
ATPase, Class VI, Type 11B (*ATP11B*)	34				6		Cell Membrane Transport	Increased expression in canine ruptured CCL	(319)
Interleukin 1 alpha (*IL1α*)	17		3	2	2		Pro-inflammatory cytokine	Increased expression in canine OA synovium and ruptured CCL. SNPs associated with human OA	(159-161,414)
Interleukin 10 (*IL10*)	7	3	2	3	3	2	Anti-inflammatory cytokine	Increased expression in canine OA synovium	(414)
Interleukin 12 (*IL12B*)	4	5	4	2			Pro-inflammatory cytokine	Increased expression in human OA synovium	(413)
Interleukin 2 (*IL2*)	19				1		T and B cell proliferation	Expressed in human OA synovium	(416)
Interleukin 4 (*IL4*)	11	1	1		4	2	Anti-inflammatory cytokine	Demonstrates role in sexual dimorphisms OA susceptibility in experimental animal model	(417)
Interleukin 6 (*IL6*)	14	2	2	1		3	Pro-inflammatory cytokine	Expressed in canine OA synovium and ruptured CCL	(414)
Leptin Receptor (*LEPR*)	5			3			Adipokine receptor	Increased agonist (Leptin) expression in human OA cartilage	(418)

Name	*CFA*	*U*	*P*	*E*	*I*	*D*	*Function*	*Justification for Evaluation Inclusion*	*Reference*
Matrix Metalloproteinase 13 (*MMP13*)	5	2	6				Collagenase (cartilage break down)	Increased expression in canine hip OA cartilage and ruptured CCL	(296,319)
Matrix Metalloproteinase 3 (*MMP3*)	5				1	1	Collagenase (cartilage break down)	Increased expression in canine stifle OA	(419)
Matrix Metalloproteinase 9 (*MMP9*)	24	2			1	1	Gelatinase	Increased expression in canine hip OA cartilage and ruptured CCL	(296,319)
Secreted protein, acidic, cysteine-rich (osteonectin / *SPARC*)	4		4		3		Matrix associated protein	Increased expression in canine ruptured CCL	(319)
Tissue inhibitor of metalloproteinase 1 (*TIMP1*)	X	1				1	Inhibition of metalloproteinase activity	Increased expression in canine hip OA cartilage	(296)
Tissue inhibitor of metalloproteinase 2 (*TIMP2*)	9				1		Inhibition of metalloproteinase activity	Reduced expression in canine hip OA cartilage and ruptured CCL	(296,319)
Tissue inhibitor of metalloproteinase 3 (*TIMP3*)	10				1	1	Inhibition of metalloproteinase activity	Increased expression in human OA cartilage	(255)
Tissue inhibitor of metalloproteinase 4 (*TIMP4*)	20	2				2	Inhibition of metalloproteinase activity	Reduced expression in canine hip cartilage OA	(296)
Tenascin C (*TNC*)	11		3	2	2	1	Extracellular matrix protein	Increased expression in ruptured CCL / OA cartilage	(296,319,420)
Tumour Necrosis Factor alpha (*TNFα*)	12	2	1	1	3	2	Pro-inflammatory cytokine	Expressed by canine OA synovium	(335,421)
Zinc finger, SWIM-type containing 2 (*ZSWIM2*)	36			2	2		Metabolism	Increased expression in OA cartilage and ligament	(319)

Table 2

The significant susceptibility SNP associations. SNP Separated by breed (GR = Golden Retriever, LR = Labrador Retriever), disease (CD=cruciate disease, ED= Elbow dysplasia, HD = Hip dysplasia), gene, SNP position, minor allele frequency in cases (%D) and controls (%C), the odds ratio (OR) for the minor allele frequency in the disease population with 95% confidence intervals (95% CI), statistical significance (P value) when compared by Chi squared analysis corrected by Monte Carlo simulation, and number of cases and controls genotyped for the general dog population and the stratified (prefixed S) dog populations. Stratification was based on neuter status (neutered cases only GR CD, LR CD) or sex (male cases only, LR ED and LR HD).

Breed / Disease	*Gene (SNP ID)*	*SNP Position*	*% D*	*% C*	*OR (95% CI)*	*P value*	*Case/ Cont*	*S % D*	*S % C*	*S OR (95% CI)*	*S P value*	*S Case/ Cont*
GR CD	***IL10* (10_14R553)**	Post-Gene	23	39	0.45 (0.24 - 0.82)	**0.017**	40 / 90	25	35	0.61 (0.24 - 1.57)	0.314	26 / 17
GR CD	**IL10 (10_13Y85)**	Post-Gene	25	41	0.48 (0.26 - 0.85)	**0.017**	40 / 91	28	35	0.71 (0.28 - 1.77)	0.469	27 / 17
GR CD	***IL10* (10_1R105)**	Pre-Gene	28	42	0.54 (0.31 - 0.93)	**0.027**	43 / 93	29	38	0.69 (0.28 - 1.72)	0.478	29 / 16
GR CD	***IL10* (10_2R420)**	Prom	27	41	0.51 (0.30 - 0.89)	**0.024**	45 / 93	28	34	0.75 (0.30 - 1.90)	0.612	30 / 16
GR CD	***ANKRD10* (ANK_I1)**	Intron 1-2	30	45	0.54 (0.31 - 0.92)	**0.025**	43 / 94	36	50	0.56 (0.23 - 1.32)	0.200	28 / 17
GR CD	***ANKRD10* (ANK_I2)**	Intron 2-3	31	46	0.54 (0.32 - 0.93)	**0.043**	43 / 94	38	53	0.53 (0.22 - 1.27)	0.202	28 / 17
LR CD	***IL12B* (12B_01M115)**	Pre-Gene	27	16	1.87 (1.13 - 3.09)	**0.019**	47 / 336	27	16	1.94 (0.86 - 4.42)	0.114	24 / 53
LR CD	***IL4* (4_12M397)**	Intron 2-3	25	35	0.62 (0.38 - 0.99)	**0.050**	50 / 340	20	31	0.57 (0.25 - 1.27)	0.180	25 / 54
LR CD	***IL4* (4_13S97)**	Intron 2-3	27	18	1.65 (1.02 - 2.68)	**0.035**	50 / 331	34	19	2.13 (1.00 - 4.54)	0.073	25 / 54
LR CD	***ZSWIM2* (ZSWIM_E8B)**	Exon 8	37	26	1.66 (1.06 - 2.59)	NOT IN HWE	49 / 328	34	19	2.27 (1.06 - 4.85)	**0.046**	25 / 54
LR ED	***IL12B* (12B_01M115)**	Pre Gene	23	16	1.58 (1.03 - 2.41)	**0.049**	77 / 336	21	17	1.27 (0.73 - 2.20)	0.452	58 / 132
LR ED	***IL12B* (12B_02W232)**	Pre Gene	8	15	0.49 (0.26 - 0.92)	**0.021**	77 / 331	9	15	0.60 (0.30 - 1.22)	0.194	58 / 131
LR ED	***IL12B* (12B_02Y190)**	Pre Gene	8	14	0.50 (0.26 - 0.93)	**0.035**	78 / 330	9	15	0.61 (0.30 - 1.23)	0.187	59 / 131
LR ED	***IL12B* (12B_03Y82)**	Prom	8	14	0.54 (0.30 - 1.00)	**0.048**	78 / 328	10	14	0.68 (0.34 - 1.35)	0.310	59 / 129
LR ED	***LEPR* (LEPRB)**	Exon 18	20	26	0.69 (0.44 - 1.07)	NOT IN HWE	76 / 273	18	29	0.51 (0.29 - 0.91)	**0.024**	57 / 106
LR ED	***SPARC* (SPARC_P5)**	Prom	17	26	0.60 (0.38 - 0.93)	NOT IN HWE	78 / 329	14	27	0.47 (0.26 - 0.83)	**0.010**	59 / 130
LR ED	***TNF* (TNF10252)**	Exon 4	1	5	0.13 (0.02 - 0.97)	**0.024**	77 / 328	1	6	0.14 (0.02 - 1.09)	0.036	58 / 130

Breed / Disease	*Gene (SNP ID)*	*SNP Position*	*% D*	*% C*	*OR (95% CI)*	*P value*	*Case/ Cont*	*S % D*	*S % C*	*S OR (95% CI)*	*S P value*	*S Case/ Cont*
LR HD	***IL4* (4_13S97)**	Intron 2-3	33	18	2.18 (1.25 - 3.81)	**0.008**	32 / 331	41	17	3.32 (1.56 - 7.05)	**0.004**	17 / 132
LR HD	***IL4* (4_8R458)**	Intron 3-4	33	18	2.15 (1.23 - 3.76)	**0.008**	32 / 330	41	17	3.38 (1.59 - 7.18)	**0.005**	17 / 131
LR HD	***IL6* (6_20R191)**	Post Gene	9	21	0.40 (0.17 - 0.95)	**0.050**	32 / 329	9	20	0.39 (0.11 - 1.32)	0.154	17 / 130

Table 3

Estimated haplotype frequencies significantly associated with the development of cranial cruciate ligament rupture (CD) in Golden Retrievers (F =Frequency [%], OR = Odds ratio, 95%CI = 95% Confidence intervals). Significant associations are presented in bold text.

Gene	*Haplotypes*	*F Controls*	*F CD*	*OR CD (+95%CI)*	*P Value*
ANKRD10	AGGG	53.7	67.0	1.75 (1.04 - 2.96)	0.051
	GAGG	**24.4**	**13.6**	**0.49 (0.25 - 0.97)**	**0.040**
	GAAC	10.9	4.5	0.39 (0.13 - 1.16)	0.117
	GAAG	8.8	13.6	1.63 (0.74 - 3.58)	0.295
	Other	2.1	1.1		
	All haplotypes				**0.029**
IL10	**AGGCCATG**	**58.5**	**73.3**	**1.75 (1.04 - 2.96)**	**0.014**
	GGACGACA	**23.9**	**12.4**	**0.45 (0.22 - 0.92)**	**0.021**
	GAATCGCA	16.5	10.0	0.56 (0.26 - 1.24)	0.194
	Other	1.1	4.2		
	All Haplotypes				**0.021**

The haplotypes for each gene were as follows; *ANKRD10* SNPs; ANK_I1, ANK_I2, ANK_I4, ANK_E5, *IL10* SNPs; 10_1R105, 10_1R218, 10_2R420, 10_6Y135, 10_10S308, 10_11R124, 10_13Y85, 10_14R553.

Table 4

Estimated haplotype frequencies significantly associated with the development of cranial cruciate ligament rupture (CD), elbow dysplasia (ED) and hip dysplasia (HD) in Labrador Retrievers (F =Frequency [%], OR = Odds ratio, 95%CI = 95% Confidence intervals). Significant associations are presented in bold text.

Gene	*Haplotype*	*F C*	*F CD*	*OR CD (95%CI)*	*P Value*	*F ED*	*OR ED (95%CI)*	*P Value*	*F HD*	*OR HD (95%CI)*	*P Value*
IL4	CGAAAG	46.3	48.3	1.07 (0.71 - 1.62)	0.762	48.8	1.11 (0.79 - 1.56)	0.584	35.9	0.65 (0.38 - 1.11)	0.113
	TGCACT	**34.9**	**23.5**	**0.57 (0.35 - 0.93)**	**0.028**	30.2	0.81 (0.56 - 1.17)	0.288	31.2	0.85 (0.49 - 1.47)	0.583
	CCAGAG	**18.4**	**27.2**	**1.69 (1.05 - 2.71)**	**0.034**	20.9	1.14 (0.81 - 1.59)	0.501	**32.8**	**2.17 (1.24 - 3.78)**	**0.006**
	Other	0.5				0.0			0.0		
	All Haplotypes				0.000			0.493			**0.033**
IL6	AGAGG	48.9	41.1	0.73 (0.48 - 1.11)	0.152	48.2	0.97 (0.69 - 1.37)	0.927	53.1	1.18 (0.71 - 1.98)	0.510
	AAAGG	30.5	37.9	1.39 (0.9 - 2.14)	0.122	31.9	1.07 (0.74 - 1.54)	0.711	37.5	1.37 (0.8 - 2.32)	0.235
	AGGAA	9.1	8.8	0.97 (0.47 - 2.02)	1.000	10.9	1.23 (0.7 - 2.15)	0.431	6.3	0.67 (0.23 - 1.90)	0.528
	GGGAG	**6.8**	8.5	1.27 (0.6 - 2.71)	0.548	5.1	0.74 (0.35 - 1.58)	0.409	**0.0**	**NA**	**0.027**
	Other	4.7				3.9			3.1		
	All Haplotypes				0.465			0.736			0.117
IL12B	ATAATGGC	51.8	43.4	0.71 (0.46 - 1.08)	0.103	45.8	0.79 (0.56 - 1.11)	0.193	56.3	1.20 (0.71 - 2.01)	0.500
	ATCATAAC	16.9	17.6	1.06 (0.61 - 1.83)	0.880	22.4	1.42 (0.93 - 2.17)	0.111	21.9	1.38 (0.74 - 2.58)	0.389
	CCAATGGC	**15.8**	**27.2**	**2.03 (1.26 - 3.29)**	**0.005**	**22.9**	**1.58 (1.04 - 2.41)**	**0.034**	12.5	0.76 (0.35 - 1.65)	0.607
	ATCTCAGT	14.5	10.7	0.71 (0.37 - 1.38)	0.369	**7.4**	**0.47 (0.25 - 0.88)**	**0.015**	7.8	0.5 (0.20 - 1.28)	0.189
	Other	1.1				1.6			1.6		
	All Haplotypes				**0.023**			**0.010**			0.344

The haplotypes for each gene were as follows: *IL4*; 4_22Y152, 4_13S97, 4_12M397, 4_8R458, 4_2M351, 4_1K110, *IL6*; 6_6R431, 6_7R485, 6_20R191, 6_20R240, 6_20R412, and *IL12B*; 12B_01M115, 12B_01Y90, 12B_02Y190, 12B_02W232, 12B_02M407, 12B_03Y82, 12B_03R196, 12B_03R462.

Concluding remarks

The link between tissue based gene expression and the genomic basis to disease has been well reported in human OA research (130,216). The majority of the human candidate gene association studies have evaluated genes which are highly expressed in the diseased tissue. This approach to selecting candidate genes for canine OA was also successful in this study.

Variability is a major problem when trying to quantifing messenger RNA. Many factors can affect the measurement of gene expression such as; sample storage, method of extraction, RNA integrity (223), method of reverse transcription (422), data normalisation (238) and data processing (423). Exhaustive analysis of all these variables was beyond the scope of this thesis so only the most pertinent issues relating to this study were evaluated.

The quality and quantity of RNA extracted from canine articular tissues using two differing methods were assessed. No differences were identified, thus validating the use of a quicker (ethanol) precipitation method. Perhaps more importantly, this part of the study highlighted that RNA integrity could not be inferred from the results of spectrophotometric tests analysing RNA purity, as one would expect. Secondly, this part of the study demonstrated that loss of sample RNA integrity was a feature of canine articular tissues collected, thus stressing the importance of checking RNA quality before use. Controversy exists as to the true downstream effects of loss of "measured" RNA integrity (423). Studies have reported that loss of integrity alters the measures of gene quantity in a tissue specific, rather than gene specific manner (424),

and other studies report that loss of integrity affects individual assays rather than sample types (223). It is plausible that both are true, but in either case, integrity should be quantified. The effect of integrity on downstream assays should have been quantified as part of this study using the gold standard evaluation of integrity, the quantification of 5' and 3' gene expression (423).

Quantitative polymerase chain reaction (qPCR) methodology was optimised in a parallel project (369). A reference gene selection algorithm (238) was applied to expression data from a set of commonly used reference genes and identify those which were most stably expressed in articular cartilage and cruciate ligament. Whilst this stratified a group of reference genes for each tissue, the overall stability of these genes was inadequate (369). Furthermore one of the genes identified by the algorithm as being most stable expressed (*SDHA*) was determined to be differentially expressed in the hip OA cartilage samples (296). This highlighted two problems; firstly, evaluation of limited sample sub-set may not reflect the stability of reference genes in larger populations; and secondly, there was a need for the identification of more stable reference genes for use in studies of canine OA tissue gene expression.

The microarray data sets generated for the identification differential gene expression between sample groups provided a source of new reference genes. New genes were easily identified by filtering the data to select the most stably expressed genes, and then these were tested using multiple stability algorithms. Whilst this idea is not unique, the method for selecting the reference genes was novel. Ultimately, the theory was proved to be correct as number of reference genes identified (such as *MRSP7* and *HIRP5*) demonstrated greater stability across multiple different tissues than those we

used in the first part of the study (such as *GAPDH*) and routinely used in other studies. The new reference genes were then utilised for the evaluation of target genes in elbow cartilage and bone samples. The major element missing from this work was the hypothesis test; i.e. to definitively prove that the "stable" genes identified by the stability algorithm were truly more appropriate for normalising data sets. This is impossible to determine with tissue samples, as there may be a natural heterogeneity of expression of the target gene as well as variation in cell quantity and cell type. The hypothesis test could have been performed by measuring the gene expression of a target gene of interest in a "constant" system such as cell culture (where the cell number is known, and thus the transcript number per cell can be calculated). This work is currently being undertaken by an MSc student under the authors supervision.

A final aspect of the RT-qPCR methodology which was developed was the use of template calibrators to quantify transcript numbers. The theory behind the technique is straightforward; a synthetic oligonucleotide corresponding to the amplicon of interest and of known transcript number can be used to calibrate the threshold cycle to a transcript number, and this accounts for differences in assay performance between runs. Furthermore, absolute values of gene expression (template number) are calculated.

The future of mRNA transcript measurement requires the simultaneous quantification of cell number in each sample, so that a meaningful measure such as the average number of transcripts per cell can be determined. Using this technique target gene expression can be directly compared in different samples, tissues and experiments. The quantification of genomic DNA by qPCR would seem to be a simple and obvious

mechanism to obtain the cell number measurement. The difficulty of achieving repeatable and consistent extraction and separation of RNA and DNA from a cell population has not been fully addressed to date. Techniques have been developed, such as spiking cell preparations with novel RNA and DNA templates to account for losses during nucleic acid extraction (425) which go some way to reducing the heterogeneity of gene expression measures using this methodology. However, co-extraction techniques have not yet been developed which can be applied to solid tissue samples. Furthermore, the heterogeneity of cell types in mixed solid tissue samples such as synovium or ligament, and variation of gene expression within a single cell type such as the chondrocytes, remain major obstacles to meaningful and accurate expression profile results (426).

A screen of cartilage matrix associated genes revealed the increase in messenger RNA expression of a number of genes; collagens (*COL1A2*, *COL2A1*, *COL3A1*, *COL5A1*), other components of the extracellular matrix (*BGN*, *CSPG2*, *LUM* and *TNC*), proteases *(CTSB*, *CTSD*, *MMP13*) and a protease inhibitor (*TIMP1*), with concurrent decrease in the expression of two protease inhibitors (*TIMP2* and *TIMP4*). These finding are not surprising, as the genes selected were derived from a review of the literature of naturally occurring human OA and experimental canine OA, and the microarray study presented in Chapter 4. Thus the broad similarities identified in the patterns of gene expression of canine OA cartilage when compared with expression profiles reported for both human end-stage OA cartilage and canine experimental (stifle) OA cartilage were in some ways self fulfilling prophecies. It seems reasonable to anticipate that the major structural components of the extracellular matrix, which contribute to the material properties of articular cartilage, would be increased in

expression in the end-stage of the OA as a response to the ongoing cartilage destruction. Similarly, protease gene expression would also be predicted to be increased in end-stage OA tissue, as they are thought to be the mediators of cartilage destruction. Thus, the gene expression changes reported may simply be affects of OA, rather than the cause of the disease.

The speed, cost and sample consumption associated with RT-qPCR dictated that the list of genes evaluated by this technique was limited. The identification of a reduction in *TIMP2* and *TIMP4* gene expression were interesting findings, as they were expressed in a manner contrary to that expected on the basis of their known biological function. However the characterisation of decreased expression of protease inhibitors is not unique in canine OA cartilage, as the expression of both *TIMP1* and *TIMP4* are decreased in human OA cartilage (255).

Subsequent analysis of the expression of a matrix associated genes in OA articular cartilage of a different joint (the elbow) using the same technique, revealed similar patterns of gene expression to those reported for end-stage hip OA. The two studies in this thesis are not truly comparable because the elbow cartilage gene expression profiles were generated from cases with ranges of severity of OA (rather than all being of end-stage), the target gene transcript numbers were calculated using template oligonucleotide calibrators, and the data were normalised using the new reference genes identified from microarray data.

Correlation of gene expression of selected genes with the radiographic measure of OA severity demonstrated that molecular measures of disease activity and associated

tissue response could be directly related to a clinical assessment of OA. These agreements must be interpreted with caution, because of the sample stratification and the generalised measure of clinical disease. However they do highlight a potential link between non-invasive measures of disease severity and molecular changes in the joint. Surprisingly, the relationship between clinical measures and gene expression have not been studied commonly in human OA research (318), and the finding presented in this thesis warrant further investigation. Clearly, the canine elbow joint can ethically provide an easily recovered source of tissue from patients with multiple grades of OA, unlike tissue taken at joint replacement surgery. Larger scale analyses with global measures of gene expression, greater numbers of samples, and more meaningful measures of disease severity, such as histological scores could provide a greater insight into the early molecular mechanisms involved in the pathogenesis of canine elbow OA.

Analysis of genome wide expression of canine OA cartilage by canine specific whole genome oligonucleotide microarrays was less definitive. The gene expression profiles were extremely heterogeneous, which made meaningful interpretation of the disease data extremely difficult. Ultimately data could only be evaluated without correction for multiple hypothesis testing, which resulted in the identification of a large number of differentially expressed genes, a significant number of which could have been false positives. The data produced is likely to be true reflection of the genes expressed in the diseased tissue (and this was reflected by the RT-qPCR results of the same samples), but evaluation of such a small number of diseased samples without other phenotypic criteria led to the inclusion of samples which probably demonstrated markedly different histological grades of OA and wide variety of cell phenotypes.

Although samples were collected in a standardised manner, the absence of additional phenotypic selection, such as the use of a histological OA score (284) may have led to the increased variability. A major limitation of investigating hip cartilage was that tissues could only be procured in an ethical manner from the joints with end-stage disease (i.e. at joint replacement).

The results of the RT-qPCR expression profiles from a much larger set of hip OA cartilage samples also demonstrated a moderate degree of heterogeneity. Thus, the variation in gene expression was probably not just restricted to microarray analysis, but because smaller sample numbers were analysed on microarrays the downstream effect on results was greater. Interestingly, the published studies of gene expression in histologically phenotyped human OA cartilage using microarray analysis of small sample numbers (<10 in disease and control groups) (248,300) have not been robustly corrected for multiple hypothesis testing. The only study published to date using microarray analysis followed a similar analysis protocol (with correction for multiple hypothesis testing) to that which we describe, required the analysis of large numbers of samples (n = 78) to achieve consistent patterns of expression (249). Thus, one may conclude that gene expression in clinical articular cartilage specimens is by its nature heterogeneous, and that this may not be an ideal tissue for analysis by microarray even though it contains a single cell type.

Traditionally, OA research has tended to focus on articular cartilage. In this study, we evaluated gene expression in other articular tissues (bone and ligament) which are involved in OA. In a parallel project conceived, planned and supervised by the author, multiple articular tissues (fat, cartilage, ligament, and synovium) were analysed for

the expression of important cytokines which mediate OA. Interestingly, the two primary sources of pro-inflammatory cytokines (*IL1* and *IL6*) identified were the synovium and ruptured cranial cruciate ligament (a tissue, which is itself covered by synovium), although various degrees of cytokine expression were detected in all tissues (414). Clearly, OA is not simply a disease of cartilage, and it would be prudent to evaluate all articular tissues for their involvement in the development and progression of the OA process.

The gene expression profile changes identified in OA bone generally mirror those observed in OA cartilage, with increased expression of the primary structural component of the extracellular matrix (*COL1*), increased expression of matrix metalloproteinases (*MMP2, -9, -13*) and one of their inhibitors (*TIMP1*), and concurrent with decrease in expression of another inhibitor (*TIMP2*). The same caveats apply for OA bone as for cartilage; namely that histological assessment of the tissue probably should have been performed concurrently with the measures of gene expression. The characterisation of gene expression in OA bone merits further analysis using microarray.

The use of canine whole genome microarray analysis to quantify gene expression in the cranial cruciate ligament yielded a large number of differential expressed genes even after correction for multiple hypothesis testing, in contrast to cartilage. This implies that the transcriptome of the ruptured CCL tissue from different individuals is relatively homogeneous in contrast to the large variety of gene expression seen between different cartilage samples with end stage OA. The general pattern of gene expression identified in ruptured CCL was broadly similar to that previously

described for articular cartilage; namely up-regulation of extracellular matrix components and proteases, with concurrent decreased expression of a protease inhibitor (*TIMP2*).

The expression profiles of normal and ruptured canine cranial cruciate ligaments yielded a number of important findings. Firstly, a large number of genes were differentially expressed, even after correcting for multiple hypothesis testing. A number of structural elements, such as *COL1A2* and *COL3A1* demonstrated increased gene expression, which was consistent with attempts at ligament repair. The marked degradation of the extracellular matrix noted in the ruptured CCL was characterised by the increased expression of a number of proteases, such as *MMP2, MMP9, CTSB, CTSD* and *CASP8*. Increased expression of *SPARC,* a gene involved in ligament development and repair, was of particular interest. Polymorphisms of this gene were therefore analysed in the genomic part of this study.

A larger number of genes were identified as being differentially expressed by RT-qPCR analysis of ruptured CCL or OA cartilage samples than by were identified microarray analysis. Although the use of small numbers of microarray analyses for transcriptome profiling identified a number of differentially expressed genes, this methodology (with limited sample numbers) does not allow the elucidation of all the finer nuances of gene transcription. Whether these can be ascertained from more controlled *in vitro* cell culture experiments remains to be seen.

A transcriptomic basis to the breed risk of CCL rupture was not determined. However, the expression profiles of a breed at low-risk of CCL rupture tended to

cluster more closely with the profiles generated from ruptured CCL. The number of genes with a significant change in expression between normal low-risk CCL and the high-risk ruptured CCL were significantly greater than would be expected by random chance. The implications are that the genetic basis to CCL rupture is present, but that our methods, or sample size, were not sufficiently sensitive to identify them. A further inference is that the anabolic response identified in CCL rupture may be present in the CCL of dogs with low-risk of rupture when compared to dogs with high-risk of rupture, and this helps to reduce risk of rupture. This theory concurs with the increased mechanical strength of the normal CCL reported in breeds at low-risk of CCL rupture when compared to breeds at high risk of rupture.

One of the major challenges in the field of transcriptomics is the ability to interpret the large quantities of transcriptomic data generated by microarray experiments. The work in this thesis utilised standard analysis techniques; namely statistical analysis of the whole data set, correction for multiple hypothesis testing and clustering of differentially expressed genes. Data analysis tools which can sort gene expression profiles on the basis of known interaction between genes within different biochemical pathways (pathway analysis) (427), and apply global literature searches (text mining) to facilitate the identification of the true interactions (428) are well described. Attempts were made to try and annotate locate differentially expressed genes to common biological pathways using a web-based pathway analysis tool (302). Ultimately this method of pathway analysis was found to be laborious and unrewarding as most of the genes had not been defined onto pathways. The major problem encountered was that the microarray platform we used was custom designed and contained gene identifiers which were unique, and therefore could not be

automatically linked into pathway analysis tools. A further issue identified was that many of the gene annotations provided with the microarray were sometimes incorrect. As a result all array spots reported as being differentially expressed had to be manually checked using the basic local alignment search tool to verify the true annotation of the sequence, which was extremely time consuming.

Tissue inhibitor of matrix metalloproteinase 2 is a major inhibitor of *MMP2* activity in the dog (429) although it can inhibit all MMPs (to a degree) and is constitutively expressed in articular cartilage (430). The consistent identification of decreased expression of *TIMP2* in OA tissues (hip and elbow articular cartilage, and ruptured cranial cruciate ligament) is perhaps the most notable feature of the work presented in this thesis. Review of the literature published on this gene reveals that the expression of *TIMP2* has been repeatedly identified to be decreased in canine OA cartilage (327) and in canine cartilage affected by other canine joint diseases such as osteochondrosis (341). Furthermore, this decrease is also identified at the protein level in synovial fluid of dogs with stifle OA (340).

The need to compare and contrast different methods of extracting genomic DNA (gDNA) arose as we noted varying intensities of the PCR bands after the first PCR reaction used in the genotyping project and on average 6% of genotyping tests failed. All gDNA samples were extracted using the phenol-chloroform method from clotted or EDTA blood and normalised on the basis of DNA measurement using a spectrophotometer. It appeared that for certain samples the quantity of gDNA measured was inadequate, or contained PCR inhibitors which reduced genotyping efficiency. Quantification of the degree of PCR inhibition of gDNA confirmed that

the spectrophotometric results underestimated the quantity of gDNA present in samples extracted from blood clots using the phenol-chloroform method, as determined using a qPCR technique.

The twenty candidate genes selected for analysis in the genomic part of the study were selected by a variety of means. Genes demonstrating differential expression in canine OA tissues (as reported in this study), human OA tissues (as reported in other studies), or demonstrating a good functional basis to be involved in the pathogenesis of OA were selected for evaluation. Initially, the optimal method for detecting SNPs was determined by comparing an *in vitro* SNP detection method with two *in silico* SNP detection methods (appendix 2). *In vitro* SNP detection was determined to be the best method as it identified nearly twice the number of SNPs which were available in the *in silico* databases, and nearly two thirds of the SNPs identified by this method were unique (not publicly available). These findings were not surprising as the canine genome has only been published relatively recently and is based primarily on the sequencing results of a single dog. The direct comparison with sequencing results of 93 dogs of multiple different breeds was likely to identify a large number of new mutations.

The rationale for candidate gene selection was similar to that used by Valdes and others (130), with selection of SNPs in genes which are differentially expressed in OA tissues. The majority of genes had a robust theoretical basis for their inclusion on their known functions in the OA joint (such as *IL1* or *MMP13*). A number of other genes, such as *ATP11B* and *ANKRD10* were selected on the basis of marked differential gene expression in canine OA tissues as determined by the microarray studies, but were

lacking any meaningful annotation of gene function. Seven candidate genes (*ATP11B, ANKRD10, SPARC, TIMP2, TIMP4, TNC* and *ZSWIM2*) were selected for *in vitro* SNP identification. Further SNPs were identified from publicly available sources for additional candidate genes (*MMP3, TIMP1* and *TIMP3*), or from previous studies (*IL1α, IL2, IL4, IL6, IL10, IL12B* and *TNFα* identified by Dr Andrea Short (401), and *MMP9, MMP13* and *LEPR* identified by Mrs Annette Barnes, personal communication).

Significant associations were identified between both the minor allele and haplotype frequencies of a number of candidate genes and the susceptibility to articular disease in dogs. Of particular interest were the association of mutations in the *IL4* gene with the risk of both hip dysplasia and CCL rupture, and the association of mutations in *IL12B* with elbow dysplasia and CCL rupture in Labrador Retrievers. These findings can be interpreted in two ways. Firstly, they may suggest that canine articular diseases do have a common pathogenesis within an individual breed. The association may be with the primary articular disease (for example, joint laxity which is a feature of CCL rupture and HD) or the OA which develops concurrently. Our disease cohorts were more likely to contain dogs with a greater risk of OA because the cases represented the extreme of the phenotype (i.e. they required articular surgery for their articular disease).

Alternatively, the data may imply that a proportion of each disease cohorts could have gone on to develop a second articular disease (for example, Labrador Retrievers with elbow dysplasia carrying the "risk" *IL12B* haplotype could have gone on to develop CCL rupture in later life which was not evident at the time of DNA collection- thus

the haplotype is a risk for one disease only). Although we cannot resolve this dilemma without further studies, this interpretation should have been made less likely by the fact that we used a general dog population of dogs as controls in this study. These controls were not phenotyped, and thus the prevalence of the disease in the control population should have made it more difficult to detect a spurious association.

The second finding of particular interest was the observation that some of the genes associated with disease (*IL4* SNPs with hip dysplasia and Labrador Retrievers, and *IL10* SNPs with CCL rupture in Golden Retrievers) mirrored similar associations for human disease (human hip and knee OA respectively). This finding is not unique, and has been reported with other canine diseases such as diabetes mellitus (401), which highlights both the value of the naturally occurring disease in the domestic dog as a model for human disease, and the commonality of candidate genes used to investigate disease in both species.

The final finding of particular interest was that the minor allele frequencies of two SNPs in *IL4* associated with the risk of hip dysplasia in Labrador Retrievers had been previously associated with a the risk of different disease (diabetes mellitus) in a different breed (Cairn Terriers) (401). Further genotyping of candidate genes close to *IL4* and investigation of the functional significance of these SNPs would be of great interest in both disease and breed cohorts. Conversely, common genetic risks were not identified for the same disease (CCL rupture) in two different breeds of dog (Labrador Retrievers and Golden Retrievers).

The candidate gene with the strongest evidence for investigation in canine OA, *TIMP2* did not have any significant associations between polymorphisms and disease status. There are a number of potential reasons for this. Firstly, the precise location of Exon 1, as determined using cloning studies (431) has not been annotated to the canine genome (277), thus polymorphisms in the promoter region and exon 1 could not be identified. Secondly, only a single SNP in this gene was evaluated. Future work would more logically concentrate on analysing the genomic polymorphisms of genes involved in *TIMP2* expression. Although these pathways have not been defined in the dog, human data have defined the *ERK1/2* and *p38 MAPK* pathways as being involved in *TIMP2* expression (432). Indeed, these pathways are of particular interest as they also control *TGFβ1* and mediate the expression of *MMP2, MMP13* and *MMP14* in a reciprocal manner to that of *TIMP2* (432,433).

The majority of associations we identified were associated with relatively small changes in the risk or protection from disease. In part this is a reflection of the quality of the control population used in this study, and the nature of complex genetic traits. With idealised controls and larger cohorts these associations may have been increased both the difference between frequencies of disease and control cohorts, and the significance and power of the findings.

The genomic gene expression study can be criticized from a number of aspects. The premise of the study, although proven to be successful, is flawed. Genes whose expression is changed in expression in diseased tissues are probably more likely the result rather than the cause of the pathological process when evaluated in end-stage disease tissue. With time, the resolution of transcriptomic technologies may allow the

gene expression pathways to be identified and dissected to more precisely identify the truly causative defective gene(s). However the full understanding the complexity of interactions involved is an enormous challenge for molecular biology, and is compounded by the heterogeneity of cell type and morphology in both normal and disease clinical tissues. The size of the disease cohorts, the lack of phenotyping of controls, the limited number of SNPs evaluated in certain candidate genes were all factors which limit the usefulness of this study.

The logical conclusions to the study were two-fold. Firstly the transcriptomic changes in canine tissues are broadly similar between different articular diseases, and may be similar to those reported in the equivalent human disease. Secondly, a genomic basis to canine OA, or articular disease *per se*, does exist. The genotyping study requires repeating with new, larger cohorts using better phenotyped controls. The greatest challenge is how these cohorts can be collected from canine populations in an ethical and economic manner. The rapid advance of genomic technologies, even during the period of this study, dictates that further studies of association will be performed using genome-wide SNP arrays. However the strong linkage disequilibrium identified in dog breeds will likely result in large haplotype blocks in which the causative mutations may not easily be identified. The race to characterise the genetic basis to OA in mammalian species is only just starting.

Reference List

1. Dingle, J.T. (1981) Catabolin: A cartilage catabolic factor from synovium. *Clinical Orthopaedics and Related Research* **156**, 219-231

2. Ushiyama, T., Chano, T., Inoue, K., Matsusue, Y. (2003) Cytokine production in the infrapatellar fat pad: another source of cytokines in knee synovial fluids. *Annals of the Rheumatic Diseases* **62**, 108-112

3. Hayashi, K., Frank, J.D., Dubinsky, C., Hao, Z., Markel, M.D., Manley, P.A., Muir, P. (2003) Histologic Changes in Ruptured Canine Cranial Cruciate Ligament. *Veterinary Surgery* **32**, 269-277

4. Rogers, J., Shepstone, L., Dieppe, P. (2004) Is osteoarthritis a systemic disorder of bone? *Arthritis and Rheumatism* **50**, 452-457

5. Brandt, K.D., Myers, S.L., Burr, D., Albrecht, M. (1991) Osteoarthritic changes in canine articular cartilage, subchondral bone, and synovium fifty-four months after transection of the anterior cruciate ligament. *Arthritis and Rheumatism* **34**, 1560-1570

6. Ness, M.G., Abercromby, R.H., May, C., Turner, B.M., Carmichael, S. (1996) A survey of orthopaedic conditions in small animal veterinary practice in Britain. *Veterinary and Comparative Orthopaedics and Traumatology* **9**, 43-45

7. Paradis, M., Sauve, F., Charest, J., Refsal, K.R., Moreau, M., Dupuis, J. (2003) Effects of moderate to severe osteoarthritis on canine thyroid function. *Candian Veterinary Journal* **44**, 407-412

8. Todhunter, R.J., Lust, G. (2002) Hip dysplasia: Pathogenesis. In: Textbook of small animal surgery. Editor(s): D.H.Slatter. 3rd Edn. WB Saunders Co, London. p.2009-2019

9. Grondalen, J., Grondalen, T. (1981) Arthrosis in the elbow joint of young rapidly growing dogs. V. A pathoanatomical investigation. *Nordisk Veterinaermedicin* **33**, 1-16

10. Bradney, I.W. (1967) Non-union of the anconeal process in the dog. *Australian Veterinary Journal* **43**, 215-216

11. Bennett, D., Tennant, B., Lewis, D.G., Baughan, J., May, C., Carter, S. (1988) A reappraisal of anterior cruciate ligament disease in the dog. *Journal of Small Animal Practice* **29**, 275-297

12. Smith, G.K., Mayhew, P.D., Kapatkin, A.S., McKelvie, P.J., Shofer, F.S., Gregor, T.P. (2001) Evaluation of risk factors for degenerative joint disease associated with hip dysplasia in German Shepherd Dogs, Golden Retrievers, Labrador Retrievers, and Rottweilers. *Journal of the American Veterinary Medical Association* **219**, 1719-1724

13. Buckwalter, J.A., Saltzman, C., Brown, T. (2004) The impact of osteoarthritis: implications for research. *Clinical Orthopaedics and Related Research* **427**, S6-15

14. Meulenbelt, I., Seymour, A.B., Nieuwland, M., Huizinga, T.W., van Duijn, C.M., Slagboom, P.E. (2004) Association of the interleukin-1 gene cluster with radiographic signs of osteoarthritis of the hip. *Arthritis and Rheumatism* **50**, 1179-1186

15. Doherty, M., Watt, I., Dieppe, P. (1983) Influence of primary generalised osteoarthritis on development of secondary osteoarthritis. *Lancet* **2**, 8-11

16. Burton-Wurster, N., Farese, J.P., Todhunter, R.J., Lust, G. (1999) Site-specific variation in femoral head cartilage composition in dogs at high and low risk for development of osteoarthritis: insights into cartilage degeneration. *Osteoarthritis and Cartilage* **7**, 486-497

17. Chalmers, H.J., Dykes, N.L., Lust, G., Farese, J.P., Burton-Wurster, N.I., Williams, A.J., Todhunter, R.J. (2006) Assessment of bone mineral density of the femoral head in dogs with early osteoarthritis. *American Journal of Veterinary Research* **67**, 796-800

18. Lust, G., Summers, B.A. (1981) Early, asymptomatic stage of degenerative joint disease in canine hip joints. *American Journal of Veterinary Research* **42**, 1849-1855

19. Madsen, J.S. (1997) The joint capsule and joint laxity in dogs with hip dysplasia. *Journal of the American Veterinary Medical Association* **210**, 1463-1465

20. Cardinet, G.H., III, Kass, P.H., Wallace, L.J., Guffy, M.M. (1997) Association between pelvic muscle mass and canine hip dysplasia. *Journal of the American Veterinary Medical Association* **210**, 1466-1473

21. Dassler, C.L. (2002) Canine hip dysplasia: Diagnosis and non-surgical treatment. In: Textbook of small animal surgery. Editor(s): D.H.Slatter. 3rd Edn. WB Sauders Co., London. p.2019-2029

22. Smith, G.K., Paster, E.R., Powers, M.Y., Lawler, D.F., Biery, D.N., Shofer, F.S., McKelvie, P.J., Kealy, R.D. (2006) Lifelong diet restriction and radiographic evidence of osteoarthritis of the hip joint in dogs. *Journal of the American Veterinary Medical Association* **229**, 690-693

23. Kealy, R.D., Olsson, S.E., Monti, K.L., Lawler, D.F., Biery, D.N., Helms, R.W., Lust, G., Smith, G.K. (1992) Effects of limited food consumption on the incidence of hip dysplasia in growing dogs. *Journal of the American Veterinary Medical Association* **201**, 857-863

24. Leighton, E.A. (1997) Genetics of canine hip dysplasia. *Journal of the American Veterinary Medical Association* **210**, 1474-1479

25. Goldsmith, L.T., Lust, G., Steinetz, B.G. (1994) Transmission of relaxin from lactating bitches to their offspring via suckling. *Biology of Reproduction* **50**, 258-265

26. Jones, D.G. (1996) BVA/Kennel Club hip dysplasia scheme. *Veterinary Record* **139**, 195

27. Corley, E.A. (1992) Role of the Orthopedic Foundation for Animals in the control of canine hip dysplasia. *Veterinary Clinics of North America Small Animal Practice* **22**, 579-593

28. Fluckiger, M.A., Friedrich, G.A., Binder, H. (1999) A radiographic stress technique for evaluation of coxofemoral joint laxity in dogs. *Veterinary Surgery* **28**, 1-9

29. Smith, G.K., Gregor, T.P., Rhodes, W.H., Biery, D.N. (1993) Coxofemoral joint laxity from distraction radiography and its contemporaneous and prospective correlation with laxity, subjective score, and evidence of degenerative joint disease from conventional hip-extended radiography in dogs. *American Journal of Veterinary Research* **54**, 1021-1042

30. Farese, J.P., Todhunter, R.J., Lust, G., Williams, A.J., Dykes, N.L. (1998) Dorsolateral subluxation of hip joints in dogs measured in a weight-bearing position with radiography and computed tomography. *Veterinary Surgery* **27**, 393-405

31. Todhunter, R.J., Grohn, Y.T., Bliss, S.P., Wilfand, A., Williams, A.J., Vernier-Singer, M., Burton-Wurster, N.I., Dykes, N.L., Wu, R., Casella, G., Acland, G.M., Lust, G. (2003) Evaluation of multiple radiographic predictors of cartilage lesions in the hip joints of eight-month-old dogs. *American Journal of Veterinary Research* **64**, 1472-1478

32. Wood, J.L., Lakhani, K.H. (2003) Hip dysplasia in labrador retrievers: the effects of age at scoring. *Veterinary Record* **152**, 37-40

33. Nielen, A.L., Knol, B.W., van Hagen, M.A., van-der-Gaag, I. (2003) Genetic and epidemiological investigation of a birth cohort of boxers. *Tijdschrift Voor Diergeneeskunde* **128**, 586-590

34. Morgan, J.P., Wind, A., Davidson, A.P. (1999) Bone dysplasias in the Labrador Retriever: a radiographic study. *Journal of the American Animal Hospital Association* **35**, 332-340

35. Kaneene, J.B., Mostosky, U.V., Padgett, G.A. (1997) Retrospective cohort study of changes in hip joint phenotype of dogs in the United States. *Journal of the American Veterinary Medical Association* **211**, 1542-1544

36. Paster, E.R., LaFond, E., Biery, D.N., Iriye, A., Gregor, T.P., Shofer, F.S., Smith, G.K. (2005) Estimates of prevalence of hip dysplasia in Golden Retrievers and Rottweilers and the influence of bias on published prevalence figures. *Journal of the American Veterinary Medical Association* **226**, 387-392

37. Dutton, R.E., Moore, G.E. (1987) Clinical review of death/euthanasia in 123 military working dog necropsies. *Military Medicine* **152**, 489-493

38. Olson, R.C. (1971) Physical evaluation and selection of military dogs. *Journal of the American Veterinary Medical Association* **159**, 1444-1446

39. Johnston, S.A. (1992) Conservative and medical management of hip dysplasia. *Veterinary Clinics of North America Small Animal Practice* **22**, 595-606

40. Olmstead, M.L., Hohn, R.B., Turner, T.M. (1983) A five-year study of 221 total hip replacements in the dog. *Journal of the American Veterinary Medical Association* **183**, 191-194

41. Duff, R., Campbell, J.R. (1977) Long term results of excision arthroplasty of the canine hip. *Veterinary Record* **101**, 181-184

42. Jacobsen, S., Sonne-Holm, S. (2005) Hip dysplasia: a significant risk factor for the development of hip osteoarthritis. A cross-sectional survey. *Rheumatology* **44**, 211-218

43. Sollazzo, V., Bertolani, G., Calzolari, E., Atti, G., Scapoli, C. (2000) A two-locus model for non-syndromic congenital dysplasia of the hip (CDH). *Annals of Human Genetics* **64**, 51-59

44. Wynne-Davies, R. (1970) Acetabular Dysplasia and Familial Joint Laxity: Two Etiological Factors in Congenital Dislocation of the Hip: A Review of 589 Patients and Their Families. *Journal of Bone and Joint Surgery - British Volume* **52-B**, 704-716

45. Granchi, D., Stea, S., Sudanese, A., Toni, A., Baldini, N., Giunti, A. (2002) Association of two gene polymorphisms with osteoarthritis secondary to hip dysplasia. *Clinical Orthopaedics and Related Research* **403**, 108-117

46. Clements, D.N., Carter, S.D., Innes, J.F., Ollier, W.E. (2006) Genetic basis of secondary osteoarthritis in dogs with joint dysplasia. *American Journal of Veterinary Research* **67**, 909-918

47. Priester, W.A., Mulvihill, J.J. (1972) Canine hip dysplasia: relative risk by sex, size, and breed, and comparative aspects. *Journal of the American Veterinary Medical Association* **160**, 735-739

48. LaFond, E., Breur, G.J., Austin, C.C. (2002) Breed Susceptibility for Developmental Orthopedic Diseases in Dogs. *Journal of the American Animal Hospital Association* **38**, 467-477

49. Janutta, V., Hamann, H., Klein, S., Tellhelm, B., Distl, O. (2005) Genetic Evaluation of Elbow Angles as Predictors of Elbow Dysplasia in German Shepherd Dogs. *Journal of Veterinary Medicine Series A* **52**, 254-261

50. Wood, J.L.N., Lakhani, K.H., Dennis, R. (2000) Heritability and epidemiology of canine hip-dysplasia score in flat-coated retrievers and

Newfoundlands in the United Kingdom. *Preventive Veterinary Medicine* **46**, 75-86

51. Reed, A.L., Keller, G.G., Vogt, D.W., Ellersieck, M.R., Corley, E.A. (2000) Effect of dam and sire qualitative hip conformation scores on progeny hip conformation. *Journal of the American Veterinary Medical Association* **217**, 675-680

52. Wood, J.L.N., Lakhani, K.H., Rogers, K. (2002) Heritability and epidemiology of canine hip-dysplasia score and its components in Labrador retrievers in the United Kingdom. *Preventive Veterinary Medicine* **55**, 95-108

53. Todhunter, R.J., Bliss, S.P., Casella, G., Wu, R., Lust, G., Burton-Wurster, N.I., Williams, A.J., Gilbert, R.O., Acland, G.M. (2003) Genetic Structure of Susceptibility Traits for Hip Dysplasia and Microsatellite Informativeness of an Outcrossed Canine Pedigree. *Journal of Heredity* **94**, 39-48

54. Houlton, J.E.F. (2005) BVA/KC hip and elbow dysplasia schemes. *The Veterinary Record* **157**, 853

55. Swenson, L., Audell, L., Hedhammar, A. (1997) Prevalence and inheritance of and selection for hip dysplasia in seven breeds of dogs in Sweden and benefit: cost analysis of a screening and control program. *Journal of the American Veterinary Medical Association* **210**, 207-214

56. Chase, K., Lawler, D.F., Adler, F.R., Ostrander, E.A., Lark, K.G. (2004) Bilaterally asymmetric effects of quantitative trait loci (QTLs): QTLs that affect laxity in the right versus left coxofemoral (hip) joints of the dog (Canis familiaris). *American Journal of Medical Genetics* **124**, 239-247

57. Todhunter, R., Mateescu, R., Lust, G., Burton-Wurster, N., Dykes, N., Bliss, S., Williams, A., Vernier-Singer, M., Corey, E., Harjes, C., Quaas, R., Zhang, Z., Gilbert, R., Volkman, D., Casella, G., Wu, R., Acland, G. (2005) Quantitative trait loci for hip dysplasia in a crossbreed canine pedigree. *Mammalian Genome* **16**, 720-730

58. Smith, G.K., Popovitch, C.A., Gregor, T.P., Shofer, F.S. (1995) Evaluation of risk factors for degenerative joint disease associated with hip dysplasia in dogs. *Journal of the American Veterinary Medical Association* **206**, 642-647

59. Hays, L., Zhang, Z., Mateescu, R.G., Lust, G., Burton-Wurster, N.I., Todhunter, R.J. (2007) Quantitative genetics of secondary hip joint osteoarthritis in a Labrador Retriever-Greyhound pedigree. *American Journal of Veterinary Research* **68**, 35-41

60. Chase, K., Lawler, D.F., Carrier, D.R., Lark, K.G. (2005) Genetic regulation of osteoarthritis: A QTL regulating cranial and caudal acetabular osteophyte formation in the hip joint of the dog (Canis familiaris). *American Journal of Medical Genetics* **135**, 334-335

61. Mateescu, R., Burton-Wurster, N.I., Lust, G., Tsai, K., Phavaphutanon, J., Todhunter, R.J. (2006) Quantitative trait loci for osteoarthritis in dogs. In: Proceedings of the 52nd Annual Meeting of the Orthopaedic Research Society p.1546

62. Meyer-Lindenberg, A., Fehr, M., Nolte, I. (2006) Co-existence of ununited anconeal process and fragmented medial coronoid process of the ulna in the dog. *Journal of Small Animal Practice* **47**, 61-65

63. Bennett, D., Duff, S.R., Kene, R.O., Lee, R. (1981) Osteochondritis dissecans and fragmentation of the coronoid process in the elbow joint of the dog. *Veterinary Record* **109**, 329-336

64. Remy, D., Neuhart, L., Fau, D., Genevois, J.P. (2004) Canine elbow dysplasia and primary lesions in German shepherd dogs in France. *Journal of Small Animal Practice* **45**, 244-248

65. Van-Ryssen, B., van-Bree, H. (1997) Arthroscopic findings in 100 dogs with elbow lameness. *Veterinary Record* **140**, 360-362

66. Read, R.A., Robins, G.M. (1982) Deformity of the proximal tibia in dogs. *Veterinary Record* **111**, 295-298

67. Kramer, A., Holsworth, I.G., Wisner, E.R., Kass, P.H., Schulz, K.S. (2006) Computed tomographic evaluation of canine radioulnar incongruence in vivo. *Veterinary Surgery* **35**, 24-29

68. Hare, W.C.D. (1961) The ages at which the centers of ossification appear roentgenographically in the limb bones of the dog. *American Journal of Veterinary Research* **22**, 825-835

69. Wolschrijn, C.F., Gruys, E., Weijs, W.A. (2005) Microcomputed tomography and histology of a fragmented medial coronoid process in a 20-week-old golden retriever. *The Veterinary Record* **157**, 383-386

70. Danielson, K.C., Fitzpatrick, N., Muir, P., Manley, P.A. (2006) Histomorphometry of Fragmented Medial Coronoid Process in Dogs: A Comparison of Affected and Normal Coronoid Processes. *Veterinary Surgery* **35**, 501-509

71. Guthrie, S., Plummer, J.M., Vaughan, L.C. (1992) Aetiopathogenesis of canine elbow osteochondrosis: a study of loose fragments removed at arthrotomy. *Research in Veterinary Science* **52**, 284-291

72. Yovich, J.C., Read, R.A. (1994) Traumatic fracture of the medial coronoid process in 2 dogs. *Veterinary and Comparative Orthopaedics and Traumatology* **7**, 173-176

73. Guthrie, S., Pidduck, H.G. (1990) Heritability of elbow osteoachondrosis within a closed population of dogs. *Journal of Small Animal Practice* **31**, 93-96

74. Clements, D.N., England, G.C.W., Clarke, S.P., Gemmill, T., Bennett, D., Carmichael, S. (2004) Owner assessment of the development of common orthopaedic disorders in Labrador Retrievers. In: Proceedings of the British Small Animal Veterinary Association Congress p.577

75. Ubbink, G.J., van de Broek, J., Hazewinkel, H.A., Wolvekamp, W.T., Rothuizen, J. (2000) Prediction of the genetic risk for fragmented coronoid process in labrador retrievers. *The Veterinary Record* **147**, 149-152

76. Lang, J., Busato, A., Baumgartner, D., Fluckiger, M. (1995) Comparison of two classification protocols in the evaluation of elbow disease in the Bernese Mountain Dog. In; Abstracts of Papers Presented at the Annual Meeting of the International Elbow Working Group (IEWG), Philadelphia, PA, August 7, 1994. *Veterinary Radiology & Ultrasound* **36**, 170-171

77. Padgett, G.A., Mostosky, U.V., Probst, C.W., Thomas, M.W., Krecke, C.F. (1995) The inheritance of osteochondritis dissecans and fragmented coronoid process of the elbow joint in labrador retrievers. *Journal of the American Animal Hospital Association* **31**, 327-330

78. Grondalen, J., Lingaas, F. (1991) Arthrosis in the elbow joint of young rapidly growing dogs - A genetic investigation. *Journal of Small Animal Practice* **32**, 460-464

79. Guthrie, S. (1989) Some radiographic and clinical aspects of ununited anconeal process. *Veterinary Record* **124**, 661-662

80. Maki, K., Liinamo, A.E., Ojala, M. (2000) Estimates of genetic parameters for hip and elbow dysplasia in Finnish Rottweilers. *Journal of Animal Science* **78**, 1141-1148

81. Studdert, V.P., Lavelle, R.B., Beilharz, R.G., Mason, T.A. (1991) Clinical-features and heritability of osteochodrosis of the elbow in Labrador retrievers. *Journal of Small Animal Practice* **32**, 557-563

82. Janutta, V., Hamann, H., Klein, S., Tellhelm, B., Distl, O. (2006) Genetic analysis of three different classification protocols for the evaluation of elbow dysplasia in German shepherd dogs. *Journal of Small Animal Practice* **47**

83. Salg, K.G., Temwitchitr, J., Imholz, S., Hazewinkel, H.A., Leegwater, P.A. (2006) Assessment of collagen genes involved in fragmented medial coronoid process development in Labrador Retrievers as determined by affected sibling-pair analysis. *American Journal of Veterinary Research* **67**, 1713-1718

84. Read, R.A., Armstrong, S.J., O'Keefe, J.D., Eger, C.E. (1990) Fragmentation of the medial coronoid process of the ulna in dogs - A study of 109 cases. *Journal of Small Animal Practice* **31**, 330-334

85. Pond, M.J., Nuki, G. (1973) Experimentally-induced osteoarthritis in the dog. *Annals of the Rheumatic Diseases* **32**, 387-388

86. Innes, J.F., Costello, M., Barr, F.J., Rudorf, H., Barr, A.R.S. (2004) Radiographic Progression of Osteoarthritis of the Canine Stifle Joint: A Prospective Study. *Vet Radiol Ultrasound* **45**, 143-148

87. Doverspike, M., Vasseur, P.B., Harb, M.F., Walls, C.M. (1993) Contra-lateral cranial cruciate ligament rupture: incidence in 114 dogs. *Journal of the American Animal Hospital Association* **29**, 167-170

88. Whitehair, J.G., Vasseur, P.B., Willits, N.H. (1993) Epidemiology of cranial cruciate ligament rupture in dogs. *Journal of the American Veterinary Medical Association* **203**, 1016-1019

89. Arnoczky, S.P., Marshall, J.L. (1977) The cruciate ligaments of the canine stifle: an anatomical and functional analysis. *American Journal of Veterinary Research* **38**, 1807-1814

90. Monahan, J.J., Grigg, P., Pappas, A.M., Leclair, W.J., Marks, T., Fowler, D.P., Sullivan, T.J. (1984) In vivo strain patterns in the four major canine knee ligaments. *Journal of Orthopaedic Research* **2**, 408-418

91. Comerford, E.J., Tarlton, J.F., Innes, J.F., Johnson, K.A., Amis, A.A., Bailey, A.J. (2005) Metabolism and composition of the canine anterior cruciate ligament relate to differences in knee joint mechanics and predisposition to ligament rupture. *Journal of Orthopaedic Research* **23**, 61-66

92. Wilke, V.L., Conzemius, M.G., Kinghorn, B.P., Mcrossan, P.E., Weiguo, C., Rothschild, M.F. (2006) Inheritance of rupture of the cranial cruciate ligament in Newfoundlands. *Journal of the American Veterinary Medical Association* **228**, 61-64

93. Vasseur, P.B., Pool, R.R., Arnoczky, S.P., Lau, R.E. (1985) Correlative biomechanical and histologic study of the cranial cruciate ligament in dogs. *American Journal of Veterinary Research* **46**, 1842-1854

94. Laros, G.S., Tipton, C.M., Cooper, R.R. (1971) Influence of physical activity on ligament insertions in the knees of dogs. *Journal of Bone and Joint Surgery (American)* **53**, 275-286

95. Aiken, S.W., Kass, P.H., Toombs, J.P. (1995) Intercondylar notch width in dogs with and without cranial cruciate ligament injuries. *Veterinary and Comparative Orthopaedics and Traumatology* **8**, 128-132

96. Comerford, E.J., Tarlton, J.F., Avery, N.C., Bailey, A.J., Innes, J.F. (2006) Distal femoral intercondylar notch dimensions and their relationship to composition and metabolism of the canine anterior cruciate ligament. *Osteoarthritis and Cartilage* **14**, 273-278

97. Slauterbeck, J.R., Pankratz, K., Xu, K.T., Bozeman, S.C., Hardy, D.M. (2004) Canine ovariohysterectomy and orchiectomy increases the prevalence of ACL injury. *Clinical Orthopaedics and Related Research* **429**, 301-305

98. Johnson, J.A., Austin, C., Breur, G.J. (1994) Incidence of canine appendicular musculoskeletal disorders in 16 veterinary teaching hospitals from 1980 through 1989. *Veterinary and Comparative Orthopaedics and Traumatology* **7**, 56-69

99. Rayward, R.M., Thomson, D.G., Davies, J.V., Innes, J.F., Whitelock, R.G. (2004) Progression of osteoarthritis following TPLO surgery: a prospective radiographic study of 40 dogs. *Journal of Small Animal Practice* **45**, 92-97

100. Aragon, C.L., Budsberg, S.C. (2005) Applications of Evidence-Based Medicine: Cranial Cruciate Ligament Injury Repair in the Dog. *Veterinary Surgery* **34**, 93-98

101. Wilke, V.L., Robinson, D.A., Evans, R.B., Rothschild, M.F., Conzemius, M.G. (2005) Estimate of the annual economic impact of treatment of cranial cruciate ligament injury in dogs in the United States. *Journal of the American Veterinary Medical Association* **227**, 1604-1607

102. Hill, C.L., Seo, G.S., Gale, D., Totterman, S., Gale, M.E., Felson, D.T. (2005) Cruciate ligament integrity in osteoarthritis of the knee. *Arthritis and Rheumatism* **52**, 794-799

103. Wada, M., Imura, S., Baba, H., Shimada, S. (1996) Knee laxity in patients with osteoarthritis and rheumatoid arthritis. *British Journal of Rheumatology* **35**, 560-563

104. Sharma, L., Lou, C., Felson, D.T., Dunlop, D.D., Kirwan-Mellis, G., Hayes, K.W., Weinrach, D., Buchanan, T.S. (1999) Laxity in healthy and osteoarthritic knees. *Arthritis and Rheumatism* **42**, 861-870

105. Wingfield, C., Amis, A.A., Stead, A.C., Law, H.T. (2000) Comparison of the biomechanical properties of rottweiler and racing greyhound cranial cruciate ligaments. *Journal of Small Animal Practice* **41**, 303-307

106. Wingfield, C., Amis, A.A., Stead, A.C., Law, H.T. (2000) Cranial cruciate stability in the rottweiler and racing greyhound: an in vitro study. *Journal of Small Animal Practice* **41**, 193-197

107. Quasnichka, H.L., Anderson-MacKenzie, J.M., Tarlton, J.F., Sims, T.J., Billingham, M.E., Bailey, A.J. (2005) Cruciate ligament laxity and femoral intercondylar notch narrowing in early-stage knee osteoarthritis. *Arthritis and Rheumatism* **52**, 3100-3109

108. Duval, J.M., Budsberg, S.C., Flo, G.L., Sammarco, J.L. (1999) Breed, sex, and body weight as risk factors for rupture of the cranial cruciate ligament in young dogs. *Journal of the American Veterinary Medical Association* **215**, 811-814

109. Wilke, V.L., Ruhe, A., Conzemius, M.G., Rothschild, M.F.Predisposition To Rupture Of The Cranial Cruciate Ligament in the Dog is Genetically Associated with Chromosome 3. In: Proceedings of the 3rd World Veterinary Orthopaedic Congress p.161

110. Ubbink, G.J., van de, B.J., Hazewinkel, H.A., Rothuizen, J. (1998) Cluster analysis of the genetic heterogeneity and disease distributions in purebred dog populations. *Veterinary Record* **142**, 209-213

111. Richman, M., Mellersh, C.S., Andre, C., Galibert, F., Ostrander, E.A. (2001) Characterization of a minimal screening set of 172 microsatellite markers for genome-wide screens of the canine genome. *Journal of Biochemical and Biophysical Methods* **47**, 137-149

112. Scott, S.W. (2002) Genetics of primary generalized osteoarthritis. *Molecular Genetics and Metabolism* **77**, 31-34

113. Risch, N., Merikangas, K. (1996) The future of genetic studies of complex human diseases. *Science* **273**, 1516-1517

114. Riley, J.H., Allan, C.J., Lai, E., Roses, A. (2000) The use of single nucleotide polymorphisms in the isolation of common disease genes. *Pharmacogenomics* **1**, 39-47

115. Hirschhorn, J.N., Lohmueller, K., Byrne, E., Hirschhorn, K. (2002) A comprehensive review of genetic association studies. *Genetics in Medicine* **4**, 45-61

116. Huizinga, T.W., Pisetsky, D.S., Kimberly, R.P. (2004) Associations, populations, and the truth: recommendations for genetic association studies in Arthritis & Rheumatism. *Arthritis and Rheumatism* **50**, 2066-2071

117. 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 Research* **14**, 2388-2396

118. Gu, C.C., Rao, D.C. (2003) Designing an optimum genetic association study using dense SNP markers and family-based sample. *Frontiers in Bioscience* **8**, S68-S80

119. Schork, N.J., Fallin, D., Lanchbury, J.S. (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clinical Genetics* **58**, 250-264

120. Krawczak, M., Reiss, J., Cooper, D.N. (1992) The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human Genetics* **90**, 41-54

121. Drazen, J.M., Yandava, C.N., Dube, L., Szczerback, N., Hippensteel, R., Pillari, A., Israel, E., Schork, N., Silverman, E.S., Katz, D.A., Drajesk, J. (1999) Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. *Nature Genetics* **22**, 168-170

122. Kwok, P.Y. (2001) Methods for genotyping single nucleotide polymorphisms. *Annual Review of Genomics and Human Genetics* **2**, 235-258

123. Brookes, A.J. (1999) The essence of SNPs. *Gene* **234**, 177-186

124. Hu, N., Wang, C., Hu, Y., Yang, H.H., Giffen, C., Tang, Z.Z., Han, X.Y., Goldstein, A.M., Emmert-Buck, M.R., Buetow, K.H., Taylor, P.R., Lee, M.P. (2005) Genome-Wide Association Study in Esophageal Cancer Using GeneChip Mapping 10K Array. *Cancer Research* **65**, 2542-2546

125. Syvanen, A.C. (2005) Toward genome-wide SNP genotyping. *Nature Genetics* **37**, S5-S10

126. Kruglyak, L. (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics* **22**, 139-144

127. Lindblad-Toh, K.A. (2007) Trait Mapping Using A Canine SNP Array: A Model For Equine Genetics. In: Proceedings of the Plant & Animal Genomes XV Conference p.101

128. Loughlin, J., Dowling, B., Chapman, K., Marcelline, L., Mustafa, Z., Southam, L., Ferreira, A., Ciesielski, C., Carson, D.A., Corr, M. (2004) Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. *Proceedings of the National Academy of Sciences* **101**, 9757-9762

129. Horton, W.E., Jr., Lethbridge-Cejku, M., Hochberg, M.C., Balakir, R., Precht, P., Plato, C.C., Tobin, J.D., Meek, L., Doege, K. (1998) An association between an aggrecan polymorphic allele and bilateral hand osteoarthritis in elderly white men: data from the Baltimore Longitudinal Study of Aging (BLSA). *Osteoarthritis and Cartilage* **6**, 245-251

130. Valdes, A.M., Hart, D.J., Jones, K.A., Surdulescu, G., Swarbrick, P., Doyle, D.V., Schafer, A.J., Spector, T.D. (2004) Association study of candidate genes for the prevalence and progression of knee osteoarthritis. *Arthritis and Rheumatism* **50**, 2497-2507

131. Meulenbelt, I., Bijkerk, C., Breedveld, F.C., Slagboom, P.E. (1997) Genetic linkage analysis of 14 candidate gene loci in a family with autosomal dominant osteoarthritis without dysplasia. *Journal of Medical Genetics* **34**, 1024-1027

132. Mabuchi, A., Ikeda, T., Fukuda, A., Koshizuka, Y., Hiraoka, H., Miyoshi, K., Haga, N., Kawaguchi, H., Kawakami, A., Yamamoto, S., Takatori, Y., Nakamura, K., Ikegawa, S. (2001) Identification of sequence polymorphisms of the COMP (cartilage oligomeric matrix protein) gene and association study in osteoarthrosis of the knee and hip joints. *Journal of Human Genetics* **46**, 456-462

133. Fernandes, J.C., Martel-Pelletier, J., Pelletier, J.P. (2002) The role of cytokines in osteoarthritis pathophysiology. *Biorheology* **39**, 237-246

134. Goldring, S.R., Goldring, M.B., Buckwalter, J. (2004) The role of cytokines in cartilage matrix degeneration in osteoarthritis. *Clinical Orthopaedics* **427**, S27-S36

135. Okazaki, K., Sandell, L.J., Eyre, D., Schurman, D.J. (2004) Extracellular matrix gene regulation. *Clinical Orthopaedics* **427**, S123-S128

136. Loughlin, J., Sinsheimer, J.S., Mustafa, Z., Carr, A.J., Clipsham, K., Bloomfield, V.A., Chitnavis, J., Bailey, A., Sykes, B., Chapman, K. (2000) Association analysis of the vitamin D receptor gene, the type I collagen gene COL1A1, and the estrogen receptor gene in idiopathic osteoarthritis. *Journal of Rheumatology* **27**, 779-784

137. Aerssens, J., Dequeker, J., Peeters, J., Breemans, S., Boonen, S. (1998) Lack of association between osteoarthritis of the hip and gene polymorphisms of VDR, COL1A1, and COL2A1 in postmenopausal women. *Arthritis and Rheumatism* **41**, 1946-1950

138. Lian, K., Zmuda, J.M., Nevitt, M.C., Lui, L., Hochberg, M.C., Greene, D., Li, J., Wang, J., Lane, N.E. (2005) Type I collagen alpha1 Sp1 transcription factor binding site polymorphism is associated with reduced risk of hip osteoarthritis defined by severe joint space narrowing in elderly women. *Arthritis and Rheumatism* **52**, 1431-1436

139. Bleasel, J.F., Holderbaum, D., Brancolini, V., Moskowitz, R.W., Considine, E.L., Prockop, D.J., Devoto, M., Williams, C.J. (1998) Five families with arginine 519-cysteine mutation in COL2A1: evidence for three distinct founders. *Human Mutation* **12**, 172-176

140. Williams, C.J., Harrison, D.A., Hopkinson, I., Baldwin, C.T., Ahmad, N.N., la-Kokko, L., Korn, R.M., Buxton, P.G., Dimascio, J., Considine, E.L., . (1992) Detection of sequence variants in the gene for human type II procollagen (COL2A1) by direct sequencing of polymerase chain reaction-amplified genomic DNA. *Human Mutation* **1**, 403-416

141. Meulenbelt, I., Bijkerk, C., De Wildt, S.C., Miedema, H.S., Breedveld, F.C., Pols, H.A., Hofman, A., van Duijn, C.M., Slagboom, P.E. (1999) Haplotype analysis of three polymorphisms of the COL2A1 gene and associations with generalised radiological osteoarthritis. *Annals of Human Genetics* **63**, 393-400

142. Ikeda, T., Mabuchi, A., Fukuda, A., Kawakami, A., Ryo, Y., Yamamoto, S., Miyoshi, K., Haga, N., Hiraoka, H., Takatori, Y., Kawaguchi, H., Nakamura, K., Ikegawa, S. (2002) Association analysis of single nucleotide polymorphisms in cartilage-specific collagen genes with knee and hip osteoarthritis in the Japanese population. *Journal of Bone and Mineral Research* **17**, 1290-1296

143. Zhai, G., Rivadeneira, F., Houwing-Duistermaat, J.J., Meulenbelt, I., Bijkerk, C., Hofman, A., van Meurs, J.B.J., Uitterlinden, A.G., Pols, H.A.P., Slagboom, P.E., van Duijn, C.M. (2004) Insulin-like growth factor I gene promoter polymorphism, collagen type II {alpha}1 (COL2A1) gene, and the prevalence of radiographic osteoarthritis: the Rotterdam Study. *Annals of the Rheumatic Diseases* **63**, 544-548

144. Uitterlinden, A.G., Burger, H., van Duijn, C.M., Huang, Q., Hofman, A., Birkenhager, J.C., van Leeuwen, J.P., Pols, H.A. (2000) Adjacent genes, for COL2A1 and the vitamin D receptor, are associated with separate features of radiographic osteoarthritis of the knee. *Arthritis and Rheumatism* **43**, 1456-1464

145. Baldwin, C.T., Cupples, L.A., Joost, O., Demissie, S., Chaisson, C., Mcalindon, T., Myers, R.H., Felson, D. (2002) Absence of linkage or association for osteoarthritis with the vitamin D receptor/type II collagen locus: the Framingham Osteoarthritis Study. *Journal of Rheumatology* **29**, 161-165

146. Loughlin, J., Irven, C., Fergusson, C., Sykes, B. (1994) Sibling pair analysis shows no linkage of generalized osteoarthritis to the loci encoding type II collagen, cartilage link protein or cartilage matrix protein. *British Journal of Rheumatology* **33**, 1103-1106

147. Vikkula, M., Nissila, M., Hirvensalo, E., Nuotio, P., Palotie, A., Aho, K., Peltonen, L. (1993) Multiallelic polymorphism of the cartilage collagen gene: no association with osteoarthrosis. *Annals of the Rheumatic Diseases* **52**, 762-764

148. Mustafa, Z., Chapman, K., Irven, C., Carr, A.J., Clipsham, K., Chitnavis, J., Sinsheimer, J.S., Bloomfield, V.A., McCartney, M., Cox, O., Sykes, B., Loughlin, J. (2000) Linkage analysis of candidate genes as susceptibility loci for osteoarthritis--suggestive linkage of COL9A1 to female hip osteoarthritis. *Rheumatology* **39**, 299-306

149. Alizadeh, B.Z., Njajou, O.T., Bijkerk, C., Meulenbelt, I., Wildt, S.C., Hofman, A., Pols, H.A., Slagboom, P.E., Duijn, C.M. (2005) Evidence for a role of the genomic region of the gene encoding for the alpha1 chain of type IX collagen (COL9A1) in hip osteoarthritis: A population-based study. *Arthritis and Rheumatism* **52**, 1437-1442

150. Loughlin, J., Mustafa, Z., Dowling, B., Southam, L., Marcelline, L., Raina, S.S., la-Kokko, L., Chapman, K. (2002) Finer linkage mapping of a primary hip osteoarthritis susceptibility locus on chromosome 6. *European Journal of Human Genetics* **10**, 562-568

151. Stefansson, S.E., Jonsson, H., Ingvarsson, T., Manolescu, I., Jonsson, H.H., Olafsdottir, G., Palsdottir, E., Stefansdottir, G., Sveinbjornsdottir, G., Frigge, M.L., Kong, A., Gulcher, J.R., Stefansson, K. (2003) Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. *American Journal of Human Genetics* **72**, 1448-1459

152. Loughlin, J., Dowling, B., Mustafa, Z., Smith, A., Sykes, B., Chapman, K. (2000) Analysis of the association of the matrillin-1 gene (CRTM) with osteoarthritis: comment on the article by Meulenbelt *et al. Arthritis and Rheumatism* **43**, 1423-1425

153. Kirk, K.M., Doege, K.J., Hecht, J., Bellamy, N., Martin, N.G. (2003) Osteoarthritis of the hands, hips and knees in an Australian twin sample--evidence of association with the aggrecan VNTR polymorphism. *Twin Research* **6**, 62-66

154. Kizawa, H., Kou, I., Iida, A., Sudo, A., Miyamoto, Y., Fukuda, A., Mabuchi, A., Kotani, A., Kawakami, A., Yamamoto, S., Uchida, A., Nakamura, K., Notoya, K., Nakamura, Y., Ikegawa, S. (2005) An aspartic acid repeat polymorphism in asporin inhibits chondrogenesis and increases susceptibility to osteoarthritis. *Nature Genetics* **37**, 138-144

155. Bergink, A.P., van Meurs, J.B., Loughlin, J., Arp, P.P., Fang, Y., Hofman, A., van Leeuwen, J.P., van Duijn, C.M., Uitterlinden, A.G., Pols, H.A. (2003) Estrogen receptor alpha gene haplotype is associated with radiographic osteoarthritis of the knee in elderly men and women. *Arthritis and Rheumatism* **48**, 1913-1922

156. Jin, S.Y., Hong, S.J., In, Y.H., Park, S.D., Yoo, M.C., Lee, H.J., Hong, M.S., Park, H.J., Yoon, S.H., Kim, B.S., Yim, S.V., Park, H.K., Chung, J.H. (2004) Estrogen receptor-alpha gene haplotype is associated with primary knee osteoarthritis in Korean population. *Arthritis Research and Therapy* **6**, R415-R421

157. Ushiyama, T., Ueyama, H., Inoue, K., Nishioka, J., Ohkubo, I., Hukuda, S. (1998) Estrogen receptor gene polymorphism and generalized osteoarthritis. *Journal of Rheumatology* **25**, 134-137

158. Keen, R.W., Hart, D.J., Lanchbury, J.S., Spector, T.D. (1997) Association of early osteoarthritis of the knee with a Taq I polymorphism of the vitamin D receptor gene. *Arthritis and Rheumatism* **40**, 1444-1449

159. Loughlin, J., Dowling, B., Mustafa, Z., Chapman, K. (2002) Association of the interleukin-1 gene cluster on chromosome 2q13 with knee osteoarthritis. *Arthritis and Rheumatism* **46**, 1519-1527

160. Smith, A.J., Keen, L.J., Billingham, M.J., Perry, M.J., Elson, C.J., Kirwan, J.R., Sims, J.E., Doherty, M., Spector, T.D., Bidwell, J.L. (2004) Extended haplotypes and linkage disequilibrium in the IL1R1-IL1A-IL1B-IL1RN gene cluster: association with knee osteoarthritis. *Genes and Immunity* **5**, 451-460

161. Smith, A.J., Elson, C.J., Perry, M.J., Bidwell, J.L. (2005) Accuracy of haplotype association studies is enhanced by increasing number of polymorphic loci examined: comment on the article by Meulenbelt *et al.* *Arthritis and Rheumatism* **52**, 675-676

162. Moos, V., Rudwaleit, M., Herzog, V., Hohlig, K., Sieper, J., Muller, B. (2000) Association of genotypes affecting the expression of interleukin-1beta or interleukin-1 receptor antagonist with osteoarthritis. *Arthritis and Rheumatism* **43**, 2417-2422

163. Stern, A.G., de Carvalho, M.R., Buck, G.A., Adler, R.A., Rao, T.P., Disler, D., Moxley, G. (2003) Association of erosive hand osteoarthritis with a single nucleotide polymorphism on the gene encoding interleukin-1 beta. *Osteoarthritis and Cartilage* **11**, 394-402

164. Forster, T., Chapman, K., Loughlin, J. (2004) Common variants within the interleukin 4 receptor alpha gene (IL4R) are associated with susceptibility to osteoarthritis. *Human Genetics* **114**, 391-395

165. Kawahara, C., Forster, T., Chapman, K., Carr, A., Loughlin, J. (2005) Genetic association analysis of the IGFBP7, ADAMTS3, and IL8 genes as the potential osteoarthritis susceptibility that maps to chromosome 4q. *Annals of the Rheumatic Diseases* **64**, 474-476

166. Fytili, P., Giannatou, E., Karachalios, T., Malizos, K., Tsezou, A. (2005) Interleukin-10G and interleukin-10R microsatellite polymorphisms and osteoarthritis of the knee. *Clinical and Experimental Rheumatology* **23**, 621-627

167. Riyazi, N., Kurreeman, F.A., Huizinga, T.W., Dekker, F.W., Stoeken-Rijsbergen, G., Kloppenburg, M. (2005) The role of interleukin 10 promoter polymorphisms in the susceptibility of distal interphalangeal osteoarthritis. *Journal of Rheumatology* **32**, 1571-1575

168. Meulenbelt, I., Bijkerk, C., Miedema, H.S., Breedveld, F.C., Hofman, A., Valkenburg, H.A., Pols, H.A.P., Slagboom, P.E., van Duijn, C.M. (1998) A genetic association study of the IGF-1 gene and radiological osteoarthritis in a population-based cohort study (the Rotterdam study). *Annals of the Rheumatic Diseases* **57**, 371-374

169. Yamada, Y., Okuizumi, H., Miyauchi, A., Takagi, Y., Ikeda, K., Harada, A. (2000) Association of transforming growth factor beta1 genotype with spinal osteophytosis in Japanese women. *Arthritis and Rheumatism* **43**, 452-460

170. Forster, T., Chapman, K., Marcelline, L., Mustafa, Z., Southam, L., Loughlin, J. (2004) Finer linkage mapping of primary osteoarthritis susceptibility loci on chromosomes 4 and 16 in families with affected women. *Arthritis and Rheumatism* **50**, 98-102

171. Sakkas, L.I., Macfarlane, D.G., Bird, H., Welsh, K.I., Panayi, G.S. (1990) Association of osteoarthritis with homozygosity for a 5.8 kb Taq I fragment of the alpha 1-antichymotrypsin gene. *British Journal of Rheumatology* **29**, 245-248

172. Hong, S.J., Yang, H.I., Yoo, M.C., In, C.S., Yim, S.V., Jin, S.Y., Choe, B.K., Chung, J.H. (2003) Angiotensin converting enzyme gene polymorphism in Korean patients with primary knee osteoarthritis. *Experimental and Molecular Medicine* **35**, 189-195

173. Min, J.L., Meulenbelt, I., Riyazi, N., Kloppenburg, M., Houwing-Duistermaat, J.J., Seymour, A.B., Pols, H.A., van Duijn, C.M., Slagboom,

P.E. (2005) Association of the frizzled-related protein gene with symptomatic osteoarthritis at multiple sites. *Arthritis and Rheumatism* **52**, 1077-1080

174. Neumann, E., Barnum, S.R., Tarner, I.H., Echols, J., Fleck, M., Judex, M., Kullmann, F., Mountz, J.D., Scholmerich, J., Gay, S., Muller-Ladner, U. (2002) Local production of complement proteins in rheumatoid arthritis synovium. *Arthritis and Rheumatism* **46**, 934-945

175. Roivainen, A., Jalava, J., Pirila, L., Yli-Jama, T., Tiusanen, H., Toivanen, P. (1997) H-*ras* oncogene point mutations in arthritic synovium. *Arthritis and Rheumatism* **40**, 1636-1643

176. Roivainen, A., Zhu, F., Sipola, E., Yli-Jama, T., Toivanen, P. (2001) Failure to verify H-*ras* mutations in arthritic synovium: comment on the article by Roivainen et al. *Arthritis and Rheumatism* **44**, 2705

177. Ikeda, T., Mabuchi, A., Fukuda, A., Hiraoka, H., Kawakami, A., Yamamoto, S., Machida, H., Takatori, Y., Kawaguchi, H., Nakamura, K., Ikegawa, S. (2001) Identification of sequence polymorphisms in two sulfation-related genes, PAPSS2 and SLC26A2, and an association analysis with knee osteoarthritis. *Journal of Human Genetics* **46**, 538-543

178. Mototani, H., Mabuchi, A., Saito, S., Fujioka, M., Iida, A., Takatori, Y., Kotani, A., Kubo, T., Nakamura, K., Sekine, A., Murakami, Y., Tsunoda, T., Notoya, K., Nakamura, Y., Ikegawa, S. (2005) A functional single nucleotide polymorphism in the core promoter region of CALM1 is associated with hip osteoarthritis in Japanese. *Human Molecular Genetics* **14**, 1009-1017

179. Nakamura, I., Okawa, A., Ikegawa, S., Takaoka, K., Nakamura, Y. (1999) Genomic organization, mapping, and polymorphisms of the gene encoding human cartilage intermediate layer protein (CILP). *Journal of Human Genetics* **44**, 203-205

180. Lorenzo, P., Bayliss, M.T., Heinegard, D. (2004) Altered patterns and synthesis of extracellular matrix macromolecules in early osteoarthritis. *Matrix Biology* **23**, 381-391

181. DiCesare, P.E., Morgelin, M., Carlson, C.S., Pasumarti, S., Paulsson, M. (1995) Cartilage oligomeric matrix protein: isolation and characterization from human articular cartilage. *Journal of Orthopaedic Research* **13**, 422-428

182. Salminen, H., Perala, M., Lorenzo, P., Saxne, T., Heinegard, D., Saamanen, A.M., Vuorio, E. (2000) Up-regulation of cartilage oligomeric matrix protein at the onset of articular cartilage degeneration in a transgenic mouse model of osteoarthritis. *Arthritis and Rheumatism* **43**, 1742-1748

183. Misumi, K., Vilim, V., Carter, S.D., Ichihashi, K., Oka, T., Sakamoto, H. (2002) Concentrations of cartilage oligomeric matrix protein in dogs with

naturally developing and experimentally induced arthropathy. *American Journal of Veterinary Research* **63**, 598-603

184. Briggs, M.D., Hoffman, S.M., King, L.M., Olsen, A.S., Mohrenweiser, H., Leroy, J.G., Mortier, G.R., Rimoin, D.L., Lachman, R.S., Gaines, E.S., . (1995) Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nature Genetics* **10**, 330-336

185. Lorenzo, P., Aspberg, A., Onnerfjord, P., Bayliss, M.T., Neame, P.J., Heinegard, D. (2001) Identification and Characterization of Asporin. A novel member of the leucine-rich repeat protein family closely related to decorin and biglycan. *Journal of Biological Chemistry* **276**, 12201-12211

186. Spector, T.D., Nandra, D., Hart, D.J., Doyle, D.V. (1997) Is hormone replacement therapy protective for hand and knee osteoarthritis in women?: The Chingford study. *Annals of the Rheumatic Diseases* **56**, 432-434

187. Spector, T.D., Campion, G.D. (1989) Generalised osteoarthritis: a hormonally mediated disease. *Annals of the Rheumatic Diseases* **48**, 523-527

188. McAlindon, T.E., Felson, D.T., Zhang, Y., Hannan, M.T., Aliabadi, P., Weissman, B., Rush, D., Wilson, P.W.F., Jacques, P. (1996) Relation of Dietary Intake and Serum Levels of Vitamin D to Progression of Osteoarthritis of the Knee among Participants in the Framingham Study. *Annals of Internal Medicine* **125**, 353-359

189. Millward-Sadler, S.J., Wright, M.O., Lee, H.S., Nishida, K., Caldwell, H., Nuki, G., Salter, D.M. (1999) Integrin-regulated Secretion of Interleukin 4: A Novel Pathway of Mechanotransduction in Human Articular Chondrocytes. *The Journal of Cell Biology* **145**, 183-189

190. Iannone, F., De, B.C., Dell'Accio, F., Covelli, M., Cantatore, F.P., Patella, V., Lo, B.G., Lapadula, G. (2001) Interleukin-10 and interleukin-10 receptor in human osteoarthritic and healthy chondrocytes. *Clinical and Experimental Rheumatology* **19**, 139-145

191. Simonet, W.S., Lacey, D.L., Dunstan, C.R., Kelley, M., Chang, M.-S., Luthy, R., Nguyen, H.Q., Wooden, S., Bennett, L., Boone, T. (1997) Osteoprotegerin: A Novel Secreted Protein Involved in the Regulation of Bone Density. *Cell* **89**, 309-319

192. Fazzalari, N.L., Kuliwaba, J.S., Atkins, G.J., Forwood, M.R., Findlay, D.M. (2001) The ratio of messenger RNA levels of receptor activator of nuclear factor kappaB ligand to osteoprotegerin correlates with bone remodeling indices in normal human cancellous bone but not in osteoarthritis. *Journal of Bone and Mineral Research* **16**, 1015-1027

193. Komuro, H., Olee, T., Kuhn, K., Quach, J., Brinson, D.C., Shikhman, A., Valbracht, J., Creighton-Achermann, L., Lotz, M. (2001) The osteoprotegerin/receptor activator of nuclear factor kappaB/receptor

activator of nuclear factor kappaB ligand system in cartilage. *Arthritis and Rheumatism* **44**, 2768-2776

194. Loeser, R.F. (2000) Chondrocyte integrin expression and function. *Biorheology* **37**, 109-116

195. Sailor, L.Z., Hewick, R.M., Morris, E.A. (1996) Recombinant human bone morphogenetic protein-2 maintains the articular chondrocyte phenotype in long-term culture. *Journal of Orthopaedic Research* **14**, 937-945

196. Nakase, T., Miyaji, T., Tomita, T., Kaneko, M., Kuriyama, K., Myoui, A., Sugamoto, K., Ochi, T., Yoshikawa, H. (2003) Localization of bone morphogenetic protein-2 in human osteoarthritic cartilage and osteophyte. *Osteoarthritis and Cartilage* **11**, 278-284

197. Trippel, S.B., Buckwalter, J. (2004) Growth factor inhibition: potential role in the etiopathogenesis of osteoarthritis. *Clinical Orthopaedics* **427**, S47-S52

198. Abe, E., Mocharla, H., Yamate, T., Taguchi, Y., Manolagas, S.C. (1999) Meltrin-alpha, a fusion protein involved in multinucleated giant cell and osteoclast formation. *Calcified Tissue International* **64**, 508-515

199. Pfander, D., Cramer, T., Deuerling, D., Weseloh, G., Swoboda, B. (2000) Expression of thrombospondin-1 and its receptor CD36 in human osteoarthritic cartilage. *Annals of the Rheumatic Diseases* **59**, 448-454

200. Amin, A.R., Attur, M., Patel, R.N., Thakker, G.D., Marshall, P.J., Rediske, J., Stuchin, S.A., Patel, I.R., Abramson, S.B. (1997) Superinduction of Cyclooxygenase-2 Activity in Human Osteoarthritis-affected Cartilage . Influence of Nitric Oxide. *Journal of Clinical Investigation* **99**, 1231-1237

201. Koki, A., Khan, N.K., Woerner, B.M., Dannenberg, A.J., Olson, L., Seibert, K., Edwards, D., Hardy, M., Isakson, P., Masferrer, J.L. (2002) Cyclooxygenase-2 in human pathological disease. *Advances in Experimental Medicine and Biology* **507**, 177-184

202. Siegle, I., Klein, T., Backman, J.T., Saal, J.G., Nusing, R.M., Fritz, P. (1998) Expression of cyclooxygenase 1 and cyclooxygenase 2 in human synovial tissue: differential elevation of cyclooxygenase 2 in inflammatory joint diseases. *Arthritis and Rheumatism* **41**, 122-129

203. Wewer, U.M., Ibaraki, K., Schjorring, P., Durkin, M.E., Young, M.F., Albrechtsen, R. (1994) A potential role for tetranectin in mineralization during osteogenesis. *The Journal of Cell Biology* **127**, 1767-1775

204. Kamper, E.F., Kopeikina, L.T., Trontzas, P., Kyriazis, N.C., Vaiopoulos, G., Stavridis, J. (1998) Comparative study of tetranectin levels in serum and synovial fluid of patients with rheumatoid arthritis, seronegative spondylarthritis and osteoarthritis. *Clinical Rheumatology* **17**, 318-324

205. Janciauskiene, S. (2001) Conformational properties of serine proteinase inhibitors (serpins) confer multiple pathophysiological roles. *Biochimica et Biophysica Acta* **1535**, 221-235

206. Lesley, J., Gal, I., Mahoney, D.J., Cordell, M.R., Rugg, M.S., Hyman, R., Day, A.J., Mikecz, K. (2004) TSG-6 Modulates the Interaction between Hyaluronan and Cell Surface CD44. *Journal of Biological Chemistry* **279**, 25745-25754

207. Bayliss, M.T., Howat, S.L.T., Dudhia, J., Murphy, J.M., Barry, F.P., Edwards, J.C.W., Day, A.J. (2001) Up-regulation and differential expression of the hyaluronan-binding protein TSG-6 in cartilage and synovium in rheumatoid arthritis and osteoarthritis. *Osteoarthritis and Cartilage* **9**, 42-48

208. Cassim, B., Naidoo, S., Ramsaroop, R., Bhoola, K.D. (1997) Immunolocalization of bradykinin receptors on human synovial tissue. *Immunopharmacology* **36**, 121-125

209. Guo, X., Day, T.F., Jiang, X., Garrett-Beal, L., Topol, L., Yang, Y. (2004) Wnt/{beta}-catenin signaling is sufficient and necessary for synovial joint formation. *Genes and Development* **18**, 2404-2417

210. Bodine, P.V.N., Zhao, W., Kharode, Y.P., Bex, F.J., Lambert, A.J., Goad, M.B., Gaur, T., Stein, G.S., Lian, J.B., Komm, B.S. (2004) The Wnt Antagonist Secreted Frizzled-Related Protein-1 Is a Negative Regulator of Trabecular Bone Formation in Adult Mice. *Molecular Endocrinology* **18**, 1222-1237

211. Plaas, A.H., West, L.A., Wong-Palms, S., Nelson, F.R. (1998) Glycosaminoglycan Sulfation in Human Osteoarthritis. Disease-related alterations at the non-reducing termini of chondroitin and dermatan sulfate. *Journal of Biological Chemistry* **273**, 12642-12649

212. Richard, M., Broquet, P., Vignon, E., Peschard, M.J., Carret, J.P., Louisot, P. (1991) Calmodulin-dependent collagenase and proteoglycanase activities in chondrocytes from human osteoarthritic cartilage. *Biochemical and Biophysical Research Communications* **174**, 1204-1207

213. Reveille, J.D. (2005) Genetic studies in the rheumatic diseases: present status and implications for the future. *Journal of Rheumatology Supplement* **72**, 10-13

214. Riyazi, N., Spee, J., Huizinga, T.W.J., Schreuder, G.M.T., de Vries, R.R.P., Dekker, F.W., Kloppenburg, M. (2003) HLA class II is associated with distal interphalangeal osteoarthritis. *Annals of the Rheumatic Diseases* **62**, 227-230

215. Moos, V., Menard, J., Sieper, J., Sparmann, M., Muller, B. (2002) Association of HLA-DRB1*02 with osteoarthritis in a cohort of 106 patients. *Rheumatology* **41**, 666-669

216. Valdes, A.M., Van-Oene, M., Hart, D.J., Surdulescu, G.L., Loughlin, J., Doherty, M., Spector, T.D. (2006) Reproducible genetic associations between candidate genes and clinical knee osteoarthritis in men and women. *Arthritis and Rheumatism* **54**, 533-539

217. Wind, A.P. (1986) Elbow incongruity and developmental elbow disease in the dog: Part I. *Journal of the American Animal Hospital Association* **22**, 711-724

218. Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E.A., Ponting, C.P., Galibert, F., Smith, D.R., deJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre, S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.P., Parker, H.G., Pollinger, J.P., Searle, S.M.J., Sutter, N.B., Thomas, R., Webber, C., Lander, E.S. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803-819

219. Bluteau, G., Conrozier, T., Mathieu, P., Vignon, E., Herbage, D., Mallein-Gerin, F. (2001) Matrix metalloproteinase-1, -3, -13 and aggrecanase-1 and -2 are differentially expressed in experimental osteoarthritis. *Biochimica et Biophysica Acta* **1526**, 147-158

220. Zafarullah, M., Martel-Pelletier, J., Cloutier, J.M., Gedamu, L., Pelletier, J.P. (1993) Distinct expression pattern of early- and late-response genes in normal and osteoarthritic human synovial membranes. *Osteoarthritis and Cartilage* **1**, 151-156

221. Hui, W., Rowan, A.D., Richards, C.D., Cawston, T.E. (2003) Oncostatin M in combination with tumor necrosis factor alpha induces cartilage damage and matrix metalloproteinase expression in vitro and in vivo. *Arthritis and Rheumatism* **48**, 3404-3418

222. Burton-Wurster, N., Mateescu, R.G., Todhunter, R.J., Clements, K.M., Sun, Q., Scarpino, V., Lust, G. (2005) Genes in Canine Articular Cartilage That Respond to Mechanical Injury: Gene Expression Studies With Affymetrix Canine GeneChip. *Journal of Heredity* **97**, 821-828

223. Bustin, S.A., Nolan, T. (2004) Pitfalls of Quantitative Real-Time Reverse-Transcription Polymerase Chain Reaction. *Journal of Biomolecular Techniques* **15**, 155-166

224. Imbeaud, S., Graudens, E., Boulanger, V., Barlet, X., Zaborski, P., Eveno, E., Mueller, O., Schroeder, A., Auffray, C. (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Research* **33**, e56

225. Bustin, S.A. (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of Molecular Endocrinology* **29**, 23-39

226. Lekanne Deprez, R.H., Fijnvandraat, A.C., Ruijter, J.M., Moorman, A.F.M. (2002) Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. *Analytical Biochemistry* **307**, 63-69

227. Nolan, T., Hands, R.E., Ogunkolade, W., Bustin, S.A. (2006) SPUD: A quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Analytical Biochemistry* **351**, 308-310

228. Manchester, K.L. (1996) Use of UV methods for measurement of protein and nucleic acid concentrations. *Biotechniques* **20**, 968-970

229. Bonini, J.A., Hofmann, C. (1991) A rapid, accurate, nonradioactive method for quantitating RNA on agarose gels. *Biotechniques* **11**, 708-710

230. Auer, H., Lyianarachchi, S., Newsom, D., Klisovic, M.I., Marcucci, u., Kornacker, K. (2003) Chipping away at the chip bias: RNA degradation in microarray analysis. *Nature Genetics* **35**, 292-293

231. Mueller, O., Lightfoot, S., Schroeder, A. (2004) RNA Integrity Number (RIN) – Standardization of RNA Quality Control. *www chem agilent com*

232. McKenna, L.A., Gehrsitz, A., Soder, S., Eger, W., Kirchner, T., Aigner, T. (2000) Effective Isolation of High-Quality Total RNA from Human Adult Articular Cartilage. *Analytical Biochemistry* **286**, 80-85

233. Fehr, J.E., Trotter, G.W., Oxford, J.T., Hart, D.A. (2000) Comparison of Northern blot hybridization and a reverse transcriptase-polymerase chain reaction technique for measurement of mRNA expression of metalloproteinases and matrix components in articular cartilage and synovial membrane from horses with osteoarthritis. *American Journal of Veterinary Research* **61**, 900-905

234. Flannery, C.R., Little, C.B., Caterson, B., Hughes, C.E. (1999) Effects of culture conditions and exposure to catabolic stimulators (IL-1 and retinoic acid) on the expression of matrix metalloproteinases (MMPs) and disintegrin metalloproteinases (ADAMs) by articular cartilage chondrocytes. *Matrix Biology* **18**, 225-237

235. Huggett, J., Dheda, K., Bustin, S., Zumla, A. (2005) Real-time RT-PCR normalisation; strategies and considerations. *Genes and Immunity* **6**, 279-284

236. Dheda, K., Huggett, J.F., Bustin, S.A., Johnson, M.A., Rook, G., Zumla, A. (2004) Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *Biotechniques* **37**, 112-119

237. Toh, Y., Oki, E., Oda, S., Tokunaga, E., Ohno, S., Maehara, Y., Nicolson, G.L., Sugimachi, K. (1997) Overexpression of the MTA1 gene in gastrointestinal carcinomas: correlation with invasion and metastasis. *International Journal of Cancer* **74**, 459-463

238. Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3**, 0034.1-0034.11

239. Akilesh, S., Shaffer, D.J., Roopenian, D. (2003) Customized Molecular Phenotyping by Quantitative Gene Expression and Pattern Recognition Analysis. *Genome Research* **13**, 1719-1727

240. Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P. (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper- Excel-based tool using pair-wise correlations. *Biotechnology Letters* **26**, 509-515

241. Andersen, C.L., Jensen, J.L., Orntoft, T.F. (2004) Normalization of Real-Time Quantitative Reverse Transcription-PCR Data: A Model-Based Variance Estimation Approach to Identify Genes Suited for Normalization, Applied to Bladder and Colon Cancer Data Sets. *Cancer Research* **64**, 5245-5250

242. Haller, F., Kulle, B., Schwager, S., Gunawan, B., von Heydebreck, A., Sultmann, H., Fuzesi, L. (2004) Equivalence test in quantitative reverse transcription polymerase chain reaction: confirmation of reference genes suitable for normalization. *Analytical Biochemistry* **335**, 1-9

243. Brinkhof, B., Spee, B., Rothuizen, J., Penning, L.C. (2006) Development and evaluation of canine reference genes for accurate quantification of gene expression. *Analytical Biochemistry* **356**, 36-43

244. Saviozzi, S., Cordero, F., Lo, M., Novello, S., Giorgio, V.S., Calogero, R. (2006) Selection of suitable reference genes for accurate normalization of gene expression profile studies in non-small cell lung cancer. *BMC Cancer* **6**, 200

245. de Brouwer, A.P., van Bokhoven, H., Kremer, H. (2006) Comparison of 12 reference genes for normalization of gene expression levels in Epstein-Barr virus-transformed lymphoblastoid cell lines and fibroblasts. *Molecular Diagnosis and Therapeutics* **10**, 197-204

246. Szabo, A., Perou, C.M., Karaca, M., Perreard, L., Quackenbush, J.F., Bernard, P.S. (2004) Statistical modeling for selecting housekeeper genes. *Genome Biology* **5**, R59

247. Wikipedia (Last accessed: 18/7/2007) Website address: http://en.wikipedia.org/wiki/Image:Microarray2.gif

248. Aigner, T., Zien, A., Gehrsitz, A., Gebhard, P.M., McKenna, L. (2001) Anabolic and catabolic gene expression pattern analysis in normal versus osteoarthritic cartilage using complementary DNA-array technology. *Arthritis and Rheumatism* **44**, 2777-2789

249. Aigner, T., Fundel, K., Saas, J., Gebhard, P.M., Haag, J., Weiss, T., Zien, A., Obermayr, F., Zimmer, R., Bartnik, E. (2006) Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. *Arthritis and Rheumatism* **54**, 3533-3544

250. Tew, S.R., Hardingham, T.E. (2006) Regulation of SOX9 mRNA in Human Articular Chondrocytes Involving p38 MAPK Activation and mRNA Stabilization. *Journal of Biological Chemistry* **281**, 39471-39479

251. Zhang, H., Liew, C.C., Marshall, K.W. (2002) Microarray analysis reveals the involvement of beta-2 microglobulin (B2M) in human osteoarthritis. *Osteoarthritis and Cartilage* **10**, 950-960

252. Zien, A., Aigner, T., Zimmer, R., Lengauer, T. (2001) Centralization: a new method for the normalization of gene expression data. *Bioinformatics* **17**, S323-331

253. Bustin, S.A. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**, 169-193

254. Wong, M.L., Medrano, J.F. (2005) Real-time PCR for mRNA quantitation. *Biotechniques* **39**, 75-85

255. Kevorkian, L., Young, D.A., Darrah, C., Donell, S.T., Shepstone, L., Porter, S., Brockbank, S.M., Edwards, D.R., Parker, A.E., Clark, I.M. (2004) Expression profiling of metalloproteinases and their inhibitors in cartilage. *Arthritis and Rheumatism* **50**, 131-141

256. Xing, W., Baylink, D., Kesavan, C., Hu, Y., Kapoor, S., Chadwick, R.B., Mohan, S. (2005) Global gene expression analysis in the bones reveals involvement of several novel genes and pathways in mediating an anabolic response of mechanical loading in mice. *Journal of Cellular Biochemistry* **96**, 1049-1060

257. Zhou, X., Dimachkie, M.M., Xiong, M., Tan, F.K., Arnett, F.C. (2004) cDNA microarrays reveal distinct gene expression clusters in idiopathic inflammatory myopathies. *Medical science monitor : international medical journal of experimental and clinical research* **10**, BR191-BR197

258. Reno, C., Marchuk, L., Sciore, P., Frank, C.B., Hart, D.A. (1997) Rapid isolation of total RNA from small samples of hypocellular, dense connective tissues. *Biotechniques* **22**, 1082-1086

259. Thompson, C.C.M., Clegg, P.D., Carter, S.D. (2001) Differential regulation of gelatinases by transforming growth factor beta-1 in normal equine chondrocytes. *Osteoarthritis and Cartilage* **9**, 325-331

260. Agilent Technologies (Last accessed: 12/3/2006) Website address: www.chem.agilent.com

261. Functional Genomics Core (Last accessed: 12/3/2006) Website address: www.dnaarrays.org/downloads.php

262. Chomczynski, P., Mackey, K. (1995) Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *Biotechniques* **19**, 942-945

263. Aigner, T., Kim, H.A. (2002) Apoptosis and cellular vitality: issues in osteoarthritic cartilage degeneration. *Arthritis and Rheumatism* **46**, 1986-1996

264. Von Euler, H., Khoshnoud, R., He, Q., Khoshnoud, A., Fornander, T., Rutqvist, L.E., Skog, S. (2005) Time-dependent RNA degradation affecting cDNA array quality in spontaneous canine tumours sampled using standard surgical procedures. *International Journal of Molecular Medicine* **16**, 979-985

265. Cancedda, R., Descalzi, C.F., Castagnola, P. (1995) Chondrocyte differentiation. *International Review of Cytology* **159**, 265-358

266. Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research* **29**, e45

267. Schoor, O., Weinschenk, T., Hennenlotter, J., Corvin, S., Stenzl, A., Rammensee, H.G., Stevanovic, S. (2003) Moderate degradation does not preclude microarray analysis of small amounts of RNA. *Biotechniques* **35**, 1192-1201

268. Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., Gaasterland, T., Glenisson, P., Holstege, F.C., Kim, I.F., Markowitz, V., Matese, J.C., Parkinson, H., Robinson, A., Sarkans, U., Schulze-Kremer, S., Stewart, J., Taylor, R., Vilo, J., Vingron, M. (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics* **29**, 365-371

269. Lawrence, R.C., Helmick, C.G., Arnett, F.C., Deyo, R.A., Felson, D.T., Giannini, E.H., Heyse, S.P., Hirsch, R., Hochberg, M.C., Hunder, G.G., Liang, M.H., Pillemer, S.R., Steen, V.D., Wolfe, F. (1998) Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis Rheum* **41**, 778-799

270. Brandt, K.D. (2002) Animal models of osteoarthritis. *Biorheology* **39**, 221-235

271. Singh, K., Masuda, K., An, H.S. (2005) Animal models for human disc degeneration. *The Spine Journal* **5**, 267S-279S

272. Murphy, G., Knauper, V., Atkinson, S., Butler, G., English, W., Hutton, M., Stracke, J., Clark, I. (2002) Matrix metalloproteinases in arthritic disease. *Arthritis Research* **4**, S39-S49

273. Nagase, H., Kashiwagi, M. (2003) Aggrecanases and cartilage matrix degradation. *Arthritis Research and Therapy* **5**, 94-103

274. Blanco Garcia, F.J. (1999) Catabolic events in osteoarthritic cartilage. *Osteoarthritis and Cartilage* **7**, 308-309

275. Baker, A.H., Edwards, D.R., Murphy, G. (2002) Metalloproteinase inhibitors: biological actions and therapeutic opportunities. *Journal of Cell Science* **115**, 3719-3727

276. Clements, D.N., Vaughan-Thomas, A., Peansukmanee, S., Carter, S.D., Innes, J.F., Ollier, W.E.R., Clegg, P.D. (2006) Assessment of the use of RNA quality metrics for the screening of normal and pathological canine articular cartilage samples. *American Journal of Veterinary Research* **67**, 1438-1444

277. Ensembl (Last accessed: 15/6/2007) Website address: www.ensembl.org

278. National Centre for Biotechnology Information (Last accessed: 15/6/2007) Website address: www.ncbi.nlm.nih.gov

279. Roche Diagnostics Ltd (Last accessed: 18/6/2007) Website address: www.roche-applied-science.com

280. Livak, K.J., Schmittgen, T.D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-δδCT Method. *Methods* **25**, 402-408

281. Benjamini, Y., Hochberg, Y. (1995) Controlling the false discovery rate; a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B* **57**, 289-300

282. Dumond, H., Presle, N., Pottie, P., Pacquelet, S., Terlain, B., Netter, P., Gepstein, A., Livne, E., Jouzeau, J.-Y. (2004) Site specific changes in gene expression and cartilage metabolism during early experimental osteoarthritis. *Osteoarthritis and Cartilage* **12**, 284-295

283. Yagi, R., McBurney, D., Laverty, D., Weiner, S., Horton, W.E., Jr. (2005) Intrajoint comparisons of gene expression patterns in human osteoarthritis suggest a change in chondrocyte phenotype. *Journal of Orthopaedic Research* **23**, 1128-1138

284. Eid, K., Thornhill, T.S., Glowacki, J. (2006) Chondrocyte gene expression in osteoarthritis: Correlation with disease severity. *Journal of Orthopaedic Research* **24**, 1062-1068

285. Diaz-Romero, J., Gaillard, J.P., Grogan, S.P., Nesic, D., Trub, T., Mainil-Varlet, P. (2005) Immunophenotypic analysis of human articular

chondrocytes: changes in surface markers associated with cell expansion in monolayer culture. *Journal of Cellular Physiology* **202**, 731-742

286. Matyas, J.R., Ehlers, P.F., Huang, D., Adams, M.E. (1999) The early molecular natural history of experimental osteoarthritis. I. Progressive discoordinate expression of aggrecan and type II procollagen messenger RNA in the articular cartilage of adult animals. *Arthritis and Rheumatism* **42**, 993-1002

287. Lorenz, H., Wenz, W., Ivancic, M., Steck, E., Richter, W. (2005) Early and stable upregulation of collagen type II, collagen type I and YKL40 expression levels in cartilage during early experimental osteoarthritis occurs independent of joint location and histological grading. *Arthritis Research and Therapy* **7**, R156-R165

288. Adams, M.E., Matyas, J.R., Huang, D., Dourado, G.S. (1995) Expression of proteoglycans and collagen in the hypertrophic phase of experimental osteoarthritis. *Journal of Rheumatology Supplement* **43**, 94-97

289. Altman, R.D., Kates, J., Chun, L.E., Dean, D.D., Eyre, D. (1992) Preliminary observations of chondral abrasion in a canine model. *Annals of the Rheumatic Diseases* **51**, 1056-1062

290. Benjamin, M., Ralphs, J.R. (2004) Biology of fibrocartilage cells. *International Review of Cytology* **233**, 1-45

291. Nishida, Y., Shinomura, T., Iwata, H., Miura, T., Kimata, K. (1994) Abnormal occurrence of a large chondroitin sulfate proteoglycan, PG-M/versican in osteoarthritic cartilage. *Osteoarthritis and Cartilage* **2**, 43-49

292. Cs-Szabo, G., Melching, L.I., Roughley, P.J., Glant, T.T. (1997) Changes in messenger RNA and protein levels of proteoglycans and link protein in human osteoarthritic cartilage samples. *Arthritis and Rheumatism* **40**, 1037-1045

293. Pfander, D., Heinz, N., Rothe, P., Carl, H.D., Swoboda, B. (2004) Tenascin and aggrecan expression by articular chondrocytes is influenced by interleukin 1{beta}: a possible explanation for the changes in matrix synthesis during osteoarthritis. *Annals of the Rheumatic Diseases* **63**, 240-244

294. Stokes, D.G., Liu, G., Coimbra, I.B., Piera-Velazquez, S., Crowl, R.M., Jimenez, S.A. (2002) Assessment of the gene expression profile of differentiated and dedifferentiated human fetal chondrocytes by microarray analysis. *Arthritis and Rheumatism* **46**, 404-419

295. Gouze, J.N., Gouze, E., Popp, M., Bush, M., Dacanay, E., Kay, J., Levings, P., Patel, K., Saran, J.P., Watson, R., Ghivizzani, S. (2006) Exogenous glucosamine globally protects chondrocytes from the arthritogenic effects of IL-1beta. *Arthritis Research & Therapy* **8**, R173

296. Clements, D.N., Carter, S.D., Innes, J.F., Ollier, W.E., Day, P.J. (2006) Analysis of normal and osteoarthritic canine cartilage mRNA expression by quantitative-PCR. *Arthritis Research and Therapy* **8**, R158

297. Goff, L.A., Bowers, J., Schwalm, J., Howerton, K., Getts, R.C., Hart, R.P. (2004) Evaluation of sense-strand mRNA amplification by comparative quantitative PCR. *BMC Genomics* **5**, 76

298. Jones, P., Jones, C., Fretwell, N., Martin, A., Soloviev, M.Design and production of a whole genome dog oligonucleotide microarray. In: Advances in Canine and Feline Genomics

299. Workman, C., Jensen, L.J., Jarmer, H., Berka, R., Gautier, L., Nielser, H.B., Saxild, H.H., Nielsen, C., Brunak, S., Knudsen, S. (2002) A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biology* **3**, 0048.1-.16

300. Sato, T., Konomi, K., Yamasaki, S., Aratani, S., Tsuchimochi, K., Yokouchi, M., Masuko-Hongo, K., Yagishita, N., Nakamura, H., Komiya, S., Beppu, M., Aoki, H., Nishioka, K., Nakajima, T. (2006) Comparative analysis of gene expression profiles in intact and damaged regions of human osteoarthritic cartilage. *Arthritis and Rheumatism* **54**, 808-817

301. Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences* **95**, 14863-14868

302. Vastrik, I., D'Eustachio, P., Schmidt, E., Joshi-Tope, G., Gopinath, G., Croft, D., de Bono, B., Gillespie, M., Jassal, B., Lewis, S., Matthews, L., Wu, G., Birney, E., Stein, L. (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biology* **8**, R39

303. Wessa P (2006) Free Statistics Software, Office for Research Development and Education, Version 1.1.20. http://www.wessa.net/

304. Sanna, M.G., da Silva Correia, J., Luo, Y., Chuang, B., Paulson, L.M., Nguyen, B., Deveraux, Q.L., Ulevitch, R.J. (2002) ILPIP, a Novel Anti-apoptotic Protein That Enhances XIAP-mediated Activation of JNK1 and Protection against Apoptosis. *Journal of Biological Chemistry* **277**, 30454-30462

305. Blanco, F.J., Guitian, R., Vazquez-Martul, E., de Toro, F.J., Galdo, F. (1998) Osteoarthritis chondrocytes die by apoptosis. A possible pathway for osteoarthritis pathology. *Arthritis and Rheumatism* **41**, 284-289

306. Antoine, K., Ferbus, D., Kolahgar, G., Prosperi, M.T., Goubin, G. (2005) Zinc finger protein overexpressed in colon carcinoma interacts with the telomeric protein hRap1. *Journal of Cellular Biochemistry* **95**, 763-768

307. Martin, J.A., Buckwalter, J.A. (2001) Telomere Erosion and Senescence in Human Articular Cartilage Chondrocytes. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **56**, B172-B179

308. Chan, C., Moore, B.E., Cotman, C.W., Okano, H., Tavares, R., Hovanesian, V., Pinar, H., Johanson, C.E., Svendsen, C.N., Stopa, E.G. (2006) Musashi1 antigen expression in human fetal germinal matrix development. *Experimental Neurology* **201**, 515-518

309. Greig, A.V., Linge, C., Terenghi, G., McGrouther, D.A., Burnstock, G. (2003) Purinergic receptors are part of a functional signaling system for proliferation and differentiation of human epidermal keratinocytes. *Journal of Investigative Dermatology* **120**, 1007-1015

310. Seibold, S., Rudroff, C., Weber, M., Galle, J., Wanner, C., Marx, M. (2003) Identification of a new tumor suppressor gene located at chromosome 8p21.3-22. *The FASEB Journal* **17**, 1180-1182

311. Dowthwaite, G.P., Bishop, J.C., Redman, S.N., Khan, I.M., Rooney, P., Evans, D.J.R., Haughton, L., Bayram, Z., Boyer, S., Thomson, B., Wolfe, M.S., Archer, C.W. (2004) The surface of articular cartilage contains a progenitor cell population. *Journal of Cell Science* **117**, 889-897

312. Gebhard, P.M., Gehrsitz, A., Bau, B., Soder, S., Eger, W., Aigner, T. (2003) Quantification of expression levels of cellular differentiation markers does not support a general shift in the cellular phenotype of osteoarthritic chondrocytes. *Journal of Orthopaedic Research* **21**, 96-101

313. Oda, T., Fukuda, A., Hagiwara, H., Masuho, Y., Muramatsu, M.A., Hisatake, K., Yamashita, T. (2004) ABT1-associated protein (ABTAP), a novel nuclear protein conserved from yeast to mammals, represses transcriptional activation by ABT1. *Journal of Cellular Biochemistry* **93**, 788-806

314. Chung, H.J., Liu, J., Dundr, M., Nie, Z., Sanford, S., Levens, D. (2006) FBPs Are Calibrated Molecular Tools To Adjust Gene Expression. *Molecular and Cellular Biology* **26**, 6584-6597

315. Gregson, H.C., Van Hooser, A.A., Ball, A.R., Jr., Brinkley, B.R., Yokomori, K. (2002) Localization of human SMC1 protein at kinetochores. *Chromosome Research* **10**, 267-277

316. Dozin, B., Malpeli, M., Camardella, L., Cancedda, R., Pietrangelo, A. (2002) Response of young, aged and osteoarthritic human articular chondrocytes to inflammatory cytokines: molecular and cellular aspects. *Matrix Biology* **21**, 449-459

317. Stove, J., Gerlach, C., Huch, K., Gunther, K.P., Brenner, R., Puhl, W., Scharf, H.P. (2001) Gene expression of stromelysin and aggrecan in osteoarthritic cartilage. *Pathobiology* **69**, 333-338

318. Stove, J., Gremmes, C., Gunther, K.P., Scharf, H.P., Schwarz, M. (2006) Metabolic activity and gene expression of osteoarthritic chondrocytes in correlation with radiological and histological characteristics. *Biomedicine and Pharmacotherapy* **60**, 644-647

319. Clements, D.N., Carter, S.D., Innes, J.F., Ollier, W.E.R., Day, P.J. (2007) Gene expression profiling of normal and ruptured canine anterior cruciate ligaments. *Osteoarthritis and Cartilage* **15**, *In Press*

320. Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A., Ploner, A. (2005) False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* **21**, 3017-3024

321. Jung, S.H., Jang, W. (2006) How accurately can we control the FDR in analyzing microarray data? *Bioinformatics* **22**, 1730-1736

322. Van Ryssen, B., van Bree, H. (1997) Arthroscopic findings in 100 dogs with elbow lameness. *Vet Rec* **140**, 360-362

323. Dedrick, D.K., Goulet, R.W., O'Connor, B.L., Brandt, K.D. (1997) Preliminary report: increased porosity of the subchondral plate in an accelerated canine model of osteoarthritis. *Osteoarthritis and Cartilage* **5**, 71-74

324. Fazzalari, N.L., Forwood, M.R., Manthey, B.A., Smith, K., Kolesik, P. (1998) Three-dimensional confocal images of microdamage in cancellous bone. *Bone* **23**, 373-378

325. Watson, P.J., Hall, L.D., Malcolm, A., Tyler, J.A. (1996) Degenerative joint disease in the guinea pig. Use of magnetic resonance imaging to monitor progression of bone pathology. *Arthritis and Rheumatism* **39**, 1327-1337

326. Hopwood, B., Gronthos, S., Kuliwaba, J.S., Robey, P.G., Findlay, D.M., Fazzalari, N.L. (2005) Identification of differentially expressed genes between osteoarthritic and normal trabecular bone from the intertrochanteric region of the proximal femur using cDNA microarray analysis. *Bone* **36**, 635-644

327. Stoker, A., Cook, J., Kuroki, K., Fox, D. (2006) Site-specific analysis of gene expression in early osteoarthritis using the Pond-Nuki model in dogs. *Journal of Orthopaedic Surgery and Research* **1**, 8

328. Crouch, D.T., Cook, J.L., Lewis, D.D., Kreeger, J.M., Tomlinson, J.L. (2000) The presence of collagen types II and X in medial coronoid processes of 21 dogs. *Veterinary and Comparative Orthopaedics and Traumatology* **13**, 178-184

329. Carvalho, R.S., Einhorn, T.A., Lehmann, W., Edgar, C., Al-Yamani, A., Apazidis, A., Pacicca, D., Clemens, T.L., Gerstenfeld, L.C. (2004) The role of angiogenesis in a murine tibial model of distraction osteogenesis. *Bone* **34**, 849-861

330. Gilbertson, E.M. (1975) Development of periarticular osteophytes in experimentally induced osteoarthritis in the dog. A study using microradiographic, microangiographic, and fluorescent bone-labelling techniques. *Annals of the Rheumatic Diseases* **34**, 12-25

331. Invitrogen (Last accessed: 29/10/2006) Website address: www.invitrogen.com

332. Mohammadi, M., Day, P.J.R. (2004) Oligonucleotides used as template calibrators for general application in quantitative polymerase chain reaction. *Analytical Biochemistry* **335**, 299-304

333. Gordon, W.J., Besancon, M.F., Conzemius, M.G., Miles, K.G., Kapatkin, A.S., Culp, W.T.N. (2003) Frequency of post-traumatic osteoarthritis in dogs after repair of a humeral condylar fracture. *Veterinary and Comparative Orthopaedics and Traumatology* **16**, 1-5

334. Gordon, W.J., Conzemius, M.C., Riedesel, E., Besancon, M.F., Evans, R., Wilke, V., Ritter, M.J. (2003) The relationship between limb function and radiographic osteoarthrosis in dogs with stifle osteoarthrosis. *Veterinary Surgery* **32**, 451-454

335. Fujita, Y., Hara, Y., Nezu, Y., Yamaguchi, S., Schulz, K.S., Tagawa, M. (2005) Direct and indirect markers of cartilage metabolism in synovial fluid obtained from dogs with hip dysplasia and correlation with clinical and radiographic variables. *American Journal of Veterinary Research* **66**, 2028-2033

336. Gouttenoire, J., Valcourt, U., Ronziere, M.C., Aubert-Foucher, E., Mallein-Gerin, F., Herbage, D. (2004) Modulation of collagen synthesis in normal and osteoarthritic cartilage. *Biorheology* **41**, 535-542

337. Bau, B., Gebhard, P.M., Haag, J., Knorr, T., Bartnik, E., Aigner, T. (2002) Relative messenger RNA expression profiling of collagenases and aggrecanases in human articular chondrocytes in vivo and in vitro. *Arthritis and Rheumatism* **46**, 2648-2657

338. Davidson, R., Waters, J., Kevorkian, L., Darrah, C., Cooper, A., Donell, S., Clark, I. (2006) Expression profiling of metalloproteinases and their inhibitors in synovium and cartilage. *Arthritis Research & Therapy* **8**, R124

339. Flannelly, J., Chambers, M.G., Dudhia, J., Hembry, R.M., Murphy, G., Mason, R.M., Bayliss, M.T. (2002) Metalloproteinase and tissue inhibitor of metalloproteinase expression in the murine STR/ort model of osteoarthritis. *Osteoarthritis and Cartilage* **10**, 722-733

340. Salinardi, B.J., Roush, J.K., Schermerhorn, T., Mitchell, K.E. (2006) Matrix metalloproteinase and tissue inhibitor of metalloproteinase in serum and synovial fluid of osteoarthritic dogs. *Veterinary and Comparative Orthopaedics and Traumatology* **19**, 49-55

341. Kuroki, K., Cook, J.L., Stoker, A.M., Turnquist, S.E., Kreeger, J.M., Tomlinson, J.L. (2005) Characterizing osteochondrosis in the dog: potential roles for matrix metalloproteinases and mechanical load in pathogenesis and disease progression. *Osteoarthritis and Cartilage* **13**, 225-234

342. Panula, H.E., Lohmander, L.S., Ronkko, S., Agren, U., Helminen, H.J., Kiviranta, I. (1998) Elevated levels of synovial fluid PLA2, stromelysin (MMP-3) and TIMP in early osteoarthrosis after tibial valgus osteotomy in young beagle dogs. *Acta Orthopaedica Scandinavica* **69**, 152-158

343. Mansell, J.P., Tarlton, J.F., Bailey, A.J. (1997) Biochemical evidence for altered subchondral bone collagen metabolism in osteoarthritis of the hip. *Rheumatology* **36**, 16-19

344. von Porat, A., Roos, E.M., Roos, H. (2004) High prevalence of osteoarthritis 14 years after an anterior cruciate ligament tear in male soccer players: a study of radiographic and patient relevant outcomes. *Annals of the Rheumatic Diseases* **63**, 269-273

345. Kannus, P., Jarvinen, M. (1989) Posttraumatic anterior cruciate ligament insufficiency as a cause of osteoarthritis in a knee joint. *Clinical Rheumatology* **8**, 251-260

346. Jacobsen, K. (1977) Osteoarthrosis following insufficiency of the cruciate ligaments in man. A clinical study. *Acta Orthopaedica Scandinavica* **48**, 520-526

347. Lo, I.K., Marchuk, L.L., Hart, D.A., Frank, C.B. (1998) Comparison of mRNA levels for matrix molecules in normal and disrupted human anterior cruciate ligaments using reverse transcription-polymerase chain reaction. *Journal of Orthopaedic Research* **16**, 421-428

348. Muir, P., Danova, N.A., Argyle, D.J., Manley, P.A., Hao, Z. (2005) Collagenolytic Protease Expression in Cranial Cruciate Ligament and Stifle Synovial Fluid in Dogs with Cranial Cruciate Ligament Rupture. *Veterinary Surgery* **34**, 482-490

349. Comerford, E.J., Innes, J.F., Tarlton, J.F., Bailey, A.J. (2004) Investigation of the composition, turnover, and thermal properties of ruptured cranial cruciate ligaments of dogs. *American Journal of Veterinary Research* **65**, 1136-1141

350. Molloy, T.J., Wang, Y., Horner, A., Skerry, T.M., Murrell, G.A. (2006) Microarray analysis of healing rat Achilles tendon: evidence for glutamate signaling mechanisms and embryonic gene expression in healing tendon tissue. *Journal of Orthopaedic Research* **24**, 842-855

351. TreeView (Last accessed: 12/12/2006) Website address: http://rana.lbl.gov/EisenSoftware.htm

352. Amiel, D., Ishizue, K.K., Harwood, F.L., Kitabayashi, L., Akeson, W.H. (1989) Injury of the anterior cruciate ligament: the role of collagenase in ligament degeneration. *Journal of Orthopaedic Research* **7**, 486-493

353. Feagin, J.A., Jr., Curl, W.W. (1976) Isolated tear of the anterior cruciate ligament: 5-year follow-up study. *American Journal of Sports Medicine* **4**, 95-100

354. Comerford, E.J., Tarlton, J.F., Avery, N.C., Bailey, A.J., Innes, J.F. (2005) Distal femoral intercondylar notch dimensions and their relationship to composition and metabolism of the canine anterior cruciate ligament. *Osteoarthritis and Cartilage* **13**, In press

355. Tremble, P.M., Lane, T.F., Sage, E.H., Werb, Z. (1993) SPARC, a secreted protein associated with morphogenesis and tissue remodeling, induces expression of metalloproteinases in fibroblasts through a novel extracellular matrix-dependent pathway. *The Journal of Cell Biology* **121**, 1433-1444

356. Everts, V., van der, Z.E., Creemers, L., Beertsen, W. (1996) Phagocytosis and intracellular digestion of collagen, its role in turnover and remodelling. *Histochemical Journal* **28**, 229-245

357. Yamaguchi, M., Ozawa, Y., Nogimura, A., Aihara, N., Kojima, T., Hirayama, Y., Kasai, K. (2004) Cathepsins B and L increased during response of periodontal ligament cells to mechanical stress in vitro. *Connective Tissue Research* **45**, 181-189

358. Goldberg, V.M., Burstein, A., Dawson, M. (1982) The influence of an experimental immune synovitis on the failure mode and strength of the rabbit anterior cruciate ligament. *Journal of Bone and Joint Surgery (American)* **64**, 900-906

359. Hart, D.A., Sciore, P., Boykiw, R., Reno, C. (1998) Pregnancy induces complex changes in the the pattern of mRNA expression in knee ligaments of the adolescent rabbit. *Matrix Biology* **17**, 21-34

360. Sarasa-Renedo, A., Chiquet, M. (2005) Mechanical signals regulating extracellular matrix gene expression in fibroblasts. *Scand J Med Sci Sports* **15**, 223-230

361. Molloy, T., Wang, Y., Murrell, G. (2003) The roles of growth factors in tendon and ligament healing. *Sports Medicine* **33**, 381-394

362. Murray, M.M., Martin, S.D., Martin, T.L., Spector, M. (2000) Histological Changes in the Human Anterior Cruciate Ligament After Rupture. *Journal of Bone and Joint Surgery* **82**, 1387

363. Barrett, J.G., Hao, Z., Graf, B.K., Kaplan, L.D., Heiner, J.P., Muir, P. (2005) Inflammatory changes in ruptured canine cranial and human anterior cruciate ligaments. *American Journal of Veterinary Research* **66**, 2073-2080

364. Ireland, M.L. (2002) The female ACL: why is it more prone to injury? *Orthopedic Clinics of North America* **33**, 637-651

365. Muir, P., Schamberger, G.M., Manley, P.A., Hao, Z. (2005) Localization of Cathepsin K and Tartrate-Resistant Acid Phosphatase in Synovium and Cranial Cruciate Ligament in Dogs with Cruciate Disease. *Veterinary Surgery* **34**, 239-246

366. Radonic, A., Thulke, S., Bae, H.G., Muller, M.A., Siegert, W., Nitsche, A. (2005) Reference gene selection for quantitative real-time PCR analysis in virus infected cells: SARS corona virus, Yellow fever virus, Human Herpesvirus-6, Camelpox virus and Cytomegalovirus infections. *Virology Journal* **2**, 7

367. Spinsanti, G., Panti, C., Lazzeri, E., Marsili, L., Casini, S., Frati, F., Fossi, C. (2006) Selection of reference genes for quantitative RT-PCR studies in striped dolphin (Stenella coeruleoalba) skin biopsies. *BMC Molecular Biology* **7**, 32

368. Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P. (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnology Letters* **26**, 509-515

369. Ayers, D., Clements, D., Salway, F., Day, P. (2007) Expression stability of commonly used reference genes in canine articular connective tissues. *BMC Veterinary Research* **3**, 7

370. Ayers, D., Clements, D.N., Salway, F., Day, P.J.R. (2006) Expression stability of commonly used Control Genes in Canine Articular Connective Tissues. *Submitted for publication*

371. Gorzelniak, K., Janke, J., Engeli, S., Sharma, A.M. (2001) Validation of endogenous controls for gene expression studies in human adipocytes and preadipocytes. *Hormone and Metabolic Research* **33**, 625-627

372. Matyas, J.R., Huang, D., Adams, M.E. (1999) A comparison of various "housekeeping" probes for northern analysis of normal and osteoarthritic articular cartilage RNA. *Connective Tissue Research* **40**, 163-172

373. Bogaert, L., Van Poucke, M., De Baere, C., Peelman, L., Gasthuys, F., Martens, A. (2006) Selection of a set of reliable reference genes for quantitative real-time PCR in normal equine skin and in equine sarcoids. *BMC Biotechnology* **6**, 24

374. Cavdar Koc, E., Blackburn, K., Burkhart, W., Spremulli, L.L. (1999) Identification of a Mammalian Mitochondrial Homolog of Ribosomal Protein S7. *Biochemical and Biophysical Research Communications* **266**, 141-146

375. Listgarten, J., Graham, K., Damaraju, S., Cass, C., Mackey, J., Zanke, B. (2003) Clinically validated benchmarking of normalisation techniques for two-colour oligonucleotide spotted microarray slides. *Applied Bioinformatics* **2**, 219-228

376. Yu, H., Luscombe, N.M., Qian, J., Gerstein, M. (2003) Genomic analysis of gene expression relationships in transcriptional regulatory networks. *Trends in Genetics* **19**, 422-427

377. Lohi, H., Young, E.J., Fitzmaurice, S.N., Rusbridge, C., Chan, E.M., Vervoort, M., Turnbull, J., Zhao, X.C., Ianzano, L., Paterson, A.D., Sutter, N.B., Ostrander, E.A., Andre, C., Shelton, G.D., Ackerley, C.A., Scherer, S.W., Minassian, B.A. (2005) Expanded Repeat in Canine Epilepsy. *Science* **307**, 81

378. Drogemuller, C., Philipp, U., Haase, B., Gunzel-apel, A.R., Leeb, T. (2007) A Noncoding Melanophilin Gene (MLPH) SNP at the Splice Donor of Exon 1 Represents a Candidate Causal Mutation for Coat Color Dilution in Dogs. *Journal of Heredity* **98**, *In press*

379. Sutter, N.B., Bustamante, C.D., Chase, K., Gray, M.M., Zhao, K., Zhu, L., Padhukasahasram, B., Karlins, E., Davis, S., Jones, P.G., Quignon, P., Johnson, G.S., Parker, H.G., Fretwell, N., Mosher, D.S., Lawler, D.F., Satyaraj, E., Nordborg, M., Lark, K.G., Wayne, R.K., Ostrander, E.A. (2007) A single IGF1 allele is a major determinant of small size in dogs. *Science* **316**, 112-115

380. Sarmiento, U.M., Storb, R.F. (1988) Restriction fragment length polymorphism of the major histocompatibility complex of the dog. *Immunogenetics* **28**, 117-124

381. Pfeiffer, I., Volkel, I., Taubert, H., Brenig, B. (2004) Forensic DNA-typing of dog hair: DNA-extraction and PCR amplification. *Forensic Science International* **141**, 149-151

382. Oberbauer, A.M., Grossman, D.I., Eggleston, M.L., Irion, D.N., Schaffer, A.L., Pedersen, N.C., Belanger, J.M. (2003) Alternatives to Blood as a Source of DNA for Large-scale Scanning Studies of Canine Genome Linkages. *Veterinary Research Communications* **27**, 27-38

383. Yue, G.H., Orban, L. (2005) A simple and affordable method for high-throughput DNA extraction from animal tissues for polymerase chain reaction. *Electrophoresis* **26**, 3081-3083

384. van der Hel, O.L., van der Luijt, R.B., Bas Bueno de Mesquita, H., van Noord, P.A.H., Slothouber, B., Roest, M., van der Schouw, Y.T., Grobbee, D.E., Pearson, P.L., Peeters, P.H.M. (2002) Quality and Quantity of DNA Isolated from Frozen Urine in Population-Based Research. *Analytical Biochemistry* **304**, 206-211

385. Short, A.D., Kennedy, L.J., Forman, O., Barnes, A., Fretwell, N., Wiggall, R., Thomson, W., Ollier, W.E.R. (2005) Canine DNA Subjected to Whole Genome Amplification is Suitable for a Wide Range of Molecular Applications. *Journal of Heredity* **96**, 829-835

386. Iovannisci, D.M., Ha, T.T., Shaw, G.M. (2006) Recovery of genomic DNA from residual frozen archival blood clots suitable for amplification and use in genotyping assays. *Genetic Testing* **10**, 44-49

387. UK DNA Archive for Companion Animals (Last accessed: 25/6/2007) Website address: http://pcwww.liv.ac.uk/DNA_Archive_for_Companion_Animals

388. Cornell Medical Genetic Archive (Last accessed: 15/6/2007) Website address: http://www.vet.cornell.edu/research/dnabank/intro.htm

389. Kramvis, A., Bukofzer, S., Kew, M.C. (1996) Comparison of hepatitis B virus DNA extractions from serum by the QIAamp blood kit, GeneReleaser, and the phenol-chloroform method. *Journal of Clinical Microbiology* **34**, 2731-2733

390. Flagstad, Roed, Stacy, Jakobsen. (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology* **8**, 879-883

391. Fahle, G.A., Fischer, S.H. (2000) Comparison of Six Commercial DNA Extraction Kits for Recovery of Cytomegalovirus DNA from Spiked Human Specimens. *Journal of Clinical Microbiology* **38**, 3860-3863

392. Deagle, B., Eveson, J.P., Jarman, S. (2006) Quantification of damage in DNA recovered from highly degraded samples - a case study on DNA in faeces. *Frontiers in Zoology* **3**, 11

393. Steffan, R.J., Goksoyr, J., Bej, A.K., Atlas, R.M. (1988) Recovery of DNA from soils and sediments. *Applied and Environmental Microbiology* **54**, 2908-2915

394. Haque, K., Pfeiffer, R., Beerman, M., Struewing, J., Chanock, S., Bergen, A. (2003) Performance of high-throughput DNA quantification methods. *BMC Biotechnology* **3**, 20

395. Qiagen (Last accessed: 25/6/2007) Website address: www1.qiagen.com

396. Gentra (Last accessed: 15/6/2007) Website address: www.gentra.com

397. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715

398. Wilke, V.L., Conzemius, M.C., Rothschild, M.F. (2005) SNP detection and association analyses of candidate genes for rupture of the cranial cruciate ligament in the dog. *Animal Genetics* **36**, 519-521

399. Valdes, A.M., Loughlin, J., Oene, M.V., Chapman, K., Surdulescu, G.L., Doherty, M., Spector, T.D. (2007) Sex and ethnic differences in the association of ASPN, CALM1, COL2A1, COMP, and FRZB with genetic susceptibility to osteoarthritis of the knee. *Arthritis and Rheumatism* **56**, 137-146

400. Spiegelman, J.I., Mindrinos, M.N., Oefner, P.J. (2000) High-accuracy DNA sequence variation screening by DHPLC. *Biotechniques* **29**, 1084-90, 1092

401. Short, A.D. The genetics of canine diabetes: A candidate gene study. Thesis sumitted for the degree of PhD at The University of Manchester; 2006

402. Broad Institute Canine SNP Database (Last accessed: 26/6/2007) Website address: http://www.broad.mit.edu/tools/data/genvar.html

403. Short, A.D., Catchpole, B., Kennedy, L.J., Barnes, A., Fretwell, N., Jones, C., Thomson, W., Ollier, W.E.R. (2007) Analysis of Candidate Susceptibility Genes in Canine Diabetes. *Journal of Heredity* **98**, 518-525

404. Sequenom Application Note (Last accessed: 3/7/2007) Website address: http://www.sequenom.com/Assets/pdfs/appnotes/8876-006.pdf

405. BCgene (Last accessed: 2/7/2007) Website address: www.bcplatforms.com

406. Sham, P.C., Curtis, D. (1995) Monte Carlo tests for associations between disease and alleles at highly polymorphic loci. *Annals of Human Genetics* **59**, 97-105

407. CLUMP (Last accessed: 2/7/2007) Website address: http://www.smd.qmul.ac.uk/statgen/dcurtis/software.html

408. Odd Ratio Calculator (Last accessed: 9/7/2007) Website address: http://www.hutchon.net/ConfidOR.htm

409. Simple Interactive Statistical Analysis (Last accessed: 16/7/2007) Website address: http://home.clara.net/sisa/

410. Petersen-Jones, S. (2005) Advances in the molecular understanding of canine retinal diseases. *Journal of Small Animal Practice* **46**, 371-380

411. Aguirre, G.K., Komaromy, A.M., Cideciyan, A.V., Brainard, D.H., Aleman, T.S., Roman, A.J., Avants, B.B., Gee, J.C., Korczykowski, M., Hauswirth, W.W., Acland, G.M., Aguirre, G.D., Jacobson, S.G. (2007) Canine and human visual cortex intact and responsive despite early retinal blindness from RPE65 mutation. *PLoS Medicine* **4**, e230

412. Bendrups, A., Hilton, A., Meager, A., Hamilton, J.A. (1993) Reduction of tumor necrosis factor alpha and interleukin-1 beta levels in human synovial tissue by interleukin-4 and glucocorticoid. *Rheumatology International* **12**, 217-220

413. Sakkas, L.I., Johanson, N.A., Scanzello, C.R., Platsoucas, C.D. (1998) Interleukin-12 Is Expressed by Infiltrating Macrophages and Synovial Lining Cells in Rheumatoid Arthritis and Osteoarthritis. *Cellular Immunology* **188**, 105-110

414. Maccoux, L.J., Salway, F., Day, P.J.R., Clements, D.N. (2007) Expression profiling of select cytokines in canine osteoarthritis tissues. *Veterinary Immunology and Immunopathology* **118**, 59-67

415. Short, A.D., Kennedy, L.J., Barnes, A., Fretwell, N., Jones, C., Thomson, W., Ollier, W.E.R. (2007) Hardy-Weinberg Expectations in Canine Breeds: Implications for genetic studies. *Journal of Heredity* **98**, esm020

416. Kardel, R., Ulfgren, A.-K., Reinholt, F.P., Holmlund, A. (2003) Inflammatory cell and cytokine patterns in patients with painful clicking and osteoarthritis in the temporomandibular joint. *International Journal of Oral and Maxillofacial Surgery* **32**, 390-396

417. Mahr, S., Menard, J., Krenn, V., Muller, B. (2003) Sexual dimorphism in the osteoarthritis of STR/ort mice may be linked to articular cytokines. *Annals of the Rheumatic Diseases* **62**, 1234-1237

418. Dumond, H., Presle, N., Terlain, B., Mainard, D., Loeuille, D., Netter, P., Pottie, P. (2003) Evidence for a key role of leptin in osteoarthritis. *Arthritis and Rheumatism* **48**, 3118-3129

419. Hegemann, N., Wondimu, A., Ullrich, K., Schmidt, M.F.G. (2003) Synovial MMP-3 and TIMP-1 levels and their correlation with cytokine expression in canine rheumatoid arthritis. *Veterinary Immunology and Immunopathology* **91**, 199-204

420. Salter, D.M. (1993) Tenascin is increased in cartilage and synovium from arthritic knees. *British Journal of Rheumatology* **32**, 780-786

421. Hegemann, N., Wondimu, A., Kohn, B., Brunnberg, L., Schmidt, M.F. (2005) Cytokine profile in canine immune-mediated polyarthritis and osteoarthritis. *Veterinary and Comparative Orthopaedics and Traumatology* **18**, 67-72

422. Peters, I.R., Helps, C.R., Hall, E.J., Day, M.J. (2004) Real-time RT-PCR: considerations for efficient and sensitive assay design. *Journal of Immunological Methods* **286**, 203-217

423. Nolan, T., Hands, R.E., Bustin, S.A. (2006) Quantification of mRNA using real-time RT-PCR. *Nature Protocols* **1**, 1559-1582

424. Fleige, S., Pfaffl, M.W. (2006) RNA integrity and the effect on the real-time qRT-PCR performance. *Molecular Aspects of Medicine* **27**, 126-139

425. Bower, N.I., Moser, R.J., Hill, J.R., Lehnert, S.A. (2007) Universal reference method for real-time PCR gene expression analysis of preimplantation embryos. *Biotechniques* **42**, 199-206

426. Day, P.J. (2006) Miniaturization applied to analysis of nucleic acids in heterogeneous tissues. *Expert Reviews in Molecular Diagnostics* **6**, 23-28

427. Kemming, D., Vogt, U., Tidow, N., Schlotter, C.M., Burger, H., Helms, M.W., Korsching, E., Granetzny, A., Boseila, A., Hillejan, L., Marra, A., Ergonenc, Y., Adiguzel, H., Brandt, B. (2006) Whole genome expression analysis for biologic rational pathway modeling: application in cancer

prognosis and therapy prediction. *Molecular Diagnosis and Therapeutics* **10**, 271-280

428. Kuffner, R., Fundel, K., Zimmer, R. (2005) Expert knowledge without the expert: integrated analysis of gene expression and literature to derive active functional contexts. *Bioinformatics* **21**, ii259-ii267

429. Bee, A., Barnes, A., Jones, M., Robertson, D.H., Clegg, P.D., Carter, S.D. (2000) Canine TIMP-2: Purification, Characterization and Molecular Detection. *The Veterinary Journal* **160**, 126-134

430. Bigg, H.F., Rowan, A.D. (2001) The inhibition of metalloproteinases as a therapeutic target in rheumatoid arthritis and osteoarthritis. *Current Opinion in Pharmacology* **1**, 314-320

431. Campbell, S.E., Nasir, L., Argyle, D.J., Gault, E.A., Duthie, S., Bennett, D. (2001) Cloning of canine IL-1ra, TNFR and TIMP-2. *Veterinary Immunology and Immunopathology* **78**, 207-214

432. Munshi, H.G., Wu, Y.I., Mukhopadhyay, S., Ottaviano, A.J., Sassano, A., Koblinski, J.E., Platanias, L.C., Stack, M.S. (2004) Differential Regulation of Membrane Type 1-Matrix Metalloproteinase Activity by ERK 1/2- and p38 MAPK-modulated Tissue Inhibitor of Metalloproteinases 2 Expression Controls Transforming Growth Factor β 1-induced Pericellular Collagenolysis. *Journal of Biological Chemistry* **279**, 39042-39050

433. Johansson, N., Ala-aho, R., Uitto, V., Grenman, R., Fusenig, N.E., Lopez-Otin, C., Kahari, V.M. (2000) Expression of collagenase-3 (MMP-13) and collagenase-1 (MMP-1) by transformed keratinocytes is dependent on the activity of p38 mitogen-activated protein kinase. *Journal of Cell Science* **113**, 227-235

434. JustBio Aligning Tool (Last accessed: 26/6/2007) Website address: www.justbio.com/aligner/index.php

435. Uitterlinden, A.G., Fang, Y., van Meurs, J.B.J., Pols, H.A.P., van Leeuwen, J.P.T.M. (2004) Genetics and biology of vitamin D receptor polymorphisms. *Gene* **338**, 143-156

436. Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M., Venter, J.C. (2003) The Dog Genome: Survey Sequencing and Comparative Analysis. *Science* **301**, 1898-1903

437. Irizarry, K., Kustanovich, V., Li, C., Brown, N., Nelson, S., Wong, W., Lee, C.J. (2000) Genome-wide analysis of single-nucleotide polymorphisms in human expressed sequences. *Nature Genetics* **26**, 233-236

438. Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., Kruglyak, L. (2004) Genetic Structure of the Purebred Domestic Dog. *Science* **304**, 1160-1164

439. Altshuler, D., Pollara, V.J., Cowles, C.R., Van Etten, W.J., Baldwin, J., Linton, L., Lander, E.S. (2000) An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* **407**, 513-516

440. Marth, G., Yeh, R., Minton, M., Donaldson, R., Li, Q., Duan, S., Davenport, R., Miller, R.D., Kwok, P.Y. (2001) Single-nucleotide polymorphisms in the public domain: how useful are they? *Nature Genetics* **27**, 371-372

Appendices

Appendix 1

The Metrics of RT-qPCR Assays

The metrics of RT-qPCR assays (Table 1) were determined from standard curves. The quantitative (real-time) reverse transcriptase polymerase chain polymerase reaction (RT-qPCR) assays were all performed in triplicate using a TaqMan™ ABI PRISM 7900 SDS (Applied Biosystems, California, USA) in 384-well plate format. Each assay well had a 10 µl reaction volume consisting of 5 µl 2X PCR master mix with uracil N-glycosylase (Universal PCR Mastermix, Applied Biosystems, California, USA), 0.1 µ each of 20 µM forward and reverse primers, 0.1 µl of 10 µM probe (ProbeLibrary, Roche Diagnostics, Lewes, UK) and 4.7 µl of sample cDNA (templates) or water (negative controls). Ten fold serial dilutions of the template were sued to generate the standard curves. R^2 values, the efficiency and dynamic range of each assay were determined from by the Sequence Detection System software (SDS 2.2.1, Applied Biosystems International, Warrington, UK) (Figure 1). Amplicon specificity was confirmed by electrophoresis of the product 1.5 % low melting temperature gel containing 0.5 µg/ml ethidium bromide, with 2 µl of 100 bp ladder (Gel Pilot 1 kb Ladder, Qiagen, Crawley, UK) for 1 hour at 120V, followed by visualisation and photography under ultraviolet light (Gel Doc 1000, Bio Rad Laboratories Limited, Hemel Hempstead, UK).

Figure 1

A standard curve generated by serial dilutions of a cDNA template using the probes and primers for COL31A.

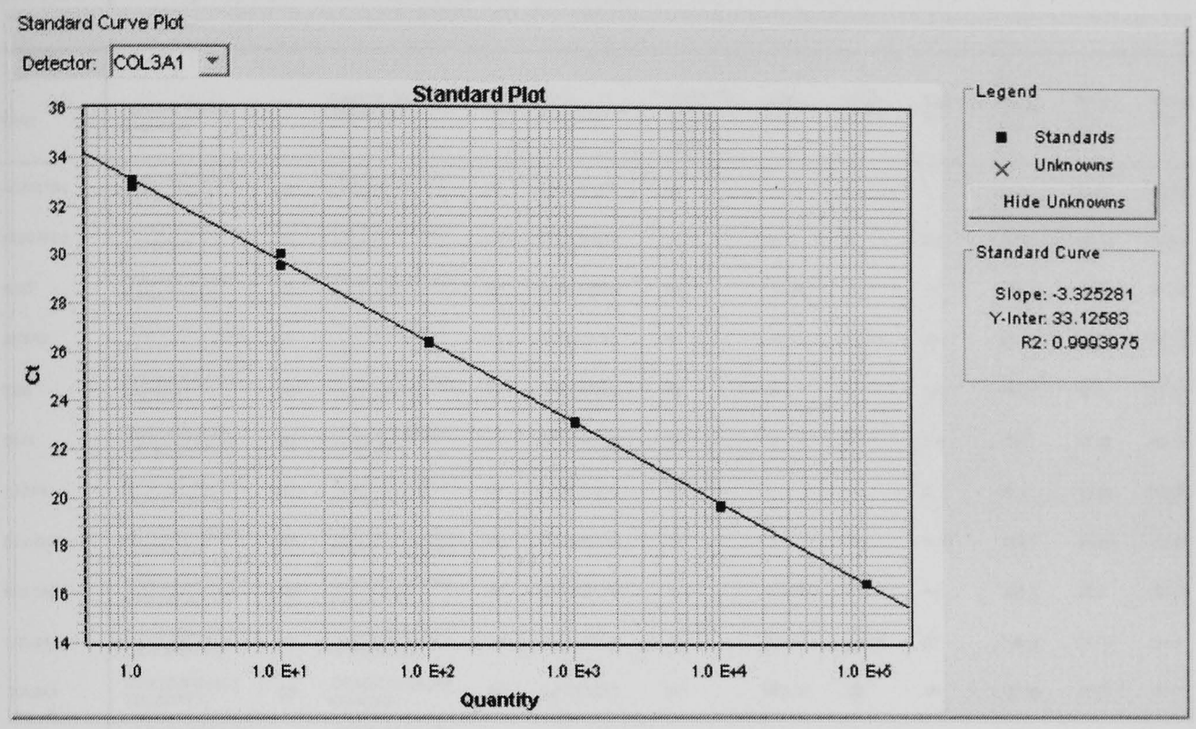


Table 1

The primer sequence, melting temperature (T_m), probe sequence, amplicon length, amplicon position within the gene (Exon; FP [Forward primer], RP [Reverse Primer] and Probe), dynamic range (upper and lower C_T values, of each quantitative reverse-transcriptase polymerase chain reaction assay.

Gene	Forward Primer Sequence	T_m	Reverse Primer Sequence	T_m	Probe Sequence	Amplicon Length (base pairs)	Exon: FP	Exon: Probe	Exon: RP	% Efficiency	Range low	Range high
ADAMTS4	GACCAGTGCAA ACTCACCTG	59	CAGGGAGTCCCA TCTACCAC	59	GGCCCTGG	83	7	7	8	87.4	30.52	34.19
ADAMTS5	TGGGTTCCCAAA TATGCAG	59	CTGTCCCATCCG TCACCT	59	CTGGGAGA	112	6	6	6-7	100.2	25.95	35.93
AGC	GGGACCTGTGT GAGATCGAC	60	GTAACAGTGGCC CTGGAACT	59	AGGAGCTG	68	12-13	13	13	101.5	18.54	34.74
ANXA2	AGAAAGTATGGC AAGTCCCTGT	59	CTTTCTGGTAGT CGCCCTTG	59	CATCCAGC	64	11	11-12	12	90.2	20.94	35.77
B2M	CCTTGCTCCTCA TCCTCCT	59	TGGGTGTCGTGA GTACACTTG	59	CAGCATCC	83	1	2	1-2	100.7	17.4	27.4
BGN	CAGAACAACGA CATCTCAGAGC	60	TCACCAGGACGA GAGCGTA	60	CTCCACCA	76	3	3	3-4	93.3	21.32	35.31
CASP8	GAGCTTCAGATA CCAGGCAGA	59	TGAAATCTGAAA AAGCATGACC	59	CTCTGCCT	63	1	1	2	87.1	23.58	31.05
C7orf28B	GCAGGAAGGGA TTCTCCAG	59	GGGTCCAGTAAG AAATCTTCCATAA	60	GCCAGGAA	75	9	10	9-10	108.7	30.49	37.13
COL1A2	CTATCAATGGTG GTACCCAGTTT	59	TGTTTTGAGAGG CATGGTTG	59	GCCTGCTG	111	43-44	44	44	100.7	23.2	33.05
COL2A1	CTGGTGAACCT GGACGAGAG	60	ACCACGATCACC CTTGACTC	59	CCTCCTGG	86	44	45-46	45	101.3	15.78	28.9
COL3A1	GGATGGTGGCT TCCAGTTT	59	CCAGCTGGACAT CGAGGA	60	GCTGCCTG	68	48-49	49	49	101.8	18.73	34.66
COL5A1	AACCTGTCGGAT GGCAAGT	60	CAGTCCAAGATC AAGGTGACAT	59	CAGCATCC	74	2-3	3	3	98.0	17.4	34.22
COL9A3	CGAGGTGCCTC AGGTGAC	60	ACCCAGCTCTCC TTTGTCC	59	GAGACCAG	105	12	13	14-15	101.4	26.33	32.91
COL10A1	ACCTGGACAACA GGGACCTA	59	CCCCTTTTCTCC TGGAAATC	59	AGCCCCAG	61	3	3	3	93.1	25.85	34.27
COX2	AAATTGCTGGCA GGGTTG	59	TCGAAGCTTTTG CTACTTGTTG	59	GGTGGCAG	69	8-9	9	9	103.3	28.91	35.32
CSPG2	TGGATGGTTTTA ATACGTTCAGG	60	GCCGTAGTCACA CGTCTCTG	59	CTGCCTTC	89	4	5	4	103.2	21.38	34.25
CTSB	CGGCCTTCACC GTGTACT	59	GTGACGTGCTGG TACACTCC	59	CTTCCTGC	61	8	8-9	8	106.9	20.37	35.83
CTSD	GGTCCACATGG AGCAGGT	59	TATGAGGGAGGT GCCTGTGT	60	TGGGCAGC	91	6-7	7	7	101.5	24.11	34.22
DCN	CGCTGTCAGTG CCATCTC	59	GGGGGAAGATCT TTTGGTACTT	59	TCCAGTGT	74	2	3	2	100	19.63	36.26
FN1	GACCAGAAGAG GCACAAGGT	59	GCTGGTTTAGGC CTTGGTC	59	GGGAGGAG	73	40	40	41	100.3	19.13	32.85
GAPDH	CTGGGGCTCAC TTGAAAGG	60	CAAACATGGGGG CATCAG	60	CTGCTCCT	72	3-4	4	4	101.9	22.23	36.99
GBP / PIAS1	GGAGACAATCA GCATTATAACAC CT	59	TGATCATCTGAC ACTGCTGCT	59	GGCTGCTG	65	12-13	13	13	99.2	28.66	38.72
Genomic DNA	AACCCTCAAAGA TGAGGTTTAGC	60	ACTCTGGGATCA CGCATGT	59	CTGCCTTC	60	NA	NA	NA	76.7	14.66	34.38
HIRP5	AATTCAGAACAT GCTGCAATTTTA	60	TGATTCATCATC CATAACCTGTTC	60	AGGTGGAG	70	6	6	6-7	93.8	21.66	32.04
TRAPPC2L	GATGATCCAGGT GTGCTGAG	59	CAATACGGTTAT GTCAACAGCACT	60	CTGGAGGA	74	5	5	5	94.5	24.45	37.95
IGF1	GGGGGTTCTAC TTCAACAAGC	59	TCATCCACGATG CCTGTCT	60	CTCCAGCA	76	2-3	3	3	100	20.22	33.51
IMP / ATIC	CGCTGCCTCTTT CAAACAT	59	TTTGGCCTCATC TTCACTGAG	60	CAGCAGGT	73	7	7-8	8	95.4	26.29	36.67
LUM	ACCTGGAAATTC TTTTAATGTATC ATC	59	CGGTATGTTTTT AAGCTTATTGTA GGA	60	TGCTGGAG	73	2	2	2	93.3	19.15	36.88
MAPK	TCTTCTTGGGAT AGCCAGTTTG	60	CCTCACCTCACA ACAAAACTGAT	60	GGTGGTGG	75	6	6	6	95.3	25.36	34.6
MMP13	CCGCGACCTTAT CTTCATCT	59	AACCTTCCAGAA TGTCATAACCA	59	AGAGGCAG	71	7	8	7-8	98.6	26.06	36.31

Gene	Forward Primer Sequence	T_m	Reverse Primer Sequence	T_m	Probe Sequence	Amplicon Length (base pairs)	Exon: FP	Exon: Probe	Exon: RP	% Efficiency	Range low	Range high
MMP2	ACCTGCAAGGC AGTGGTC	59	TCCAAATTTCAC GCTTTTCA	59	AGCTGGAG	89	11-12	12	12	95.1	15.34	32.24
MMP9	CACGCATGACAT CTTCCAGT	59	CGAGAATTCACA CGCCAGTA	59	CTTCTGCC	75	12	13	13	102.7	21.49	34.85
MRPS25	TGAAGGTCATGA CGGTGAAC	59	TGGATCTGAGGT ATGTTGAAAAAC	59	GCCAGGAA	88	1	1-2	2	85.6	20.4	40
NCK2	CAGACGCTCTAC CCGTTCA	59	GTCTCGCCCTTC TCGAAGTT	60	AGGAGGAG	62	2	2	2	93.3	30.45	33.94
NOS2A	GGCTCAAATCAC AACGGAAT	59	AGAGCTCGACCA GGAGAGTG	59	CCAGCCGC	66	17	18	17-18	106.8	26.43	36.14
NOS3	TCCTGTACCCTG CTTCATCA	59	TAAGATGCAAGG CAGACTGG	59	CTGCCACC	73	22	23	23	92.2	27.7	34.74
ORMDL2	ATGGACTACGG GCTCCAAT	59	CTGGCCAGGAG GTAGAGTACA	59	CTCCTCCC	77	2	2	3	106.2	23.28	34.45
PTDSS1	ACTCAGAATGCG ACGATGG	59	TCAGAACCTTTT GAACCTTTCG	60	CTGGTCTC	73	12	12	12-13	101.9	23.8	30.36
RPL13A	CTGCCCCACAA GACCAAG	60	GGGATCCCATCA AACACCT	59	CCAGGCTG	65	5	5	5	96.4	19.45	35.84
MRSPS7	AGTGCAGGGAG AAGAAGCAC	59	CAGCAGCTCGTG TGACAACT	60	GGATGCTG	75	5	5	5	101.7	22.39	36.02
SDHA	GGTGGCACTTCT ACGACACC	60	ATGTAGTGGATG GCGTCCTG	60	CTGGCTGG	89	4	4	4	102.5	21.56	34.62
TBP	TCCACAGCCTAT CCAGAACA	59	CTGCTGCTGTTG TCTCTGCT	59	CTGGAGGA	67	2	2	2	97.4	16.48	29.99
TIMP1	TGCATCCTGCTG TTGCTG	60	AACTTGGCCCTG ATGACG	59	CCCAGCAG	110	2	2-3	2	93.5	22.96	37.02
TIMP2	ATGGGCTGTGA GTGCAAGAT	60	CACTCATCCGGA GACGAGAT	60	CTGCCCCA	68	3-4	4	4	102.3	21.95	36.81
TIMP4	GCAGAGAGAAA GTCTGAATCATC A	60	GGCACTGTATAG CAGGTGGTAA	59	TGTGGCTG	72	4	5	4	95.1	17.89	31.63
TKT	CAACTTCTGTGG CTCCCACT	60	CCAGATCTTCCA GAGCCATC	59	TGGGGAAG	74	8	8-9	9	106.9	23.47	32.84
TNC	TGGATGGGACA GTCAAGGA	60	GCTCAGCTCTGC CAGGTTA	59	CCACCTCC	68	21-22	22	22	102.8	21.86	36.86
VIM	TACAGGAAGCT GCTGGAAGG	60	CCTCAGGTTCAG GGAAGAAA	59	GAGCAGGA	75	7	8	7-8	101.1	17.91	34.35
VZV	CAAAGCAGAACA TCGAGCAC	59	GGGTGTCACAG GGTGACTAAG	59	TCCTGCTG	67	NA	NA	NA	100.8	23.01	32.97

Appendix 2

A comparison of methods for identifying canine single nucleotide polymorphisms

ABSTRACT

Introduction

The identification of canine single nucleotide polymorphism (SNPs) has been facilitated by the publication of the canine genome and subsequent development of a canine SNP database. However the relative usefulness of different methods for identifying canine SNPs is undefined.

Materials and Methods

SNP frequencies from three publicly available sources were compared to determine the SNP frequencies in the different genomic regions of 20 genes. A comparison of *in silico* SNP identification was made with *in vitro* screening by evaluating the SNP frequencies in seven of the genes.

Results

The open-access SNP database provided nearly seven times the number of SNPs per gene compared to the canine genome, and the number of SNPs identified by *in vitro* evaluation was nearly double again. The majority of SNPs identified in the publicly available SNP database, and by *in vitro* screening were unique to their source. Nearly a third of SNPs identified by *in vitro* screening had been previously reported in a publicly available source.

Conclusions

Open-access sources of canine SNPs identify a large number of SNPs, but a considerable number of SNPs will be missed unless *in vitro* screening is performed. The true utility of different methods of SNP identification requires testing through genotyping.

INTRODUCTION

Single nucleotide polymorphisms (SNPs) are single nucleotide changes within the genome of in which the most common allele occurs with less than 99% frequency in the population at large (119). The functional significance of SNPs within and around genes are that in a coding region they may directly impact on the protein, in the intronic regions SNPs may alter splicing (120), and in the promoter region a SNP may influence gene expression (121). Additionally, SNPs with and without functional significance may show distinctive patterns of linkage disequilibrium which can be utilised in genetic linkage and direct association studies. The advantages of using SNPs, rather than other genetic polymorphisms such as variable nucleotide tandem repeats (VNTR's) or microsatellite markers for investigating disease are the multiple. Allelic discrimination of SNPs is straight forward, and multiple methods of genotyping SNP exist (122). SNPs are less mutable than other types of polymorphism (123), which should make them more reliable for assessing linkage disequilibrium, allelic associations and co-segregation phenomena as associations are unlikely to be confounded by mutation between generations (119).

Full release of the canine genome sequence in July 2004 (218) has transformed the field of canine genomic research. In addition to the draft sequence of the complete canine (Boxer) genome, and large numbers of SNPs were identified in the host genome and subsequently published. At present the *in vitro* identification of SNPs is both expensive and time-consuming, requiring the direct sequencing of DNA. The characterisation of the relationships between different SNPs and the canine phenotype are beginning to be investigated (398).

The aim of this pilot study was to evaluate the SNP frequencies in 20 genes, by comparing SNPs published in an on-line canine SNP database, the canine genome, and by active SNP mining through the evaluation of published mRNA sequences. The frequency of SNP identification, and their usefulness (as determined by their recognition in more than one breed) was further tested by comparing the SNP identifications in seven genes which had been screened *in vitro* for polymorphisms.

MATERIALS AND METHODS

Candidate gene selection

A selection of candidate genes for evaluation in association study with osteoarthritis was collected. The selection methods were to identify candidate gene selection were; reporting successful human candidate gene association studies, changed gene expression in diseased *in vivo* tissue, or changed gene expression in *in vitro* disease models.

Sequence information

The published canine genome sequence annotated for each gene and transcript were identified in the canine genome (277) and stored, including the 10 kb sequence up and downstream to the gene. The genomic locations of the promoter region (1000 bp upstream of the gene), each intron and exon boundary, and the 10 kb up- and down-stream of the gene of interest were recorded.

Public Available SNP identification

SNPs were identified from two publicly available SNP sources; the Ensembl canine genome (277) and the Broad Institute Canine SNP database (402). Further SNPs were identified by aligning all the published messenger RNA (mRNA) sequences in the National Centre for Biotechnology Information Nucleotide database (278) for each gene against the canine genome sequence (NCBI SNPs).

SNPs recorded in the canine genome were identified for each gene in the "transcript structure" report. The genomic location (chromosome and sequence number) was checked for each SNP by aligning the flanking sequence. SNPs recorded in the Broad database were identified by position, and again checked by aligning flanking sequence with the canine genome. Messenger RNA transcripts for each gene were identified using the following search strategy; "[gene of interest- both the abbreviation and full name separately]", "canis" or "dog", and "mRNA". Each sequence was recorded and the aligned using a web based sequence aligning tool (434). Discrepancies between each sequence and the canine genome sequence were recorded as SNPs, providing at least 95% homology was identified between the two sequences. Where poor homology was identified, the aberrant NCBI mRNA sequence was aligned with the canine genome gene sequence to try and locate any mis-annotated sequence.

The following information was recorded about each SNP;

- genomic position,
- base change,
- sequence 20bp up- and downstream of each SNP,

- The number of breed(s) of dog in which SNPs were identified in.

Alignment of the genomic positions allowed identification of SNPs common to more than one database. The SNP detection rates (SNP per 10000 base pairs [bp]) were recorded for each gene by dividing the number of SNPs detected by the length of sequence analysed for; the entire sequence length evaluated the promoter region, the intronic region, and exonic regions. The proportion of SNPs common to more than one database and the proportion of SNPs common to more than one breed were also recorded.

In vitro SNP Identification

SNP identification had been previously performed by separate investigators (Andrea Short, University of Manchester, and Annette Barnes, University of Liverpool) on seven of the genes selected prior to this study (401). Briefly, 93 DNA samples from eleven breeds of dog (Golden Retriever [n = 8], Rottweiler [n = 8], Shih Tzu [n = 8], Labrador [n = 16], Yorkshire Terrier [n = 8], American Cocker Spaniel [n = 7], Cavalier King Charles Spaniel [n = 2], Doberman [n = 9], West Highland White Terrier [n = 9], German Shepherd dog [n = 9] and Beagle [n = 9]) had targeted areas of each gene of interest amplified by PCR. The PCR product was assessed for the presence of a polymorphic product using denaturing high performance liquid chromatography (dHPLC) (400), and amplicons with melt curve analyses suggestive of a SNP were sequenced. SNPs identified were annotated a genomic position by alignment of the sequence with the canine genome.

The breed and adjacent sequence were recorded, and the data aligned with that of the other databases in an Excel spreadsheet (Microsoft, UK). The SNP detection rates were recorded as previously described.

RESULTS

Candidate gene selection

The candidate genes selected were; Aggrecan (*AGR*) (129), Bone morphogenic protein 2 (130), cartilage intermediate layer protein (*CILP*) (130), type II collagen alpha I chain (45) (*COL2A1*), Type IX collagen alpha 3 chain (*COL9A3*) (142), cyclooxygenase 2 (*COX2*) (130), decorin (*DCN*) (251), oestrogen receptor alpha (*ERα*) (130), insulin like growth factor 1 (*IGF1*), interleukin 1 alpha (*IL1α*) (159), Interleukin 1 beta (*IL1β*) (162), interleukin 1 receptor antagonist (*IL1ra*) (159), interleukin 4 (*IL4*) (164), matrix metalloproteinase-2 (*MMP2*) (248), -3 (*MMP3*) (248), -9 (*MMP9*) (255), osteoprotegrin (OPG) (130), tissue growth factor beta (*TGFβ*) (169), tissue necrosis factor alpha (*TNFα*) (133) and vitamin D receptor (*VDR*) (435). The candidate genes previously evaluated by *in vitro* analysis were *IL1α, IL1β, IL4, MMP9, TGFβ, TNF* and *VDR*.

Public Available SNP identification

The results of the SNP identification rates are presented in Figures 1 and 2. When all 20 candidate genes were evaluated, the mean SNP detection rate for the SNP database was 9.7 SNPs per 10 kb (range 2.9 - 19.6), and for the canine genome was 1.3 SNPs per 10 kb (range 0.0 – 5.1). The mean percentage of SNPs confirmed as being present in more than one breed for the SNP database was 14% (range 0 – 55%) and for the

canine genome was 14% (range 0 - 50%). The mean percentage of SNPs unique for the SNP database were 94% (range 55% – 100%) and for the canine genome was 45% (range 0% – 100%). The mean SNP detection rates in the promoter region for the SNP database was 10.0 SNPs per 10 kb (range 0.0 – 50.0), and for the canine genome was 0.5 SNPs per 10 kb (range 0.0 – 10.0). The mean number of SNPs in intronic sequences for the SNP database was 9.4 SNPs per 10 kb (range 2.9 - 30.4) and for the canine genome was 1.2 SNPs per 10 kb (range 0.0 - 29.7). The mean number of SNPs in exonic sequences for the SNP database was 3.9 SNPs per 10 kb (range 0.0 - 250.4), for the canine genome was 0.7 SNPs per 10 kb (range 0.0 - 14.2), and for mRNA sequences was 32.2 (range 0.0 - 250.3).

When the 7 candidate genes were evaluated, the mean SNP detection rates for the SNP database were 9.4 SNPs per 10 kb (range 2.9 – 15.2), for the canine genome were 2.1 SNPs per 10 kb (range 0.0 – 5.1), and for *in vitro* identification was 15.4 (range 2.3 – 32.1). The mean percentage of SNPs confirmed in more than one breed for the SNP database was 18.0% (range 0 – 55%), for the canine genome was 14.3% (range 0 – 44.4%) and for *in vitro* identification was 100%. The mean percentage of SNPs unique to a database for the SNP database was 84.9% (range 55 – 100%), for the canine genome was 28.3% (range 0.0 – 50%) and for *in vitro* identification was 72.5% (range 30 – 100). The mean SNP detection rates for the promoter region of the SNP database was 7.1 SNPs per 10 kb (range 0.0 – 30.0), for the canine genome was 0.0 SNPs per 10 kb and for *in vitro* identification was 15.7 (range 0 – 30). The mean number of SNPs in intronic sequences for the SNP database was 9.4 SNPs per 10 kb (range 2.9 - 30.4), for the canine genome was 5.2 SNPs per 10 kb (range 0.0 - 29.7) and for *in vitro* identification was 19.7 (range 2.5 – 59.5). The mean number of SNPs

in exonic sequences for the SNP database was 3.4 SNPs per 10 kb (range 0.0 – 14.2), for the canine genome was 2.4 SNPs per 10 kb (range 0.0 - 14.2), and for in vitro detection was 11.9 (range 0.0 – 42.7).

DISCUSSION

The canine genome sequencing project identified nearly 770,000 SNPs in the genome of a single Boxer dog. The observed mutation rate in this individual was approximately 1 SNP per 1700 bases (218). Comparison with a second canine genome sequence from a single Standard Poodle dog (436) increased the observed mutation rate to 1 SNP per 900 bases. Subsequent comparison with partial genome sequences of a number of other breeds revealed that 1 SNP per 900 bases is the approximate mutation rate when comparing an area of the Boxer genome to that of another breed (218). Thus it is unsurprising that the mutation rate recorded from the SNP database was similar to this (approximately 1SNP per 1000 bases), as the database is hosted by the institution responsible for sequencing the canine genome and reporting on the SNPs identified in project (and SNPs subsequently identified in other breeds of dog).

The SNP database provided nearly seven times the number of SNPs per gene compared to the canine genome sequence in the candidate genes we evaluated. The reasons for this were not entirely clear, as the reported mutation rate for the canine (Boxer) genome suggests that the difference should have been only two fold (218). Furthermore a number of the SNPs recorded in the canine genome are simply differences in sequence between the two canine genome sequencing projects, and thus

may not represent functional SNPs in individual breed populations, although evidence reported from the canine genome sequencing project suggested that they may well be informative, as on average 73% of the distinct SNP are informative (polymorphic) within a given breed (218).

The *in vitro* SNP detection method identified nearly double the number of SNPs across the seven genes of interest than the SNP database, which reflects both the number of individuals screened and the diversity of breeds evaluated. Thus *in vitro* investigation identified a number of SNPs which could not be defined electronically. Furthermore all these SNPs were present in multiple breeds, in contrast to the *in silico* SNPs, reflecting that these SNPs have been tested in a large number of individuals unlike the *in silico* SNPs. *In vitro* detection of SNPs identified similar mutation rates across all regions of the in the seven candidate genes assessed unlike the *in silico* SNP detection methods which identified lower detection rates in exonic regions. A lower mutation rate in exonic regions when compared to other genomic regions would be expected, but the *in vitro* SNP detection focussed primarily on exonic regions which may have artificially increased the SNP rates identified in these areas.

The largest numbers of SNPs were identified by the evaluation of mRNA sequences from the public databases. Nearly 10x the number of exonic SNPs were identified when compared to the canine SNP database. One would anticipate that most of these sequences would have been generated from a single individual, and there was an average of 2.5 sequences reported per gene. Thus the mutation detection rate was considerably higher than would be expected which suggests that proportion of the SNPs identified by this method may have been sequencing error. Genotype analysis

of each of these mutations would have determined the true utility of this method of mutation detection in dogs. Absolute verification of human expressed sequence tag (EST) polymorphisms in coding regions can be performed using a combination of Bayesian inference to weigh evidence for true polymorphism versus sequencing error, misalignment or ambiguity, miss-clustering or chimeric EST sequences. This type of statistical analysis was not evaluated in this study, but when applied to select human SNPs predicted by EST analysis assessment of the resulting SNP profiles indicated that over 70% of the filtered SNPs are valid (437), suggesting the EST / mRNA databases can be a valuable source of SNPs.

The majority of SNPs identified in the publicly available SNP database, and by *in vitro* screening were unique to that source, which is surprising given that the majority of SNPs available from the canine genome should have been available in the canine SNP database. The reason for the discrepancy was not determined, but clearly both databases need to be searched to maximise the chances of identifying all freely available SNPs. Indeed, when SNPs identified by *in vitro* determination were also considered then the majority of SNPs present in the canine genome could be identified elsewhere. This study is purely descriptive and makes no attempt to identify the most useful SNPs for canine studies. Clearly, breed (438) and geographical differences may have significant effects of the usefulness of different SNPs which remain unresolved until all SNPs detected have been genotyped across a population of interest.

Nearly a third of SNPs identified by *in vitro* screening had been previously identified in a publicly available source, however there was no clear way of identifying those

SNPs from the general SNPs listed without recourse to pilot genotyping studies. Evaluation of the human public domain SNP databases identifies similar variations in *in-silico* SNP utility. For example 80% of the candidate SNPs in databases such as "The SNP Consortium" (439) are polymorphic and roughly 50% of these SNPs to be common (minor allele frequency of ≥20%) (440). Conversely, only 15% of candidate SNPs in the dbSNP (278) are thought to have been proven to be polymorphic in any population (440). Our results suggest that the same is true of electronic canine SNP sources, with the SNP database being a more useful source than the canine genome.

CONCLUSIONS

Although publicly available SNP databases provide a ready source of freely available SNPs, a large proportion of these were found to be breed specific in nature. At present, open-access canine SNP sources are not sufficiently detailed to preclude the use of *in vitro* SNP screening for the investigation of genetic traits suspected of having a similar aetiology in more than one breed. Messenger RNA sequences could provide a large potential source of exonic SNPs, but their utility has yet to be determined in the dog.

Figure 1

A comparison of SNP numbers in 20 candidate genes identified from an in-silico canine SNP database (in silico database), the canine genome sequence (in-silico canine genome) and in-silico mRNA files (in-silico mRNA). The SNP metrics evaluated were; the percentage of SNPs present in more than one breed (% SNPs >1 breed), the percentage of SNPs which were unique to the database (%unique), the number of SNPs per 10kb of sequence evaluated (SNP/10kb), the number of SNPs per 10kb intronic region (Intronic SNP/10kb) and the number of SNPs per 10kb exonic region (Exonic SNP/10kb. The mean value and standard deviations are presented

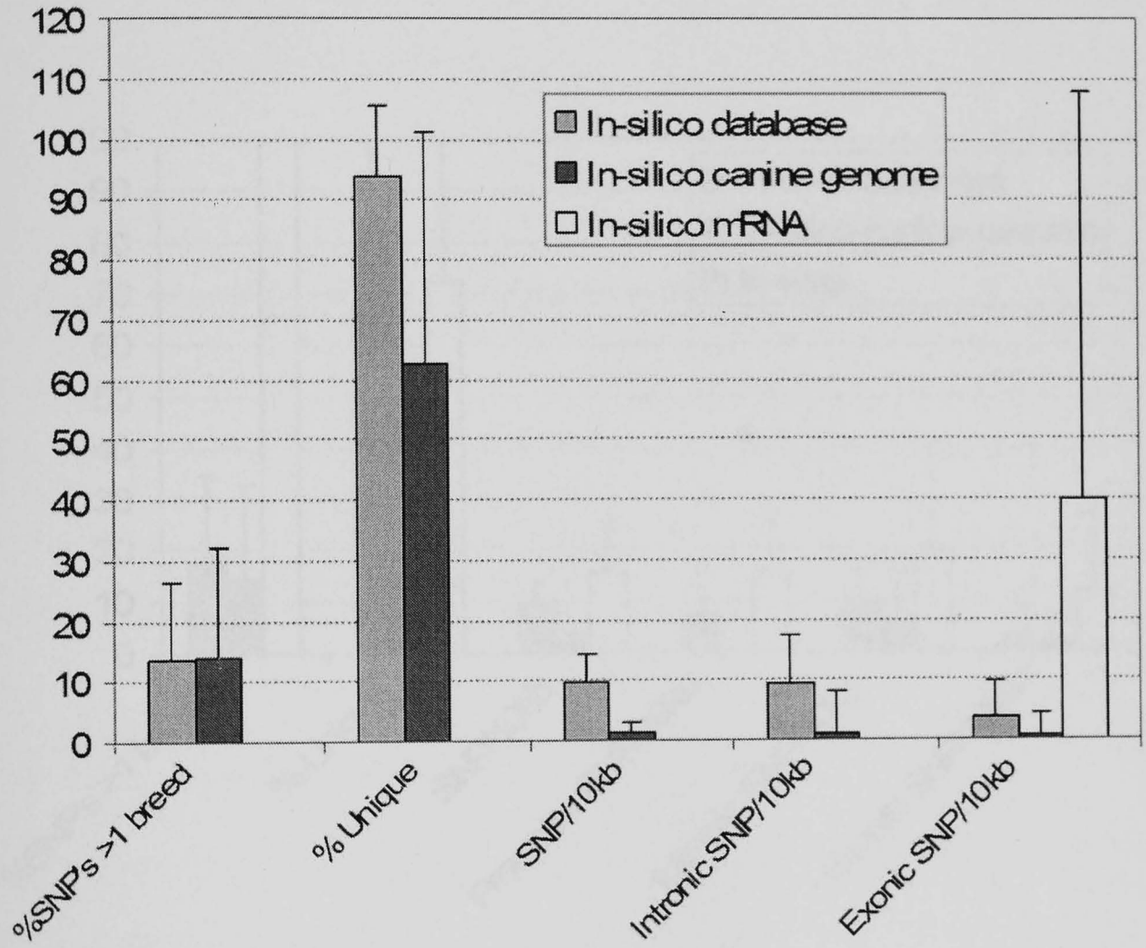
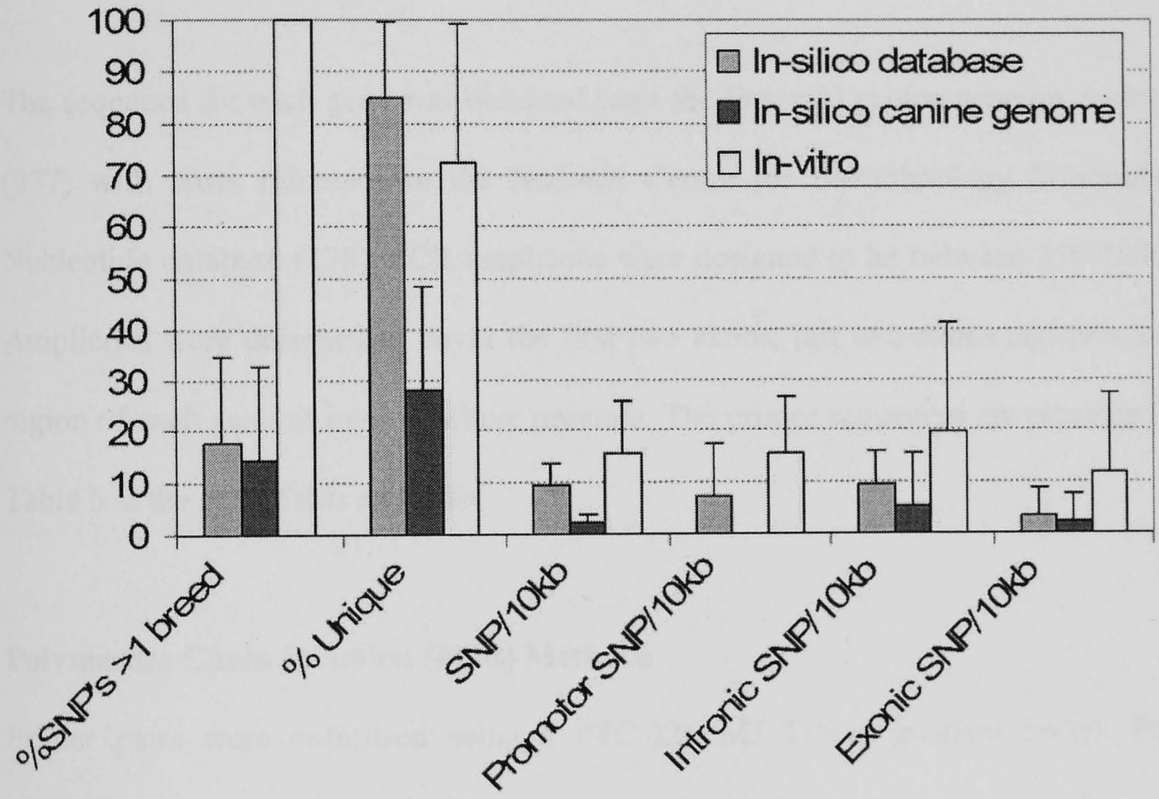


Figure 2

A comparison of SNP numbers in 7 candidate genes identified from an in-silico canine SNP database (in-silico database), the canine genome sequence (in-silico canine genome) and in vitro evaluation (in vitro). The SNP metric evaluated were; the percentage of SNPs present in more than one breed (% SNPs >1 breed), the percentage of SNPs which were unique to the database (%unique), the number of SNPs per 10kb of sequence evaluated (SNP/10kb), the number of SNPs per 10kb promoter region, (Promotor SNP/10kb), the number of SNPs per 10kb intronic region (Intronic SNP/10kb) and the number of SNPs per 10kb exonic region (Exonic SNP/10kb). The mean value and standard deviations are presented



Appendix 3

Supplementary Data for Chapter 8

Single Nucleotide Polymorphism Detection and Genotyping

Primer design for SNP identification

The sequence for each gene was obtained from the Ensembl canine genome database (277) with cross reference to the National Centre for Biotechnology Information Nucleotide database (278). PCR amplicons were designed to be between 350-700bp. Amplicons were designed to cover the first two exons, last two exons and promoter region of each gene of interest where possible. The primer sequences are presented in Table 5 at the end of this appendix.

Polymerase Chain Reaction (PCR) Methods

Primer pairs were optimised using a PTC-225 MJ Tetrad gradient cycler. PCR reactions were performed in 25 µl volumes containing 25 pmol forward and reverse primer pairs, 1x Qiagen buffer (10x stock containing 15 mM $MgCl_2$), 0.5 units Qiagen hot start Taq, 0.8 mM dNTPs (ABgene), 25 ng DNA and deionised water. A negative control reaction was included for each primer pair. Initially, optimisation was performed using a touchdown PCR. The cycling conditions were denaturation at 95°C for 10 minutes, 30 seconds at 95°C, 1 minute at 64°C, decreasing by 1°C per cycle for 10 cycles, and extension of 1 minute at 72°C, followed by 30 cycles of 30 cycles of 30 seconds at 95°C, 1 minute at 55°C and 1 minute at 72°C. A final extension of 72°C for

10 minutes completed the reaction. Amplicons which did not amplify using the standard PCR protocol were optimised using temperature $MgCl_2$ gradients. The $MgCl_2$ gradients varied from 1.5 mM to 4.5 mM in 1 mM increments, by adding $MgCl_2$ 50mM (Qiagen) at the appropriate volume to each reaction.

Agarose Gel Electrophoresis Methods

A 2% agarose gel was used for electrophoresis of DNA fragments. Agarose gels were prepared with 60 ml of 0.5% TBE (1 in 20 dilution of 10xTBE Buffer [135 mM Tris-HCl, 45 mM boric acid, 2.5 mM EDTA pH 8.0] with deionised water) and adding 2g agarose (Sigma), allowing the mixture to stand for 3 minutes, before melting in a microwave using the standard power level. When fully melted, the remaining 40 mls of 0.5x TBE were added and mixed. Ethidium bromide was added to a final concentration 0.5 µg/ml, and the solution was mixed and poured into a gel tray containing and appropriate comb. The gel was left to set for 15 minutes, before adding 5 µl of each sample to 1 µl of 6x Gel loading buffer (ABgene), and 2 µl of 100 bp DNA ladder (Superladder-low, ABgene) the immersed in a gel tank containing enough 0.5x TBE to cover the gel, and run at 120V for 60 minutes.

Denaturing High-Performance Liquid Chromatography (dHPLC) Mutation Detection Methods

For dHPLC WAVE screening, PCR reactions were performed in 25 µl volumes containing 25 pmol of forward and reverse primer, 1x Qiagen buffer (10x stock containing 15 mM $MgCl_2$), 0.5 units Qiagen Hot Start Taq, 0.8 mM dNTPs (ABgene),

25 ng DNA and deionised water. Negative (no template) reactions were included for each amplicons. dHPLC was carried out on WAVE DNA Analysis Equipment (3500HT, Transgenomic Inc., Elancourt, France). PCR products were screened for heteroduplexes by subjecting the 25 µl PCR reaction to a denaturing step (95°, for 5 minutes) and a gradually annealing gradient of 1°C / 90 seconds down to a final temperature of 4°C. Routinely 5 µl of the PCR product were separated though a 2% linear acetonitrile gradient at the optimal temperature. The standard WAVE buffers were Buffer A, Buffer B, Buffer C, Buffer D and syringe wash solution (Transgenomic Inc.).

The DNA sequence of each amplicon was entered into WAVEMAKER software (Transgenomic Inc.), and the temperatures for optimal heteroduplex separation determined. Each PCR product was analysed at least two different temperatures to allow detection of polymorphism along the entire length of the amplicon. Temperatures at which the helix fraction of the DNA sequence was between 50-95% were used to screen samples. A low range mutation standard was included in each run to verify the column resolution.

Sample patterns were analysed and numbered using the WAVEMAKER software (Transgenomic Inc.). A total of 23 samples (12 Labrador Retrievers, and 11 Golden Retrievers) were analysed for each amplicon. Every sample which demonstrated a different amplicon elution pattern (usually 2 or 3 samples per amplicon) was sequenced to identify polymorphisms.

DNA Sequencing

Five microlitre of PCR product were treated with 2U shrimp alkaline phosphatase (SAP, ABgene) and 10U *ExoI* (Sigma). The plate was placed on a PTC-225 MJ PCR machine, heated to 37°C for 30 minutes followed by 15 minutes at 85°C. The products and primers were sent to Lark Technologies Inc. (Saffron Walden, UK) for DNA sequencing in forward and reverse orientation. The sequence traces were analysed using Trev (version 1.9, MRC) and aligned with a web-based multiple alignment tool (434). The position of all SNPs were checked with reference to the canine genome and annotated appropriately (277); where SNPs had been previously reported, the RefSeq (rs) number was used. Approximately 50% of all detected mutation had been previously reported.

Further SNP Identification

Further single nucleotide polymorphisms were identified in candidate genes by evaluating the canine genome (277) and an open access canine SNP database (402). Where possible candidate SNPs where chosen which had been sequenced in more than one breed. A full list of all the SNPs genotyped are presented in Table 6.

Sequenom Genotyping

Sequenom matrix-assisted laser desorption / ionisation time of flight mass spectrometry (MALDI-TOF MS) to genotype genomic DNA. The principle of the assay is that a short sequence containing the SNP location is amplified by PCR, and the reaction cleaned with SAP (as for DNA sequencing). Oligonucleotide probes are designed to anneal adjacent to the SNP and added to the PCR reaction. DNA polymerase and terminator nucleotides extend the primer through the polymorphic

site generating allele-specific extension products (dependent on which SNPs are present), each with a unique molecular mass. These masses are analysed by MALDI-TOF MS and the genotypes assigned on the basis of the detected mass. The resolution of the instrument allows up to 35 different SNPs to be detected in a single reaction.

Primers and probes for each SNP were designed using the Assay Design Software (Sequenom), and are presented in Table 7. Primers were diluted to 100 μM and plexes pooled to contain 500 nm of each forward and reverse primer (appendix 3). Probes were diluted to 400 μM and the probe pools split into 50% high and 50% low mass probes (on the basis of the probe molecular weight), and pooled to contain 7 μM (low mass) or 14 μM (high mass) probes.

For each PCR reaction 20 ng DNA was plated into a 384 well plate and dried down by heating to 80°C for 15 minutes on a PCR machine with no lid. The PCR was performed in 5 μl volumes with each well containing 1.25x Qiagen buffer (10x stock containing 15 mM $MgCl_2$), 0.5 units Qiagen Hot Start Taq, 0.5 mM dNTPs (ABgene), 1.625 mM $MgCl_2$, 100 nM primer mix. The reaction conditions were; 95°C for 15 minutes, 35 cycles or 95°C for 20 seconds, 56°C for 30 seconds, 72°C for 1 minute, followed by and 72°C for 3 minutes on a Dyad PCR machine. The reaction was then maintained at 4°C. The reactions were then treated with 0.3U shrimp alkaline phosphatase (SAP) to inactivate any dNTPs left over from the reaction. Reactions were incubated at 37°C for 40 minutes, and denatured at 85°C for 5 minutes.

iPLEX primer extension was carried with 0.22x iPLEX buffer, 1x iPLEX termination mix, 0.625 μM low mass probe, 1.25 μM high mass probe, 1x iPLEX enzyme assed

to each well. The reaction were amplified by cycling at 94°C for 30 seconds, 40 cycles of 94°C for 5 seconds, 5 cycles of 52°C for 5 seconds and 80°C for 5 seconds, and a final extension of 72°C for 3 minutes. Samples were diluted with 25 µl water and desalted with 6 mg Resin before being centrifuged at 4000 rpm in a Jouan CR4 centrifuge, and spotted onto a spectroCHIP using a Sequenom MASSarray nanodispenser (Samsung). The spectoCHIP was run in the MASSarray system and genotype data exported from the instrument in an Excel spreadsheet.

Data Analysis

Genotype and phenotype data were imported into BCgene software (405), which was used to calculate genotyping rates, minor allele frequencies (MAF), Hardy-Weinberg equilibrium (HWE) for each control population (Table 7). SNPs were not analysed further if the call rates were below 80% or if the control population was not in HWE. Case-control comparison of MAFs was performed by Chi^2 (χ^2) comparison and odds ratio (OR) calculation using the BC gene software (Tables 8, 9, 10 and 11). Data analysis was repeated after stratification of each population on the basis of previously reported diseases risk factors (neuter status for cruciate disease [evaluation of neutered animals only], and sex for ED and HD [evaluation of male animals only]) (Tables 12). Significant differences ($P < 0.05$) were checked for multiple permutations using Monte Carlo simulation (406), using a freely available software program (T1 statistic, CLUMP) (407). Any SNPs with MAF demonstrating sex or neuter associations within the control population were removed from further analysis (3 SNPs eliminated because they demonstrated a sex bias).

Haplotype frequencies were estimated for each cohort (control populations, disease population and control population stratified on sex or neuter status). SNPs were

considered for haplotype analysis if the minor allele frequencies were greater than 5% and the SNPs were in HWE the population analysed. Thus haplotypes were calculated for 13 genes in GRs (*ANKRD10, ATP11B, IL1α, IL4, IL6, IL10, IL12, LEPR, SPARC, TIMP3, TIMP4, TNC, TNFα* and *ZSWIM2*) and for 10 genes in LRs (*ANKRD10, MMP9, IL4, IL6, IL10, IL12, SPARC, TIMP3, TNC* and *TNFα*). Maximum likelihood haplotype frequencies were computed using an expectation-maximisation algorithm, using HelixTree version 4.1.0 software (GoldenHelix, Inc., Bozeman, USA). Haplotype frequency estimates were multiplied by the number of chromosomes in diseases and control groups to generate contingency tables. Frequency estimates were compared between controls and cases by χ^2 analysis (Table 13, 14 and 15) checked for multiple permutations using CLUMP (T1 statistic) and ORs and CIs calculated for the haplotypes of each gene using a web-based statistical calculator (408). Contingency tables containing values less than 5 were analysed using web-based Fishers exact test calculator (409). The significant associations are presented in Chapter 8, Tables 3 and 4.

SUPPLEMENTARY RESULTS

Table 5

The metrics of the primer pairs used in WAVE screening. Gene location, primer sequences, melting temperatures (Tm), amplicons lengths, $MgCl_2$ concentrations, and the WAVE melting temperatures used are listed..

Gene	Gene Location	Forward Primer	Tm	Reverse Primer	Tm	Amplicon Length	[$MgCl_2$] (mM)	Melt Temp 1	Melt Temp 2
TIMP2	Promotor 1	GTCTCCTTCATCCTTGTGTC	58.6	TCTATCCTACTACTCCTAATCTGTCC	58.8	422	3.5	62.1	
TIMP2	Exon 1	GCACTTAGGCGTGACATC	59.3	GTTACTCAAGAAGGCAGGAC	57.9	698	4.5	57.9	62.1
TIMP2	Exon 2	GTAGAGGAGCAGTTGAGAAGTC	58.7	GTGAGCATTTCCAGCATC	59.3	659	4.5	61.9	59.9
TIMP2	Exon 3	CACAGACCAGGCATTGAC	60.3	GAAGGAGGTGATGAGGTAAAG	59.2	587	4.5	61.2	62.4
TIMP2	Exon 4	CAAACTGCCCTTGAACATC	60.7	GAAGCCTGGGACACACAC	61.9	519	4.5	61.0	62.9
TIMP4	Promotor	AGACTTGGAGCCCTATGC	59.3	CTATTTCCCTGCCTGAGC	59.9	714	4.5	58.5	61.1
TIMP4	Exon 1	TTGTTTCACTTTCTTTCTGC	57.1	TCCCTTACACTTTCATTTCC	58.0	687	4.5	61.2	63.8
TIMP4	Exon 2	ATGGCGAGAGATAAACTACC	57.3	CCACTTCAGCATCAATACC	58.0	692	4.5	57.9	
TIMP4	Exon 3	CATCTGTGTGTGCCAAATCC	64.0	TCATAGGCAGAGGGCAGAAC	64.0	616	4.5	54.9	58.0
TIMP4	Exon 4	GACGACAACTTACTACTTTCATC	56.1	CCCAACTTCATACAGAGC	55.4	625	4.5	56.3	59.2
TIMP4	Exon 5	TGCTCAAGGTCAGTGGTAG	58.9	TGAAGGGATGTGATGGTC	59.3	471	4.5	58.3	61.9
SPARC	Promotor	CACAAACAGTCATACATTCATC	56.9	TAATCTTCACCAGCGAGTC	57.9	428	4.5	57.4	61.5
SPARC	Promotor 2	GGGCACAATAACAATACC	55.1	GCTGCTACCTACACAGACTC	56.7	518	4.5	58.2	61.2
SPARC	Exon 1	AGATTGTTCCAGATGATTCC	57.8	CAACCAGAGAAGGGAGAG	57.2	389	4.5	62.4	63.5
SPARC	Exon 2	TCCAGAGTTCCAATGAGC	58.6	TACAGACCCACCGAGATG	59.6	579	4.5	60.8	
SPARC	Exon 9	AAATCAAGAGTCGGAGAGC	58.5	TAACGAGGCACAGAGAGG	59.0	698	4.5	57.3	63.2
SPARC	Exon 10	GTTTTCTCATCTTGCTTTG	54.9	GTTTTCCTAACACAGACTCAG	55.4	642	4.5	57.8	61.6
TNC	Promotor 1	TGCCCTCTTACTGTGTCC	58.7	ATAGATGTTCAGGTGGTTCC	57.8	604	4.5	57.7	58.9
TNC	Promotor 2	AGAAATAAAGCAGGGAGAAG	56.7	TGGAAGCACAGAGTAGAATC	56.8	682	4.5	57.0	58.9
TNC	Exon 1	TTTCCCTCACCTTCTTGTAG	58.3	TCATCACTTCTCCCATCTG	58.7	653	4.5	62.2	
TNC	Exon 3	ACAAGTGAAGGGAAACAAAC	58.8	GCAAAGGTGAAGCAGTAAG	57.6	452	4.5	63.1	
TNC	Exon 27	TGCTTTCTCCACCTTATTTC	58.3	CTGTTTAGTTGCGGTTCTG	58.4	592	4.5	59.0	62.1
TNC	Exon 28	CAACAACTAAGCGGCAAC	59.4	CCAAGAGAAGGACAACCAC	59.7	471	4.5	60.2	
ANKRD10	Promotor	CCCAGACACCTGCTTACC	60.7	TGGCACAGACACACATAGAG	59.9	522	4.5	57.3	59.2
ANKRD10	Exon 2	ACGCAGTTTTGAGAGAAGTC	58.6	AGCCGAAAGGTGAAAGTATC	59.7	596	4.5	55.3	60.7
ANKRD10	Exon 3	GAAGTAGGCTGTGTGTGGTC	59.7	GGTGGAGGAATAAGTGAAGG	59.9	696	4.5	58.0	
ANKRD10	Exon 4	TAATGAGATGTGCGTAAGAGC	59.4	CCAGCCGTAAAAGAGAGC	60.1	494	4.5	59.6	
ANKRD10	Exon 5	AAGGAACAACCATAAAGGAG	57.1	TAGAATACACCCAAAACACG	57.2	641	4.5	55.7	61.4
ANKRD10	Exon 6	AATGCTGCTGTTGGGTTC	61.3	AGGAGGGAGGAGTGTGTCC	62.3	721	4.5	56.7	61.2
ATP11A	Promotor 1	CCACATACCCTTTCACATTC	59.2	GTTGCTTACGAGTTCAGTTTC	58.2	618	4.5	54.3	55.5
ATP11A	Promotor 2	ACAACTTATCCCATCTTACAACC	59.7	ACTGCCAAGGACCATCAC	61.2	746	4.5	57.8	59.8
ATP11A	Exon 6	AAATAGGAGTTTAGGAAGACC	54.7	AATGAATGAATGTGACTGTG	55.7	719	4.5	52.4	55.3
ATP11A	Exon 7	ATAAAGCCTCGGTAAATGG	58.1	GTCTCAGCAAAATCACAAAC	57.3	619	4.5	51.9	58.1
ATP11A	Exon 12	TGTTCTACCTTCTTTCTTTGC	57.7	GCCTGTATTCTCACACTCTATG	57.9	699	4.5	53.2	56.5
ATP11A	Exon 13	CCTGCCTGGTTATTATTTCC	59.8	CACACCTCCTGCTTTAGTTG	59.9	713	4.5	54.2	58.5
ZSWIM2	Promotor	GGTTTTTCTTTCTCTTTAGTTAC	54	TGACCTACACTCACTTTTTG	55.2	519	4.5	54.4	
ZSWIM2	Exon 1	TATGAAGAGGCTGGTGTC	55.9	TCAAGATGAAATAAAATACTGC	55.2	422	4.5	53.2	55.3
ZSWIM2	Exon 2	GACTGACTGGCGTATCTC	55.4	ATTACATTTCTTATTTTTCCTG	53.4	687	3.5	53.2	
ZSWIM2	Exon 7	CTAAACACAAGGCGGAAG	58.3	CCCTACAAGGAAAGTCAAAC	57.6	561	4.5	52.8	56.9
ZSWIM2	Exon 8/1	AATAATGAAACTGGGAGAAAC	56.2	TTGTGAGCAATAGAAGGAAG	57.1	554	4.5	53.8	55.2
ZSWIM2	Exon 8/2	CTCAATGTAATTCTGGAAAACTG	59.2	CAAGTAGGTAAAATGATGACAAATAAG	59.8	704	4.5	53.3	58.5

Table 6

A list of the SNPs genotyped. The gene, SNP identity (SNP_ID), canine chromosome number (CFA), gene position (Exonic SNPs are also denoted as coding for a synonymous (Synon) or non-synonymous (Non-synon), and the amino acid identity and number are stated), base pair number, source (AS = (401), AB = Annette Barnes (personal communication), DC = Identified within this study as described above, rs = RefSeq (278)), minor alleles (MA LR = minor allele in Labrador Retrievers, MA GR = minor allele Golden Retrievers), major allele (MJR LR = major allele in Labrador Retrievers, MA GR = major allele Golden Retrievers), and the 20 base pair flanking sequence either size of the SNP (Left above, Right below). ***Please note that all SNPs designated AS should be referred to the original source (401), and the associated restrictions on their use.***

Gene	*SNP_ID*	*CFA*	*Gene Position*	*Base Number*	*Source*	*MA LR*	*MJA LR*	*MA GR*	*MJA LR*	*Left and Right flanking sequences*
ANKRD10	ANK_E5	22	EXON 5 (Synon: Val - 254)	62018474	rs23105168	G	C	C	G	ACAGCAGCAAACTCTCTGTC ACCAGCATTTTGTCAGCTTC
ANKRD10	ANK_I4	22	INTRON 4-5	62021737	DC	G	A	G	A	AAGTTGTCTTGATGGTTCAG AGCATTAGTTGTTCCTGAGG
ANKRD10	ANK_E4B	22	EXON 4 (Non synon: Asp - 163 [A] to Tyr [C])	62027111	rs23065428	A	C	A	C	ACCCTGGGTTTGAGCAATGT TGCTGCCGTCAGGCCACTGG
ANKRD10	ANK_E4	22	EXON 4 (Synon: Thr -160)	62027118	rs23065425	A	C	A	C	GTTTGAGCAATGTCTGCTGC GTCAGGCCACTGGCATTTCT
ANKRD10	ANK_I3	22	INTRON 3-4	62030377	rs23059739	C	T	C	T	GCAGAAGATCTCTTAATACC GATTCTGGCACTCACCAACA
ANKRD10	ANK_I2	22	INTRON 2-3	62038934	DC	A	G	G	A	TTGGGTTTTAAGTACAACAC TTTGTGCTGAGATCAGCATA
ANKRD10	ANK_I1	22	INTRON 1-2	62043702	rs23039786	A	G	A	G	AAAAAGTCTGTGTAGCTATA TACGCTCTTCAGAGACTATT
ATP11B	ATP_P1	34	PROMOTOR	18872976	rs23875728	G	A	G	A	AACTTTAGGAGAGGAGGACA GGTTTTTTGTTTAACCATAA
ATP11B	ATP_I5	34	INTRON 5-6	18880829	rs23881177	A	C	A	C	TGTTTCTTCCCTAATGAGAA AAATTCGATTATATCACTAT
ATP11B	ATP_I6	34	INTRON 6-7	18882628	rs23843542	G	A	G	A	ACTGGATTGTAACAGGCAAG ATGGAAAGCTCTGGATGGAT
ATP11B	ATP_I7	34	INTRON 7-8	18885412	rs23880628	T	G	T	G	ACATTCTGAATTTTAACCAA TACATTAGGCCATGGAATTT
ATP11B	ATP_I8	34	INTRON 8-9	18887779	rs23845156	A	C	A	C	TGTTAGAAAAGAAGCTAATT CTGTCTGGATTCACCTCTAT
ATP11B	ATP_I8B	34	INTRON 8-9	18888975	rs23882292	G	A	G	A	GTTCTAAAAGGATATAGACC AAAATAGTGGTCTTCTCTAG
ATP11B	ATP_I10	34	INTRON 10-11	18892361	rs23868830	A	C	A	C	TCCCTCTGCCTCTGCCCCTT CCCTGTTCATCATGCTCTCT
ATP11B	ATP_I15	34	INTRON 15-16	18918027	DC	T	G	T	G	TTTAAAGTGGTTATTGTAGG TTTTTTTTAATGCTTTTTGG
IL1α	IL1AE7X255	17	EXON 6 (Synon: 3' UTR)	40082405	rs8968929	G	A	G	A	ACGTGTACTATGTACATGGA GAGTCCAATCCTTTACTCAT
IL1α	IL1AE7X221	17	EXON 6 (Synon: 3' UTR)	40082438	AS	C	T	C	T	CTGGAGTCTATAACTTGTGA GTGTTGACAGTCCACGTGTA
IL1α	IL1A12227	17	INTRON3	40088450	AS	T	C	T	C	GTGCCTTTTATCCTTGTGAC GAAAGCAGTTACATACTACT
IL1α	IL1A11235	17	INTRON1	40089442	AS	G	C	G	C	TAATGTGGTCATTAAAACAA TGCAGAGATGTAACAAACAG
IL1α	IL1A10084	17	PROMOTOR	40090593	rs22530526	A	C	C	A	GCAGATAACACAAGGGAGTG AAAGAAGAATGGGGAAAATG
IL1α	IL1AA	17	PROMOTOR	40090951	AS	T	G	T	G	TGCCTGGTTTGGTGTGTGAG TATAATCACGGTCAGATTTC
IL1α	IL1AB	17	PROMOTOR	NA Repeat of IL1A10084	rs22530526	A	C	C	A	
IL1α	IL1AC	17	EXON 6	NA Repeat of IL1AE7X221	AS	C	T	C	T	
IL1α	IL1AD	17	EXON 6	NA Repeat of IL1AE7X255	rs8968929	G	A	G	A	
IL2	2_12Y206	19	INTRON 3-4	20764513	AS	T	C	T	C	ACAAATGAGGAAAACCTTGG TTATATGCTATCATCACTTG
IL4	4_25Y336	11	PRE GENE	23974956	AS	T	C	T	C	TGAAAAATCCTCAAAAAACT AAAATACAAATGCTACATGA

Gene	SNP_ID	CFA	Gene Position	Base Number	Source	MA LR	MJA LR	MA GR	MJA LR	Left and Right flanking sequences
IL4	4_22Y152	11	PROMOTOR	23976639	rs22146864	C	T	C	T	CATGTAGCCTTTTGTATCTG CTTCTTTCACTTACCCTAGT
IL4	4_12M397	11	INTRON 2-3	23982900	AS	A	C	A	C	TCCAGTTAGCTCCCCCACCC CCTCCATGGGAGGTGGCAAG
IL4	4_13S97	11	INTRON 2-3	23983033	AS	G	C	G	C	CATTTGTACTACCCCTTCCA ATTTTTTATAGTGAATTTAT
IL4	4_8R458	11	INTRON 3-4	23984915	AS	A	G	A	G	GAGGAAGCTTCTGGAAGAGG TGCAGTTGAGCTGGGCCATG
IL4	4_7S246	11	INTRON 3-4	23985620	rs22124242	G	C	C	G	ATGGAAGAATTGGGGACATT ATCCCCTTGCTGAGCCTGTC
IL4	4_2M351	11	POST GENE	23987559	AS	A	C	A	C	CATTTGTACTACCCCTTCCA ATTTTTTATAGTGAATTTAT
IL4	4_1K110	11	POST GENE	23988181	rs22189535	G	T	G	T	CATGTAGCCTTTTGTATCTG CTTCTTTCACTTACCCTAGT
IL6	6_6R431	14	PRE GENE	39431169	AS	A	G	A	G	TTTTGCAAGCATCACAGTGG GCTGGGAGAGGTGGCTTCAT
IL6	6_7S166	14	PRE GENE	39431770	AS	C	G	C	G	CGTGATTCAGAGCCTCAGAG CTTGTCTGTGTTTGGAGATT
IL6	6_7R485	14	PROMOTOR	39432089	AS	G	A	A	G	GTCTTCACCAGGGCCCCTGC CAGAGAGCAGGGCTGACGCT
IL6	6_8R289	14	PROMOTOR	39432418	AS	G	A	A	G	TCACAAATATGAATTAACTG AATGCTAAATCCTAGCCCGC
IL6	6_8W328	14	PROMOTOR	39432457	AS	T	A	T	A	GCTAATCTGGTAATTAAAGT TTTTTTTAATCATAGCCTTA
IL6	6_10Y257	14	EXON 2 (Synon: Asp -40)	39433189	AS	T	C	C	T	GCAGGAGATTCCAAGGATGA GCCACTTCAAATAGTCTACC
IL6	6_18R120	14	EXON 4 (Non synon: Arg - 196 [G] to Gin [A])	39436643	AS	G	A	G	A	AACAATTCACCTCATCCTGC GAGTCTGGAGGATTTCCTGC
IL6	6_20R191	14	POST GENE	39437322	AS	A	G	A	G	CTGTACACACTTTTATGGAC TAGGAGAAGGGACTTCCCAA
IL6	6_20R240	14	POST GENE	39437371	AS	G	A	G	A	AGCTAGAAGGTAAGGCACAG CCCAGATTTTAAATCCAGGT
IL6	6_20R412	14	POST GENE	39437543	AS	G	A	G	A	TTGAAATGACAACCACTTAT TAAGACCTAGCAATGTGCAC
IL10	10_14R553	7	POST GENE	8886928	rs24430111	G	A	G	A	GCACTTATTTTGAGCCAGCC GTGCTAGTTCTGTACATGTC
IL10	10_13Y85	7	POST GENE	8887853	AS	T	C	T	C	AGTCTTCCTATAAACTCAGT CTTTAAGACATTATCCTTAA
IL10	10_11R124	7	INTRON 4-5	8889092	AS	A	G	A	G	GCTTTTTAGGCCAACCCCGC CCGCCTCTCCCAAGGCACTG
IL10	10_10S308	7	INTRON 4-5	8889378	rs24433849	C	G	C	G	TTCCCAGAACAGGCGGCCTC GCCCTCTGCGGGGCTGAGCC
IL10	10_9R210	7	INTRON 3-4	8889995	AS	G	A	A	G	CCCCCAAGTGCCAGGGACAC GGAGCTGGGAGCCGTGGCAT
IL10	10_6R426	7	EXON 3 (Non synon: Ser - 77 [A] to Gly [G])	8891031	AS	A	G	A	G	CCCAACGCTYTTGCCTTTAG GTTACCTGGGTTGCCAAGCC
IL10	10_6Y135	7	EXON 2 (Synon: Thr -68)	8891320	rs24427871	C	T	C	T	AAGTCCTCCAGCAGGGACCC GTCAGCAGTATGTTGTCCAG
IL10	10_4Y100	7	EXON 1 (Non synon: Leu - 27 [T] to Pro [C])	8892274	AS	T	C	T	C	CCGACACCAGAGCACCCTAC TGAGGACGACTGCACCCACT
IL10	10_3M171	7	PROMOTER	8892673	AS	C	A	C	A	CTGGAAAGTTATTTTAAAAC GAGAGAGAGGTAGCTCATCC
IL10	10_2R420	7	PROMOTER	8892882	AS	G	A	G	A	AGCAAGGAAAAGCCTTGGGT TTCAATCCAGGTTGGGGAGG
IL10	10_1R218	7	PRE GENE	8893568	rs24382557	G	A	G	A	TCCTTTCCTTATTAGAGGTA AGCAACTTTCCTCACTGCAC
IL10	10_1R117	7	PRE GENE	8893668	rs24382554	A	G	A	G	GGGAGGTAGGAAAAGCTCCT TAGAAGGAGAAGGTCAAGGT
IL10	10_1R105	7	PRE GENE	8893680	AS	A	G	A	G	CTGTCTTCACTGGGGAGGTA GAAAAGCTCCTATAGAAGGA
IL12B	12B_01Y90	4	PRE GENE	54384501	AS	T	C	T	C	GGCACGACTTTTTACCCTAC ATTGTACACAAAACAGACAT
IL12B	12B_01M115	4	PRE GENE	54384525	AS	A	C	A	C	GTACACAAAACAGACATATC GATATTTCCTTTATCTCTTC
IL12B	12B_02Y146	4	PRE GENE	54385070	AS	C	T	C	T	TAGTCAGCGGCTTCTAACCA GTGTCAGAGAACATGGATGC
IL12B	12B_02Y190	4	PRE GENE	54385114	AS	T	C	T	C	CTGAGATGGATGGAGATGTT CAGGATGAGATGAAATGATA
IL12B	12B_02W232	4	PRE GENE	54385156	AS	A	T	A	T	ATATCTCTACCTAATTCAGA GTAGGGTACAGTTTTCACAT
IL12B	12B_02M407	4	PROMOTER	54385331	AS	A	C	C	A	AAGAACCTCTTGATTTTCAG GCTTATGGGCTTGAACATGG
IL12B	12B_03Y82	4	PROMOTER	54385528	AS	C	T	C	T	TTCCAGGTTACTTTGATGTG

Gene	SNP_ID	CFA	Gene Position	Base Number	Source	MA LR	MJA LR	MA GR	MJA LR	Left and Right flanking sequences
										ACTCAAGCTTGAGAATCACT
IL12B	12B_03R196	4	PROMOTER	54385642	AS	G	A	A	G	CTGGCCACCAGATCATTGCC TAATTTGAAATCACCTCTAA
IL12B	12B_03R462	4	PROMOTER	54385908	AS	G	A	G	A	CAGAGCCTTTTACATAGTCA TACCAAGTATATAATTGCTA
IL12B	12B_10R105	4	EXON 2 (Non synon: Val - 32 [G] to Ile [A])	54389975	AS	G	A	A	G	CCCTGTGCTCTCCAGTTTAT TTGTAGAGTTGGACTGGCAC
IL12B	12B_12Y142	4	EXON 3 (Non synon: Tyr - 138 [T] to His [C])	54390840	AS	T	C	T	C	TGAAATGTGAGGCAAAGAAT ATTCTGGACGTTTCACATGC
LEPR	LEPRC	5	EXON18 (Synon: Ser -1166)	47690381	AB	T	A	A	T	GAAAATAATGGTGAAAGGTC GTCTATTATTTAGGGGTCAC
LEPR	LEPRB	5	EXON18 (Non synon: His - 960 [A] to Pro [C])	47690733	AB	A	C	A	C	TATTTGTATCCGTGATCAGC CAACAGTGCTAACTTCTCTG
LEPR	LEPRA	5	EXON18 (Non synon: Thr - 926 [C] - Iso [T])	47690835	rs24233146	C	T	T	C	AGAAGATATCAGTGTTGACA ATCATGGAAGAATAAAGATG
LEPR	IL1AE	5		NA Repeat of LEPRA	rs2433146	C	T	T	C	
MMP13	MMP13A	5	PRE GENE	31860952	AB	A	G	A	G	TTGAGCCCAATTCTAATCTC AACTCATTGGAAATATTTAT
MMP13	MMP13B	5	PRE GENE	31861018	AB	A	G	A	G	TGTAATAATTGTCCTCCGCC ATAGCTTACATTTCATAAGC
MMP13	MMP13D	5	PROMOTOR	31861589	rs24280642	G	A	G	A	CTTGGAGTCAAATGCTGACA TATCATTTATKAACTTTTGA
MMP13	MMP13E	5	PROMOTOR	31861600	AB	T	G	T	G	ATGCTGACARTATCATTTAT AACTTTTGACCTTTATTAGA
MMP13	MMP13F	5	PROMOTOR	31861824	AB	G	A	G	A	GAACCTGTGGGGAATCCATG AGGATTCCATGTGTATCGAG
MMP13	MMP13G	5	PROMOTOR	31861947	AB	T	C	T	C	CTCATGTCTCCTTCTGTAAT GCTCTAGGGAAAATGATGTT
MMP13	MMP13H	5	PROMOTOR	31862036	AB	A	G	G	A	TTTAAGGAAGTGAGAGCATC TCTTGATACTCTTGTCTGAA
MMP13	MMP13J	5	PROMOTOR	31862069	AB	A	G	G	A	TGTCTGAAAAGARTAAAAGT GCTGCTTTTCTACAGAGGGA
MMP13	MMP13K	5	PROMOTOR	31862231	AB	C	T	T	C	ATGCCCTCATTTTATATTTC CTCAAATTCTACCACAAACC
MMP3	MMP3B	5	INTRON 9-10	31965257	rs24241537	A	G	A	G	TCAAGATCTGAGCCAAAATC AGTCAGACGCTTAAGAGACT
MMP3	MMP3A	5	POST GENE	31968448	rs24241538	A	C	A	C	GAGGGGATTGTGGCTGGACT ACGCAGTCACACTGCCTAGC
MMP9	MMP9_U2	24	UPSTREAM	36242295	rs23174490	G	T	G	T	AACCAGGGATTCTTAATCTG AAGGTCTCCCAGGCTTCAGG
MMP9	MMP9D	24	PROMOTOR	36249069	AB	C	T	C	T	TGTAAGCCCTTTCTTTGCTT CTCATGCTGGGGCCGCCCCC
MMP9	MMP9C	24	PROMOTOR	36249070	AB	A	G	A	G	TTTTATTTATTCATTCATGA AATCAGAGAGAGAGAGGGAG
MMP9	MMP9B	24	PROMOTOR	36249096	AB	A	G	A	G	GCCCCCTTGAGTAACGCTGC GCCGGTCCAGGGAGCTCCTG
MMP9	MMP9A	24	PROMOTOR	36250525	AB	A	T	A	T	TCAGATAAATAAATATTTTT WAAAAGATTTTATTTATTCA
MMP9	MMP9_I8	24	INTRON 8	36254234	rs23174499	G	A	G	A	ACCTCTCAATTTTCTCACCT TTAAATGGCACCTGCTCATA
MMP9	MMP9_D	24	DOWNSTREAM	36263716	rs23203666	C	T	C	T	CCTTTTNTTTTCTCACCGGA GGAATAGGTCCTCACATTGT
MMP9	MMP9_U1	24	UPSTREAM	NA Repeat of MMP9_I8	rs23174499	G	A	G	A	
MMP9	MMP9_P4	24	PROMOTOR	NA Repeat of MMP9A	AB	A	G	A	G	
MMP9	MMP9_P1	24	PRE GENE	NA Repeat of MMP9B	AB	A	T	A	T	
SPARC	SPARC_P2	15	PROMOTOR	60862114	rs24137749	A	G	A	G	AACCGCTGAGCCACCCAGGG TCCCAATATTCTCATAAATG
SPARC	SPARC_P3	15	PROMOTOR	60862154	rs24137747	A	T	A	T	GTTCAACCAAGACCTGTTGC CTGAGAAGTTCATGTAGCAT
SPARC	SPARC_P4	15	PROMOTOR	60862736	rs24117472	G	A	G	A	CAGTTCCCCTTGGAACAAGG GAAGATTGTTCCAGATGATT
SPARC	SPARC_P5	15	PROMOTOR	60862814	DC	G	A	G	A	GTGTAGGTAGCAGCCCCAGA CCCCTGAGGCAGCCAGKGTT
SPARC	SPARC_P7	15	PROMOTOR	60862846	rs24126270	A	T	T	A	GCCAGKGTTGGGGAGCCTAG CAGGACAGCACACCAGAAGA
SPARC	SPARC_I2	15	INTRON 2-3	60874270	DC	G	C	G	C	CTGTGTTTGGCTTCTCGTTC CCCTCATCACCCCACCCAGG

Gene	SNP_ID	CFA	Gene Position	Base Number	Source	MA LR	MJA LR	MA GR	MJA LR	Left and Right flanking sequences
SPARC	SPARC_I2B	15	INTRON 2-3	60874302	DC	G	A	A	G	CCACCCAGGGCAGTCCCTGA GCATAACCTTGGGGATCTCC
SPARC	SPARC_I3	15	INTRON 3-4	60874578	DC	G	A	G	A	GTGTATCCCCCAATCCTGAC CATTCCTGAATGCTTCCCAA
TIMP1	TIMP1A	X	PRE GENE	41126881	rs24610470	A	G	A	G	GGGTAGTAAAACTGGCAGGA GCCTATTTAGCTTTTTACAA
TIMP1	TIMP1B	X	POST GENE	41342699	rs24858175	C	T	C	T	AGAGGGTCTGTGCTGGCCTT CTTGTCCACAGGTTCAGGCA
TIMP2	TIMP2A	9	PRE GENE	5387352	rs22654395	G	A	G	A	GGGAGCCTGAAGTGGGACTC ATCCCAGGTCTCCAGGATCA
TIMP2	TIMP2_I1	9	INTRON 1-2	5498774	rs22640922	C	A	C	A	GGTAGGACGTGCTTCTTTCC YCTGGCCGGTTTCTGCCTCC
TIMP2	TIMP2_I1C	9	INTRON 1-2	5510716	rs22642411	A	G	A	G	TGTGGTTCTTCCAGTGAGTG AAGATGGCCCCGAGATGTGG
TIMP2	TIMP2_I2	9	INTRON 2-3	5521105	DC	C	A	C	A	GCCTCCCCCATCAAGTAGCC ACGCTTGATCAAAGTTCCTG
TIMP2	TIMP2_I3	9	INTRON 3-4	5523257	rs22668880	T	A	T	A	CCTGGAGGGAGGTGCACACA GTTCCCACTGTGGGGACAGG
TIMP3	TIMP3A	10	POST GENE	33749146	rs22007745	C	T	T	C	ATTCAGGTAGGTATTTGGCA CAGGTACTGATATTTGCAAC
TIMP3	TIMP3B	10	INTRON 2-3	33774314	rs22007793	C	A	A	C	TCAGGAGTGCTCAGAACACA CCCCCTGCTAAGTGCTCCAG
TIMP3	TIMP3C	10	INTRON 1-2	33807011	rs22007843	A	G	A	G	CGGGACTCGCTACGCCCTCC GAGCTGCACCCCAGTCCGGG
TIMP4	TIMP4_D	20	DOWNSTREAM	9373496	rs22864090	C	T	C	T	TTACACATGAGGAGATGAGG TAAGCTAAAGAAGCAAAGGA
TIMP4	TIMP4D	20	PRE GENE	9373496	rs22865393	C	T	C	T	AGTGGAGTGCTGCCAGGACC GGGGAGGCCCCTGTGCTATC
TIMP4	TIMP4_U	20	UPSTREAM	9402311	rs22918153	A	C	A	C	GACTCAGCTCTCTTTCAATT TTTTCTTTTCAAACATCACT
TIMP4	TIMP4_I1	20	INTRON 1-2	9404522	DC	A	G	G	A	CAGTTTTTGATTAATGCCCA ATGGCGAGAGATAAACTACC
TIMP4	TIMP4A	20	POST GENE	9534137	rs8874264	G	A	G	A	TGGTGTGAGGATTAAATAAG ATGTGGCAACATGCTTTAAA
TNC	TNC_D	11	DOWNSTREAM	72107869	DC	G	A	G	A	AGCAAAATCTTCCAGTGGCG CGTCACGGTGGTTGTCCTTC
TNC	TNC_E25	11	EXON 25 (Non synon: Val - 2118 [G] to Ile [A])	72112377	DC	G	A	G	A	ACGGGAAGACAGCCTACGCC TCTATGACAGGTTCAGCGTG
TNC	TNC_I24	11	INTRON 24-25	72112527	DC	C	T	C	T	TGCTGCTCCTATAGGGTGGA GGAATAGAAACAATAGCCTC
TNC	TNC_I4	11	INTRON 4-5	72165326	DC	T	G	G	T	CCACTTGTGAGTGCGCAGAA GCTCCCCCAGCCTCCACAGC
TNC	TNC_E2	11	EXON 2 (Synon: Arg -167)	72170870	DC	G	A	A	G	AATCTTGTGTCTTCCCTGCG GAGCAGTGCACCACGGGAGC
TNC	TNC_E1	11	EXON 1 (Synon: Ser -7)	72171352	rs22191688	C	G	G	C	CTACTCTGTGCTTCCACGTC CCACATCTGCCTGAGACTGA
TNC	TNC_P5	11	PROMOTOR	72171378	rs22191687	G	T	T	G	AGAGCACGAACAAGCTGATC GGATTCTACTCTGTGCTTCC
TNC	TNC_P4	11	PROMOTOR	72171896	DC	C	T	C	T	TCATTTGAGGCTCTGCCTCC AGTGCAGAGAGAGGAACCAC
TNC	TNC_P2	11	PROMOTOR	72171947	DC	C	T	C	T	TTAAGGAGCTGGGCTGGAAT AAAGTTCACCRTGCCTTCTA
TNC	TNC_P1	11	PROMOTOR	72172131	DC	T	C	T	C	GAGGCTCACGGACATTCTTT TGAGCCATCTCTTACCCTCC
TNFα	TNF6547	12	PRE GENE	4075592	AS	C	A	C	A	TTCGCTCTGTAGAAAAATCC GAAAAAAAAAATTGGTTTCA
TNFα	TNF7178	12	PRE GENE	4076221	AS	T	A	T	A	CCCAATAAACCTCTTTTCTC GAAATGCTGTCTATGTCTGT
TNFα	TNF8647	12	PROMOTOR	4077693	rs22216187	A	C	A	C	ACAAGGCCCCAGGGCTCTAC GTCTCCCCACTGGACTTGAG
TNFα	TNFEXON 1AB	12	EXON1 (Non synon: Val - 40 [G] to Ile [A])	4077844	AS	A	G	A	G	GCCTCTTCTCCTTCCTCCTC TCGCAGGGGCCACCACACTC
TNFα	TNF9367	12	INTRON 1-2	4078413	AS	T	C	T	C	TGCATAAGCTGTTTCTCCTA AGGGGTGACTTGCTCTGATG
TNFα	TNF9585	12	INTRON 2-3	4078631	AS	T	C	T	C	TGGAGGTCAAAGTAGTGGGA CTTTAAGGATCTCACCATTT
TNFα	TNF10252	12	EXON 4 (Non synon: Glu - 187 [A] to Val [T])	4079289	AS	A	T	A	T	TTGCCAAAGGGAGACCCCAG GGGGACCGAGGCCAAGCCCT
TNFα	TNFEXON 4AAB	12	EXON 4 (Synon: Ser -213)	4079368	AS	C	T	C	T	GAGAAGGGTGATCGACTCAG GCTGAGATCAATCTGCCTAA
TNFα	TNF10411	12	POST GENE	4079449	AS	A	G	A	G	AATCATTGCCCTGTAAGGGG TAGGACGTCCATTCTTGCCC

Gene	*SNP_ID*	*CFA*	*Gene Position*	*Base Number*	*Source*	*MA LR*	*MJA LR*	*MA GR*	*MJA LR*	*Left and Right flanking sequences*
TNFα	TNF10513	12	POST GENE	4079551	AS	A	G	A	G	AATTAAGGGCTCAGGGCTGG CCTCAAGCTTAGAACTTTAA
ZSWIM2	ZSWIM_E8A	36	EXON 8 (Synon: Leu -509)	32020319	DC	T	C	C	T	CACAATCCAAAGAAGACTCT GGTACTAGAATAAAAGAAGA
ZSWIM3	ZSWIM_E8B	36	EXON 8 (Synon: Lys -326)	32020865	rs9057634	G	A	G	A	GACAATTGGTTATTCTACAA TGCAACTCATGCCCTATTGA
ZSWIM4	ZSWIM_I6	36	INTRON 6-7	32023672	rs23938094	A	T	A	T	ATCTTCAGGGCCCACCTCTA TTCTAGTTTTTTTGCTGTTT
ZSWIM5	ZSWIM_I6B	36	INTRON 6-7	32023767	rs23938095	T	C	T	C	GGGCTGGAATCAACTTCTTC TAACTCTTGTTAATATTAAT
ZSWIM6	ZSWIM_I6C	36	INTRON 6-7	32028732	rs23941364	A	G	G	A	ATTTTATCCTGTCCTTTCAA ATGATCTTCTTTTGTATATG

Table 7

Primer and probe sequences used for Sequenom iPLEX genotyping.

PLEX No.	*SNP Identity*	*Forward Primer / Reverse Primer Sequences*	*Probe Sequence*	*Allele 1*	*Allele 2*
1	10_11R124	ACGTTGGATGTCGCTAGCCACGCTTTTTAG ACGTTGGATGTGAAGGATGGACCCAGGCAA	TAGCCACGCTTTTTAGGCCAACCCCGC	A	G
1	10_1R117	ACGTTGGATGGTCCCTTGATGTACCTTGAC ACGTTGGATGTGCTCTTCCTAGTTACTGTC	TACCTTGACCTTCTCCTTCTA	G	A
1	10_1R218	ACGTTGGATGCGCCCTCTCCTTTCCTTATT ACGTTGGATGTGTGTGTGTGTTTGAGGGTG	CTCCTTTCCTTATTAGAGGTA	A	G
1	10_4Y100	ACGTTGGATGACTGCTCTGTTGCTGCCTG ACGTTGGATGTGGGAAGTGGGTGCAGTCG	CAGCCGACACCAGAGCACCCTAC	C	T
1	IL-1014R553	ACGTTGGATGACAGCCGATGAGATGTTGAC ACGTTGGATGAATCCCATACCCTATGGCTG	TACTGAGCACTTATTTTGAGCCAGCC	A	G
1	12_03R196	ACGTTGGATGTGGTGGTGGGAGACAATTAG ACGTTGGATGGGAGAGAAACTAAACCTGGC	GACAATTAGAGGTGATTTCAAATTA	G	A
1	12_01Y90	ACGTTGGATGCAGCCAGGCACGACTTTTTA ACGTTGGATGATGTCAGCTTGTACCAAGGG	GCACGACTTTTTACCCTAC	C	T
1	12_02M407	ACGTTGGATGCCACACTTTGAGAACCACTG ACGTTGGATGGTCTTCTCCCAAAGAACCTC	TGTTCAAGCCCATAAGC	C	A
1	12_02Y190	ACGTTGGATGATGCTCTCTGAGATGGATGG ACGTTGGATGATGTGAAAACTGTACCCTAC	GAGATGGATGGAGATGTT	C	T
1	12_03Y82	ACGTTGGATGTAACAAGGCTTCCAGGTTAC ACGTTGGATGGCTCCAAACTCAAAGGTTAC	CCAGGTTACTTTGATGTG	C	T
1	12_10R105	ACGTTGGATGTGAGGACCACCATTTCTCCG ACGTTGGATGACAATCCAGTTCTCCACTCC	CCAGTCCAACTCTACAA	G	A
1	12_12Y142	ACGTTGGATGGATCTTTCTGAAATGTGAGGC ACGTTGGATGCAAATCAGTACTGATTGCCG	TTCTGAAATGTGAGGCAAAGAAT	C	T
1	4_25Y336	ACGTTGGATGGAATTACTGGATCATGTAGC ACGTTGGATGAAACTGGTGCAGCCACTATG	GATCATGTAGCATTTGTATTTT	C	T
1	4_7S246	ACGTTGGATGAAGAATCAGGTGACAGGCTC ACGTTGGATGGGAAGAGCTCAGAGTAGATG	GGCTCAGCAAGGGGAT	C	G
1	4_8R458	ACGTTGGATGCTGGTGCCCAGAAAATTGAG ACGTTGGATGGAACTCCTGATCTTCTGCTC	AAATTGAGCGCAGAGCAGTG	G	A
1	6_20R191	ACGTTGGATGCTTCTAGCTGGGTGACTTTG ACGTTGGATGTATGATGCTCAATCCCAGCC	GGTGACTTTGGGAAGTCCCTTCTCCTA	G	A
1	TNF10252	ACGTTGGATGATCAAGAGCCCTTGCCAAAG ACGTTGGATGTTCTCCAGTTGGAAGACCCC	GCCCTTGCCAAAGGGAGACCCCAG	A	T
1	TNF10411	ACGTTGGATGAGTGAGTGATCAAAGGGTCG ACGTTGGATGGGCAGGTGTACTTTGGAATC	GGTTTGGGCAAGAATGGACGTCCTA	G	A
1	TNF7178	ACGTTGGATGATCTGCACCTTCAACGAAGC ACGTTGGATGAAAATTCTCCCCTCCCAGAC	CGAAGCCCAATAAACCTCTTTTCTC	A	T
1	TNFEXON4AAB	ACGTTGGATGACTCGGCAAAGTCCAGATAG ACGTTGGATGGGTCTTCCAACTGGAGAAGG	TTAGGCAGATTGATCTCAGC	T	C
2	IL1A10084	ACGTTGGATGGAATGACTTAGCCCACACTC ACGTTGGATGGGAGGCAGATACATATGCAG	ATTTTCCCCATTCTTCTTT	C	A
2	IL1A11235	ACGTTGGATGACCGTGTGTGTTACCAAAGC ACGTTGGATGCTGTCAAACAAGATAATGAG	AAAGCTAATGTGGTCATTAAAACAA	C	G
2	IL1AE7X221	ACGTTGGATGTACATAGTACACGTGGACTG ACGTTGGATGCTTTCGGTTACTGGAAACCC	TACATAGTACACGTGGACTGTCAACAC	T	C
2	10_10S308	ACGTTGGATGCACCCTCTTCCCAGAACAG ACGTTGGATGGGGAGCAGGCCCTGCCCG	CAGAACAGGCGGCCTC	C	G
2	10_13Y85	ACGTTGGATGTACAGACGCCATAGTCTTCC ACGTTGGATGCCTTAGTCTTGAAAACCAGC	CGCCATAGTCTTCCTATAAACTCAGT	C	T
2	10_9R210	ACGTTGGATGAAGTGTTAATGCCACGGCTC ACGTTGGATGGAGTCTGGGCCCTTTTTCAG	CACGGCTCCCAGCTCC	G	A
2	12_01M115	ACGTTGGATGATGTCAGCTTGTACCAAGGG ACGTTGGATGGGCACGACTTTTTACCCTAC	GGGAAGAGATAAAGGAAATATC	C	A
2	12_02Y146	ACGTTGGATGTCTCCATCCATCTCAGAGAG ACGTTGGATGCTTCTTATGATTTAGTCAG	ATCCATGTTCTCTGACAC	T	C
2	4_22Y152	ACGTTGGATGCTCTCCCTACTGATTTCCTC ACGTTGGATGAATATGGTTGCAGGGCCTTC	TTGTGAAGGACAGAATCCA	C	T
2	4_2M351	ACGTTGGATGGTGAGGGTTCACTTCATTTG ACGTTGGATGGCACAGGTAATACAAGATCTG	TTGTACTACCCCTTCCA	C	A
2	IL-41K110	ACGTTGGATGGCCACTTCTGGATGTTTCAT ACGTTGGATGCGCTACAATATGGATGAACC	GAATCATGTAGCCTTTTGTATCTG	G	T

PLEX No.	*SNP Identity*	*Forward Primer / Reverse Primer Sequences*	*Probe Sequence*	*Allele 1*	*Allele 2*
2	6_18R120	ACGTTGGATGCTGAACTGCAGGAAATCCTC ACGTTGGATGTATCTTGCAGTCGCAGGATG	TGCAGGAAATCCTCCAGACTC	G	A
2	6_20R240	ACGTTGGATGTCACCCAGCTAGAAGGTAAG ACGTTGGATGGGGACCCTAAAGGTTAAGAG	AGCTAGAAGGTAAGGCACAG	A	G
2	6_20R412	ACGTTGGATGTTGGAAGTGCACATTGCTAG ACGTTGGATGAGGGAATGCATGTAAAGATG	AAGTGCACATTGCTAGGTCTTA	G	A
2	6_6R431	ACGTTGGATGAGCAATCCCACACTACAGAG ACGTTGGATGCTCTCCTGCGCTGAATGAAG	GAGGCTTTTTGCAAGCATCACAGTGG	A	G
2	6_7R485	ACGTTGGATGACTCTCTTGCTCACCTCTTC ACGTTGGATGAGATCCAAGTCTTCACCAGG	CAGCGTCAGCCCTGCTCTCTG	G	A
2	6_7S166	ACGTTGGATGTGTTTTGAGTCCAGAGGTGC ACGTTGGATGAAGAAAACCTAGGGCAAGCG	ATCTCCAAACACAGACAAG	G	C
2	6_8R289	ACGTTGGATGTTACCAGATTAGCGGGCTAG ACGTTGGATGGAAGCTCAGGTCTAAACGTC	GGGCTAGGATTTAGCATT	G	A
2	TNF6547	ACGTTGGATGCAGAATGGAGGCAAAATGGG ACGTTGGATGTGTCTTCTTTGGAGCCTTCG	GTCTTGAAACCAATTTTTTTTTTC	C	A
2	TNFEXON1AB	ACGTTGGATGTTCTGCCTCAGCCTCTTCTC ACGTTGGATGATCACTCCAAAGTGCAGCAG	TCTTCTCCTTCCTCCTC	A	G
3	IL1A12227	ACGTTGGATGATCCTTGTGACAGAAAGCAG ACGTTGGATGGTAACTTACAAAGGAGCATAG	ACATACTACTCATAAGCTATGTT	C	T
3	IL1AE7X255	ACGTTGGATGGCCTTGACTCTGGAGTCTAT ACGTTGGATGGCAAGTGACTATGAGTAAAGG	GGCCACGTGTACTATGTACATGGA	A	G
3	10_1R105	ACGTTGGATGTGCTCTTCCTAGTTACTGTC ACGTTGGATGGTCCCTTGATGTACCTTGAC	GGTCTTCACTGGGGAGGTA	A	G
3	10_2R420	ACGTTGGATGAATAATTGGATCCCCTCCCC ACGTTGGATGGAAACTGAGGCTCTTCCCAG	GAGACCCAACCTGGATTGAA	G	A
3	10_3M171	ACGTTGGATGGTTCACCCCAGGAAATCAAC ACGTTGGATGATTTTAGGATGAGCTACCTC	CCAGCTGGAAAGTTATTTTAAAAC	C	A
3	10_6R426	ACGTTGGATGACTGGATCATCTCCGACAGG ACGTTGGATGCAGCTCTTCCGCCCAGTCA	TGGCAACCCAGGTAAC	G	A
3	10_6Y135	ACGTTGGATGGCAGCAAATGAAGGACAAGC ACGTTGGATGGCTCTCACCTTAAAGTCCTC	GGACAACATACTGCTGAC	C	T
3	12_02W232	ACGTTGGATGTTACTATCCAGGGTTTGTGC ACGTTGGATGCAGGATGAGATGAAATGAT	TGTGAAAACTGTACCCTAC	T	A
3	12_03R462	ACGTTGGATGGCAGGAACATGACTTATTGG ACGTTGGATGTCTCGCTCAGAGCCTTTTAC	AACATTTAGCAATTATATACTTGGTA	G	A
3	2_12Y206	ACGTTGGATGGAATTCTTGTGTTCACTGAG ACGTTGGATGGTTGATACAAGTGATGATAGC	GACAAATGAGGAAAACCTTGG	C	T
3	4_12M397	ACGTTGGATGCTGGATATTGGTGCTTTGGG ACGTTGGATGCTTTGCAGACACTTGCCACC	GTTAGCTCCCCCACCC	C	A
3	4_13S97	ACGTTGGATGAGATCAGAGGAAGCTTCTGG ACGTTGGATGCTATACCTCCTAGGCCAAAG	GAAGCTTCTGGAAGAGG	C	G
3	6_10Y257	ACGTTGGATGTTTGCAGAGGTGAGTGGTAG ACGTTGGATGATGGCTACTGCTTTCCCTAC	GGGTGTAGACTATTTGAAGTGGC	T	C
3	6_8W328	ACGTTGGATGGGTGAGAAGCTAAGGCTATG ACGTTGGATGAATGCTAAATCCTAGCCCGC	GAAGCTAAGGCTATGATTAAAAAAA	T	A
3	TNF10513	ACGTTGGATGCTCACATCCCTGGATCTTAG ACGTTGGATGCCCTTCAGGCTTAGAAAGAG	GTTTAAAGTTCTAAGCTTGAGG	G	A
3	TNF8647	ACGTTGGATGCTAATATACAAGGCCCCAGG ACGTTGGATGCTTTCAGTGCTCATGGTGTG	AGGCCCCAGGGCTCTAC	C	A
3	TNF9367	ACGTTGGATGGGATGGATGGGAGAGAAAAC ACGTTGGATGAGGAGGTTTAGCATCAGAGC	CTGCATAAGCTGTTTCTCCTA	C	T
3	TNF9585	ACGTTGGATGTTCAGGCACTTGTTTGAGGG ACGTTGGATGGGTGAGATCCTTAAGCTTCC	GGAGGTCAAAGTAGTGGGA	C	T
4	MMP9D	ACGTTGGATGGACGTGGTGTAAGCCCTTTC ACGTTGGATGAAGTAGGGGCTGACTCAGG	AGCCCTTTCTTTGCTT	C	T
4	TIMP4D	ACGTTGGATGAGGAGAAAGCCAGTGGAGTG ACGTTGGATGCTGTTTAGCAGCACTGAAGG	GAGTGCTGCCAGGACC	C	T
4	TIMP3C	ACGTTGGATGTTTTGCGGTGCCCGGTCTTG ACGTTGGATGAAAGAAACGGGACTCGCTAC	GACTGGGGTGCAGCTC	G	A
4	MMP13B	ACGTTGGATGGGATCTGTTTTTGGACAAGG ACGTTGGATGCGGAACCAGACAAGCTTATG	CATAATTGTCCTCCGCC	A	G
4	IL1AA	ACGTTGGATGTCTTGTCTGTGCCTGGTTTG ACGTTGGATGCCTCAGTTCCTCATCAGTAG	CTGGTTTGGTGTGTGAG	G	T
4	TIMP2A	ACGTTGGATGTTCAGCTCAGGGTATGATCC ACGTTGGATGACAAGCAGAGAGAGAAGCAG	ATCCTGGAGACCTGGGAT	G	A

PLEX No.	*SNP Identity*	*Forward Primer / Reverse Primer Sequences*	*Probe Sequence*	*Allele 1*	*Allele 2*
4	MMP13D	ACGTTGGATGAAGTCTACTGGCTTGGAGTC ACGTTGGATGAGAATAATCTGAGCACAGAG	TGGAGTCAAATGCTGACA	A	G
4	MMP13G	ACGTTGGATGCAGGTACTCATGTCTCCTTC ACGTTGGATGCTAGAACCTGGTGACTTCTG	TCATGTCTCCTTCTGTAAT	C	T
4	MMP2A	ACGTTGGATGCTCAGAGCTTGTTTTGCCAG ACGTTGGATGCCCCAGCCCCCTTGAGTAA	TTTAGCTCCCTGGACCGGC	G	A
4	LEPRB	ACGTTGGATGCTCCAGCTCAGAGAAGTTAG ACGTTGGATGCTCTGCTTTTGACAACTCCG	GGAGAAGTTAGCACTGTTG	C	A
4	IL1AC	ACGTTGGATGCTTTCGGTTACTGGAAACCC ACGTTGGATGCATGTACATAGTACACGTGG	CTGGAGTCTATAACTTGTGA	C	T
4	IL1AB	ACGTTGGATGGAATGACTTAGCCCACACTC ACGTTGGATGGGAGGCAGATACATATGCAG	GAATTTTCCCCATTCTTCTTT	C	A
4	TIMP3A	ACGTTGGATGTACACCACGGAAATGAGAAC ACGTTGGATGAGCCATAGTTGGAATTCAGG	AGTTGCAAATATCAGTACCTG	T	C
4	MMP3A	ACGTTGGATGATACTCATCAGGCTAGGCAG ACGTTGGATGGGAGGTTTTGAGGGAAGTTG	GGCGAGGCAGTGTGACTGCGT	C	A
4	MMP13H	ACGTTGGATGTCCCTCTGTAGAAAAGCAGC ACGTTGGATGGAAGCAGGACTAAGTATCTC	GTTTCAGACAAGAGTATCAAGA	G	A
4	TIMP1A	ACGTTGGATGTTCTCTGTCCCTTCCAGATG ACGTTGGATGCCAAGGCCCTTAAGAACAAC	ATCTTGTAAAAAGCTAAATAGGC	G	A
4	MMP9C	ACGTTGGATGTCTCTCTCTCCCTCTCTCTC ACGTTGGATGATGCTCTCCCTCTCTCTCAG	GGATCTCCCTCTCTCTCTCTGATT	G	A
4	LEPRC	ACGTTGGATGTTTTTGATTGAGGTGACCCC ACGTTGGATGGACCTTTTGAAACTTGAGGG	TGAGGTGACCCCTAAATAATAGAC	T	A
4	LEPRA	ACGTTGGATGGGTCCTCTTCTTTTGGAACC ACGTTGGATGGTTGTTGGCACCATCTCATC	GGGAAGAAGATATCAGTGTTGACA	C	T
4	MMP13K	ACGTTGGATGTTTGAGACCCCGCTGAAATG ACGTTGGATGTCCCAAGAAAGTGTGGTTTG	AAGAGATGCCCTCATTTTATATTTC	C	T
4	TIMP4A	ACGTTGGATGATCTCATGATTGGTGTGAGG ACGTTGGATGCAGGTGCCCTGAGTGTTTTG	CTGCTTGGTGTGAGGATTAAATAAG	A	G
4	MMP13F	ACGTTGGATGTTCCATATAGGCTCCTACCC ACGTTGGATGGTGACTCTCGATACACATGG	CTTACTGAACCTGTGGGGAATCCATG	A	G
4	MMP13A	ACGTTGGATGCCTTGTCCAAAAACAGATCC ACGTTGGATGTTTGTGTTGCTGGTTGAGCC	CCTTACAATAAATATTTCCAATGAGTT	G	A
5	TIMP1B	ACGTTGGATGCAGACCTTTCTCCTTTTGCC ACGTTGGATGAAGAACAGGGAGTTAGAGGG	TGAACCTGTGGACAAG	T	C
5	MMP3B	ACGTTGGATGATCTGTCATTACATGGGCCG ACGTTGGATGCTGGAAATCAAGATCTGAGC	CTCTTAAGCGTCTGACT	G	A
5	IL1AD	ACGTTGGATGGACAGTCCACGTGTACTATG ACGTTGGATGCAAAGGCTCAGCACATTCTC	CGTGTACTATGTACATGGA	A	G
5	IL1AE	ACGTTGGATGGGTCCTCTTCTTTTGGAACC ACGTTGGATGGTTGTTGGCACCATCTCATC	AGAAGATATCAGTGTTGACA	C	T
5	MMP13J	ACGTTGGATGGGGCAAAAATATTTTTCCCTC ACGTTGGATGCTCTTTAAGGAAGTGAGAGC	TTCCCTCTGTAGAAAAGCAGC	G	A
5	MMP13E	ACGTTGGATGAGAATAATCTGAGCACAGAG ACGTTGGATGTGGCTTGGAGTCAAATGCTG	ACCTCTAATAAAGGTCAAAAGTT	G	T
5	TIMP3B	ACGTTGGATGACTGTCATGCAGGCACATAG ACGTTGGATGGGATAAAATTTCCCATGGGC	AGGTCAGGAGTGCTCAGAACACA	C	A
5	MMP9A	ACGTTGGATGATGCTCTCCCTCTCTCTCAG ACGTTGGATGTCTCTCCCTCTCTCTCTG	CTCTCTCTCAGATAAATAAATATTTTT	A	T
6	SPARC_I3	ACGTTGGATGCACAGCACAGATGTGAATCC ACGTTGGATGAGTGTGCCATGGAAGGTTTG	TCCCCCAATCCTGAC	A	G
6	SPARC_P7	ACGTTGGATGTCTCTCCTCCTCTTCTTCTG ACGTTGGATGCATGAGTCTGTGTAGGTAGC	TGGTGTGCTGTCCTG	T	A
6	MMP9_P4	ACGTTGGATGGACGTGGTGTAAGCCCTTTC ACGTTGGATGAAGTAGGGGCTGACTCAGG	AGCCCTTTCTTTGCTT	C	T
6	ATP_I10	ACGTTGGATGCTCACTCAGGAGTCTGCTTC ACGTTGGATGGCAAGAGAGAGCATGATGA	AGCTGCCTCTGCCCCTT	C	A
6	SPARC_I2	ACGTTGGATGAGCTCACTGGAGATCCCCAA ACGTTGGATGCTGTGTTTGGCTTCTCGTTC	GATCCCCAAGGTTATGC	G	C
6	ANK_E4	ACGTTGGATGGTCACACAGCCTGAGAAATG ACGTTGGATGAGAAACTGGGTGCACTCTTG	AGTGCCAGTGGCCTGAC	C	A
6	MMP9_I8	ACGTTGGATGTAGGTGACCATGGAAGTTAC ACGTTGGATGAACCTCCCAACAACCCTATG	CCTCTCAATTTTCTCACCT	A	G
6	TIMP4_I1	ACGTTGGATGCTGCACAGAGGTAGTTTATC ACGTTGGATGTTAGACCAGGCACAGAGAAG	GTAGTTTATCTCTCGCCAT	G	A

PLEX No.	SNP Identity	Forward Primer / Reverse Primer Sequences	Probe Sequence	Allele 1	Allele 2
6	TNC_E2	ACGTTGGATGGAGCTGGAGAATCTTGTGTC ACGTTGGATGTGTACCTTCGGCAGGCTGGA	TTCGTGTGTCTTCCCTGCG	A	G
6	TNC_P5	ACGTTGGATGGGGAGTTCTTCAAGAAGAGC ACGTTGGATGTCAGTCTCAGGCAGATGTGG	GAGCACGAACAAGCTGATC	G	T
6	ANK_I1	ACGTTGGATGGCACAAAATAGTCTCTGAAG ACGTTGGATGCCACTTTCAAGCACCTCTAC	AGGAGTCTCTGAAGAGCGTA	G	A
6	ZSWIM_E8A	ACGTTGGATGATGGCCAAAGACTGTAAACG ACGTTGGATGTTCTGGGAGGAAAGTACTTG	CCACAATCCAAAGAAGACTCT	C	T
6	TNC_E25	ACGTTGGATGTTGTAGCGAGTCTTGGCATC ACGTTGGATGAGTACGAGCTCCGGGTAGAC	ACACGCTGAACCTGTCATAGA	G	A
6	TIMP2_I1C	ACGTTGGATGTCTCTGACCTGCTGTGGTTC ACGTTGGATGCAGGATACTGTCCACATCTC	TTGGGGTTCTTCCAGTGAGTG	A	G
6	ATP_I6	ACGTTGGATGGACTGTCATAGACCTATCAC ACGTTGGATGGACATCTCGTCTATCCATCC	CAACGGGATTGTAACAGGCAAG	A	G
6	TIMP4_D	ACGTTGGATGATCTAGGAACCTGTGCTCTC ACGTTGGATGTCTTGTAGACCAAACATCTC	CCCCCCTTTGCTTCTTTAGCTTA	T	C
6	TIMP2_I1	ACGTTGGATGTATCAAGACAACCCGGTAGG ACGTTGGATGCCAAGCTAGACCAAGAACAC	CCCCGTAGGACGTGCTTCTTTCC	C	A
6	ANK_I3	ACGTTGGATGATCCTCCTCTCAAGCAGAAG ACGTTGGATGCGATGAAGGTACATAAGTGG	CAAGCAGAAGATCTCTTAATACC	C	T
6	ATP_I8B	ACGTTGGATGTATAACACCTAGAGAAGACC ACGTTGGATGGCAGAAGGAAAGTTCTAAAAG	GACCTAGAGAAGACCACTATTTT	G	A
6	MMP9_D	ACGTTGGATGTGGGTTGACCACAATGTGAG ACGTTGGATGGAGCTATTCTCGATGCTTCC	GGGAACAATGTGAGGACCTATTCC	T	C
6	ATP_I15	ACGTTGGATGACTTTTGGTTTTAAAGTGG ACGTTGGATGCTGGAGGAGATGAAGCAGAA	TGGTTTTAAAGTGGTTATTGTAGG	G	T
6	TNC_P4	ACGTTGGATGGCATGTTCATCATTTGAGGC ACGTTGGATGTGAGGCTGTAGGCACAATAG	GTTAATCATTTGAGGCTCTGCCTCC	C	T
6	TNC_D	ACGTTGGATGCACCCAAGAGAAGGACAACC ACGTTGGATGCTCCACAGGTGTCTGTTTTG	GGATAGAAGGACAACCACCGTGACG	G	A
6	ZSWIM_I6C	ACGTTGGATGGATTAATTTTATCCTGTCC ACGTTGGATGATTAGCCTATTAGGTACAC	GTATTAATTTTATCCTGTCCTTTCAA	A	G
6	TNC_I24	ACGTTGGATGACAGGCAAGGGTTACAGAAG ACGTTGGATGTTATAGCCCAACTGCTGCTC	GGGGATGAGGCTATTGTTTCTATTCC	T	C
6	TIMP4_U	ACGTTGGATGCCCCAAATCTCATCCTTCTC ACGTTGGATGCCCATTTTGATGTGATAGTG	TCCCTCAGACTCAGCTCTCTTTCAATT	C	A
6	ATP_I7	ACGTTGGATGTACTCTCAAACACTTCTCAC ACGTTGGATGCCCTCTGAAAATTCCATGGC	AATGCTCACATTCTGAATTTTAACCAA	G	T
6	ZSWIM_I6	ACGTTGGATGGGAAGTAACTGCAGATGTGG ACGTTGGATGATTCAGTCAATCTTCAGGGC	GGGATGAAAACAGCAAAAAAACTAGAA	T	A
6	SPARC_P3	ACGTTGGATGGGCTAAGTCCATCACAATTC ACGTTGGATGCATAAATGTTCAACCAAGACC	GAAGTCTTATGCTACATGAACTTCTCAG	T	A
7	MMP9_U2	ACGTTGGATGCTATACAGGTTCCCTGAAGC ACGTTGGATGTCCTCAGCCTTGGGAAATAC	AGCCTGGGAGACCTT	G	T
7	TNC_I4	ACGTTGGATGAAGAAGAGCATTCCCGTCAG ACGTTGGATGCAAACCAAGGGCAATAGAGG	TGTGAGTGCGCAGAA	G	T
7	SPARC_I2B	ACGTTGGATGTGGAGATCCCCAAGGTTATG ACGTTGGATGTCCAATGAGCCCTGTGTTTG	GTGGGGTGATGAGGG	G	A
7	MMP9_U1	ACGTTGGATGAACCTCCCAACAACCCTATG ACGTTGGATGTAGGTGACCATGGAAGTTAC	AGCAGGTGCCATTTAA	G	A
7	ANK_E4B	ACGTTGGATGAGAAACTGGGTGCACTCTTG ACGTTGGATGGTCACACAGCCTGAGAAATG	ATGGGTTTGAGCAATGT	C	A
7	ATP_P1	ACGTTGGATGAAGCTTAACTTTAGGAGAGG ACGTTGGATGTCCACCTAGAGATGGAAGAG	TTTAGGAGAGGAGGACA	A	G
7	ZSWIM_I6B	ACGTTGGATGCATTCGTGAGGGCTGGAATC ACGTTGGATGGAACATTATTCATGGAACAG	GCTGGAATCAACTTCTTC	C	T
7	ANK_I4	ACGTTGGATGCCTTTCTCCTCAGGAACAAC ACGTTGGATGCTGGGAGTGTACTTAATGTG	TCAGGAACAACTAATGCT	G	A
7	ATP_I8	ACGTTGGATGGACGTTATAGAGGTGAATCC ACGTTGGATGGCAAAGGTGAAAATTAGGAGG	TAGAGGTGAATCCAGACAG	C	A
7	SPARC_P4	ACGTTGGATGCCACAGGAATCATCTGGAAC ACGTTGGATGAGCCTTTGGGTAATCCCTTG	AATCATCTGGAACAATCTTC	G	A
7	TNC_P1	ACGTTGGATGTTGGAGATCGGAGAACACAC ACGTTGGATGAGTCCCAACTATGATGGAGG	GAGGCTCACGGACATTCTTT	C	T
7	TNC_P2	ACGTTGGATGCCACCTAACAGTTAAGGAGC ACGTTGGATGAGGCAGAGCCTCAAATGATG	AGACGGAGCTGGGCTGGAAT	C	T

PLEX No.	*SNP Identity*	*Forward Primer / Reverse Primer Sequences*	*Probe Sequence*	*Allele 1*	*Allele 2*
7	SPARC_P5	ACGTTGGATGTCTTCTTCTGGTGTGCTGTC ACGTTGGATGCATGAGTCTGTGTAGGTAGC	GGGGTCTGGCTGCCTCAGGGG	G	A
7	TNC_E1	ACGTTGGATGAGAAGAGCACGAACAAGCTG ACGTTGGATGGGTTTCTGCGTTCTTGATCC	GTCGACTCTGTGCTTCCACGTC	C	G
7	ZSWIM_E8B	ACGTTGGATGGGAAATGTATTGACAATTGG ACGTTGGATGCTTGTCCATCAATAGGGCAT	GTGACAATTGGTTATTCTACAA	A	G
7	SPARC_P2	ACGTTGGATGAAGGCAGGCGCTAAACCGCT ACGTTGGATGCAACAGGTCTTGGTTGAACA	CTAACCCGCTGAGCCACCCAGGG	A	G
7	TIMP2_I2	ACGTTGGATGGGGTATTGGCAGGAACTTTG ACGTTGGATGTGAGAAGTCAAACCCCACTG	GGTCAGGAACTTTGATCAAGCGT	C	A
7	TIMP2_I3	ACGTTGGATGTGACGGGAGTTCAACAGGAG ACGTTGGATGCTAGTTTCTTCTCCTGTCCC	GGGACTGGAGGGAGGTGCACACA	A	T
7	ANK_E5	ACGTTGGATGCGATGTGGAAGATGAAGCTG ACGTTGGATGAACAGAGAACTGTCCACTCC	CTCTTAAGCTGACAAAATGCTGGT	G	C
7	ANK_I2	ACGTTGGATGAGAGGATAATTGTTTTGGG ACGTTGGATGCACTCCACATGGAAGTCAAG	ACTTGTTGGGTTTTAAGTACAACAC	A	G
7	ATP_I5	ACGTTGGATGGGTAAATCCTAGACTCCCTG ACGTTGGATGGATTATTCCTGTTTCTTCCC	ACAGTATAGTGATATAATCGAATTT	C	A
7	MMP9_P1	ACGTTGGATGATGCTCTCCCTCTCTCTCAG ACGTTGGATGTCATGAATGAATAAATAAAA	CTCTCTCTCAGATAAATAAATATTTTT	A	T

Table 8

Comparisons of each SNP in the control and cranial cruciate ligament rupture (CCLR) Golden Retriever populations. The SNP identity, genotyping rate (call rate), Hardy Weinberg equilibrium P value (HWE) (by Chi squared analysis), minor allele frequencies (MAF) for control (C) and cranial cruciate ligament disease (CCLD) populations, and the two populations stratified on the basis of neuter status (neutered only are denoted S) are listed. Case - control comparison by Chi squared analysis (P value), with Monte Carlo correction for significant associations (denoted CV) are also listed. SNPs are noted to be heterozygous if the MAF > 1% (YES), low (LOW) if MAF < 1%, and not (NO) if homozygous. SNPs were the genotyping test failed (FAIL), the reasons for failure (waters genotyping, low genotyping rate or no probe in the reaction are listed), and SNPs out of Hardy Weinberg Equilibrium are also indicated (OUT OF HWE).

SNP Identity	*Call Rate*	*HWE*	*MAF C*	*MAF CCLR*	*OR (95% CI)*	*P Value*	*MAF C S*	*MAF CCLR S*	*OR S (95% CI)*	*P Value S*	*Heterozygous*
10_10S308	0.950	0.496	0.231	0.159	0.63 (0.32 - 1.25)	0.181	0.176	0.185	1.06 (0.35 - 3.24)	0.918	YES
10_11R124	0.705	0.681	0.225	0.167	0.69 (0.3 - 1.56)	0.367	0.231	0.139	0.54 (0.14 - 2)	0.350	YES
10_13Y85	0.942	0.503	0.412	0.250	0.48 (0.26 - 0.85)	0.012, CV 0.017	0.353	0.278	0.71 (0.28 - 1.77)	0.457	YES
10_14R553	0.935	0.999	0.394	0.225	0.45 (0.24 - 0.82)	0.008, CV 0.017	0.353	0.250	0.61 (0.24 - 1.57)	0.304	YES
10_1R105	0.978	0.484	0.419	0.279	0.54 (0.31 - 0.93)	0.026, CV 0.027	0.375	0.293	0.69 (0.28 - 1.72)	0.426	YES
10_1R117	0.691	0.000	0.152	0.083	0.51 (0.18 - 1.43)	0.193	0.136	0.075	0.51 (0.09 - 2.79)	0.434	OUT OF HWE
10_1R218	0.957	0.807	0.172	0.105	0.56 (0.25 - 1.24)	0.149	0.176	0.089	0.46 (0.13 - 1.63)	0.221	YES
10_2R420	0.993	0.378	0.414	0.267	0.51 (0.3 0- 0.89)	0.017, CV 0.024	0.344	0.283	0.75 (0.30 - 1.9)	0.549	YES
10_3M171	0.986	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
10_4Y100	0.978	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
10_6R426	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
10_6Y135	1.000	0.941	0.176	0.111	0.59 (0.28 - 1.25)	0.165	0.176	0.100	0.52 (0.15 - 1.76)	0.286	YES
10_9R210	0.950	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
12B_01M115	0.950	0.040	0.044	0.000	NA	0.054	0.059	0.000	NA	0.071	OUT OF HWE
12B_01Y90	0.964	0.000	0.022	0.011	0.49 (0.05 - 4.44)	0.516	0.059	0.017	0.27 (0.02 - 3.11)	0.264	OUT OF HWE
12B_02M407	1.000	0.022	0.298	0.333	1.18 (0.69 - 2.02)	0.549	0.382	0.317	0.75 (0.31 - 1.8)	0.518	OUT OF HWE
12B_02W232	0.986	0.378	0.414	0.318	0.66 (0.39 - 1.13)	0.128	0.441	0.293	0.53 (0.22 - 1.27)	0.150	YES
12B_02Y146	0.935	0.915	0.011	0.013	1.13 (0.1 - 12.61)	0.923	0.029	0.019	0 62 (0.04 - 10.3)	0.738	YES
12B_02Y190	0.993	0.440	0.415	0.318	0.66 (0.39 - 1.12)	0.124	0.441	0.293	0.53 (0 22 - 1.27)	0.150	YES

SNP Identity	*Call Rate*	*HWE*	*MAF C*	*MAF CCLR*	*OR (95% CI)*	*P Value*	*MAF C S*	*MAF CCLR S*	*OR S (95% CI)*	*P Value S*	*Heterozygous*
12B_03R196	0.986	0.015	0.296	0.352	1.3 (0.76 - 2.22)	0.346	0.382	0.328	0.79 (0.33 - 1.9)	0.594	OUT OF HWE
12B_03R462	0.993	0.141	0.293	0.330	1.19 (0.69 - 2.05)	0.534	0.176	0.379	2.85 (1.02 - 7.98)	0.041, CV 0.042	YES
12B_03Y82	0.978	0.378	0.414	0.314	0.65 (0.38 - 1.11)	0.114	0.441	0.293	0.53 (0.22 - 1.27)	0.150	YES
12B_10R105	1.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
12B_12Y142	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
2_12Y206	1.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
4_12M397	1.000	0.309	0.420	0.333	0.69 (0.41 - 1.17)	0.165	0.353	0.333	0.92 (0.38 - 2.22)	0.847	YES
4_13S97	1.000	0.204	0.059	0.056	0.95 (0.32 - 2.81)	0.921	0.029	0.017	0.56 (0.03 - 9.24)	0.681	YES
4_1K110	0.950	0.462	0.423	0.329	0.67 (0.39 - 1.16)	0.149	0.353	0.315	0.84 (0.34 - 2.09)	0.711	YES
4_22Y152	0.957	0.462	0.423	0.333	0.68 (0.4 - 1.17)	0.164	0.353	0.321	0.87 (0.35 - 2.14)	0.758	YES
4_25Y336	0.957	0.098	0.478	0.439	1.39 (0.83 - 2.35)	0.213	0.438	0.389	0.82 (0.34 - 1.99)	0.657	YES
4_2M351	0.957	0.462	0.423	0.333	0.68 (0.4 - 1.17)	0.164	0.353	0.321	0.87 (0.35 - 2.14)	0.758	YES
4_7S246	0.993	0.547	0.479	0.386	0.69 (0.41 - 1.15)	0.151	0.382	0.345	0.85 (0.35 - 2.05)	0.717	YES
4_8R458	0.978	0.213	0.060	0.057	0.95 (0.32 - 2.82)	0.923	0.029	0.017	0.58 (0.04 - 9.57)	0.699	YES
6_10Y257	1.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
6_18R120	0.914	0.000	0.022	0.027	1.22 (0.22 - 6.82)	0.819	0.000	0.042	NA	0.242	FAIL LOW GENOTYPING RATE
6_20R191	0.957	0.186	0.394	0.465	1.33 (0.79 - 2.24)	0.274	0.412	0.466	1.24 (0.53 - 2.93)	0.617	YES
6_20R240	0.957	0.502	0.341	0.452	1.6 (0.94 - 2.71)	0.080	0.353	0.446	1.48 (0.61 - 3.56)	0.382	YES
6_20R412	0.957	0.317	0.330	0.452	1.68 (0.99 - 2.85)	0.054	0.353	0.446	1.48 (0.61 - 3.56)	0.382	YES
6_6R431	0.950	0.460	0.072	0.071	0.99 (0.36 - 2.7)	0.981	0.059	0.089	1.57 (0.29 - 8.57)	0.601	YES
6_7R485	0.942	0.539	0.060	0.063	1.04 (0.35 - 3.09)	0.949	0.029	0.093	3.37 (0.38 - 30.15)	0.252	YES
6_7S166	0.950	0.916	0.011	0.012	1.11 (0.1 - 12.43)	0.932	0.000	0.018	NA	0.433	YES
6_8R289	0.950	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
6_8W328	1.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
ANK_E4	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
ANK_E4B	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
ANK_E5	0.978	0.935	0.108	0.047	0.4 (0.13 - 1.22)	0.099	0.125	0.054	0.4 (0.08 - 1.9)	0.234	YES
ANK_I1	0.986	0.607	0.447	0.302	0.54 (0.31 - 0.92)	0.024, CV 0.025	0.500	0.357	1.8 (0.76 - 4.28)	0.182	YES
ANK_I2	0.986	0.333	0.457	0.314	0.54 (0.32 - 0.93)	0.025, CV 0.043	0.471	0.375	1.88 (0.79 - 4.45)	0.152	YES

SNP Identity	*Call Rate*	*HWE*	*MAF C*	*MAF CCLR*	*OR (95% CI)*	*P Value*	*MAF C S*	*MAF CCLR S*	*OR S (95% CI)*	*P Value S*	*Heterozygous*
ANK_I3	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
ANK_I4	0.971	0.961	0.207	0.186	0.88 (0.46 - 1.68)	0.695	0.188	0.214	1.18 (0.4 - 3.53)	0.764	YES
ATP_I10	0.993	0.611	0.309	0.273	0.84 (0.48 - 1.47)	0.544	0.412	0.293	0.59 (0.24 - 1.44)	0.245	YES
ATP_I15	0.978	0.611	0.309	0.250	0.75 (0.42 - 1.34)	0.326	0.412	0.259	0.5 (0.2 - 1.25)	0.135	YES
ATP_I5	0.993	0.779	0.301	0.289	0.94 (0.54 - 1.64)	0.835	0.375	0.317	0.77 (0.31 - 1.9)	0.573	YES
ATP_I6	0.281	1.000	0.000	0.036	NA	0.179	0.000	0.042	NA	0.557	FAIL LOW GENOTYPING RATE
ATP_I7	0.986	0.611	0.309	0.256	0.77 (0.43 - 1.37)	0.373	0.412	0.268	0.52 (0.21 - 1.29)	0.157	YES
ATP_I8	0.978	0.933	0.310	0.273	0.84 (0.48 - 1.47)	0.532	0.375	0.293	0.69 (0.28 - 1.72)	0.426	YES
ATP_I8B	0.978	0.644	0.312	0.233	0.67 (0.37 - 1.2)	0.179	0.412	0.250	0.48 (0.19 - 1.19)	0.108	YES
ATP_P1	0.856	0.001	0.266	0.275	1.05 (0.57 - 1.92)	0.880	0.133	0.278	2.5 (0.75 - 8.38)	0.129	FAIL LOW GENOTYPING RATE
IL1A10084	0.957	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1A11235	0.942	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1A12227	1.000	0.630	0.250	0.189	0.7 (0.37 - 1.3)	0.257	0.294	0.150	0.42 (0.15 - 1.18)	0.095	YES
IL1AA	0.921	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1AB	0.957	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1AC	0.957	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1AD	0.906	0.827	0.023	0.039	1.77 (0.39 - 8.09)	0.458	0.033	0.038	1.16 (0.1 - 13.36)	0.905	YES
IL1AE	0.971	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1AE7X221	0.849	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
IL1AE7X255	0.978	0.831	0.022	0.034	1.59 (0.35 - 7.25)	0.547	0.031	0.034	1.11 (0.1 - 12.71)	0.935	YES
LEPRA	0.950	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
LEPRB	0.978	0.874	0.016	0.012	0.72 (0.07 - 7)	0.774	0.029	0.018	0.6 (0.04 - 9.92)	0.718	YES
LEPRC	0.957	0.148	0.242	0.288	1.26 (0.7 - 2.28)	0.435	0.235	0.278	1.25 (0.46 - 3.37)	0.659	YES
MMP13A	0.950	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
MMP13B	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
MMP13D	0.950	1.000	0.000	0.013	NA	NA	0.000	0.000	NA	NA	NO
MMP13E	0.914	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
MMP13F	0.978	0.958	0.005	0.000	NA	NA	0.000	0.000	NA	NA	LOW
MMP13G	0.964	1.000	0.000	0.023	NA	NA	0.000	0.018	NA	0.447	NO
MMP13H	0.971	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
MMP13J	0.899	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	LOW
MMP13K	0.964	0.002	0.277	0.417	1.86 (1.08 - 3.2)	0.023	0.235	0.375	1.95 (0.75 - 5.09)	0.169	OUT OF HWE
MMP2A	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
MMP3A	0.971	0.973	0.145	0.131	0.89 (0.42 - 1.89)	0.756	0.088	0.161	1.98 (0.5 - 7.89)	0.327	YES
MMP3B	0.986	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO

SNP Identity	*Call Rate*	*HWE*	*MAF C*	*MAF CCLR*	*OR (95% CI)*	*P Value*	*MAF C S*	*MAF CCLR S*	*OR S (95% CI)*	*P Value S*	*Heterozygous*
MMP9_D	0.993	0.002	0.223	0.159	0.66 (0.34 - 1.28)	0.216	0.147	0.103	0.67 (0.19 - 2.38)	0.534	OUT OF HWE
MMP9_I8	0.770	0.000	0.027	0.029	1.08 (0.19 - 6.02)	0.934	0.059	0.050	0.84 (0.11 - 6.32)	0.867	OUT OF HWE
MMP9_P1	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9_P4	0.784	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
MMP9_U1	0.763	0.000	0.021	0.014	0.67 (0.07 - 6.58)	0.731	0.029	0.024	0.8 (0.05 - 13.36)	0.879	OUT OF HWE
MMP9_U2	1.000	0.219	0.484	0.489	1.02 (0.62 - 1.69)	0.940	0.412	0.450	1.75 (0.75 - 4.09)	0.198	YES
MMP9A	0.281	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE
MMP9C	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9D	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
SPARC_I2	0.986	0.523	0.176	0.163	0.91 (0.46 - 1.81)	0.795	0.147	0.143	0.97 (0.29 - 3.24)	0.956	YES
SPARC_I2B	0.986	0.369	0.151	0.125	0.81 (0.38 - 1.7)	0.572	0.147	0.138	0.93 (0.28 - 3.1)	0.903	YES
SPARC_I3	0.964	0.203	0.231	0.233	1.01 (0.55 - 1.85)	0.974	0.235	0.232	0.98 (0.36 - 2.69)	0.973	YES
SPARC_P2	0.647	0.059	0.194	0.232	1.26 (0.59 - 2.7)	0.553	0.231	0.211	0.89 (0.27 - 2.95)	0.847	FAIL WATERS GENOTYPING RATE
SPARC_P3	0.978	0.130	0.382	0.360	0.91 (0.54 - 1.55)	0.736	0.353	0.357	1.02 (0.42 - 2.48)	0.968	YES
SPARC_P4	0.986	0.097	0.349	0.352	1.01 (0.6 - 1.72)	0.964	0.382	0.362	0.92 (0.38 - 2.2)	0.846	YES
SPARC_P5	0.993	0.214	0.266	0.284	1.1 (0.62 - 1.93)	0.752	0.265	0.276	1.06 (0.41 - 2.75)	0.908	YES
SPARC_P7	0.986	0.065	0.479	0.477	1.2 (0.72 - 1.99)	0.494	0.471	0.482	1.05 (0.45 - 2.46)	0.915	YES
TIMP1A	0.964	0.000	0.043	0.000	NA	0.052	0.059	0.000	NA	0.071	OUT OF HWE
TIMP1B	0.971	0.000	0.044	0.000	NA	0.046	0.063	0.000	NA	0.054	OUT OF HWE
TIMP2_I1	0.978	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I1C	0.986	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I2	0.971	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I3	0.986	0.180	0.226	0.307	1.52 (0.86 - 2.68)	0.149	0.382	0.345	0.85 (0.35 - 2.05)	0.717	YES
TIMP2A	0.324	0.928	0.016	0.038	2.52 (0.15 - 41.87)	0.505	0.000	0.071	NA	0.345	FAIL LOW GENOTYPING RATE
TIMP3A	0.935	0.604	0.378	0.325	0.79 (0.45 - 1.38)	0.414	0.313	0.327	1.07 (0.42 - 2.75)	0.891	YES
TIMP3B	0.914	0.693	0.354	0.316	0.84 (0.48 - 1.49)	0.557	0.233	0.327	1.6 (0.57 - 4.45)	0.370	YES
TIMP3C	0.000	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL NO PROBE
TIMP4_I1	0.993	1.000	0.000	0.033	NA	0.012	0.000	0.050	NA	NA	FAIL WATER GENOTYPING

SNP Identity	*Call Rate*	*HWE*	*MAF C*	*MAF CCLR*	*OR (95% CI)*	*P Value*	*MAF C S*	*MAF CCLR S*	*OR S (95% CI)*	*P Value S*	*Heterozygous*
TIMP4_U	0.964	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
TIMP4A	0.971	0.790	0.027	0.036	1.34 (0.31 - 5.75)	0.692	0.059	0.056	0.94 (0.15 - 5.94)	0.949	YES
TIMP4_D	0.964	0.830	0.022	0.000	1.14 (0.2 - 6.36)	0.880	0.000	0.000	NA	0.185	YES
TIMP4D	0.957	1.000	0.000	0.025	NA	1.000	0.029	0.037	1.27 (0.11 - 14.56)	0.848	YES
TNC_D	0.986	0.000	0.064	0.023	0.35 (0.08 - 1.6)	0.157	0.059	0.036	0.59 (0.08 - 4.41)	0.606	OUT OF HWE
TNC_E1	0.986	0.490	0.435	0.477	1.42 (0.85 - 2.36)	0.176	0.353	0.483	1.71 (0.72 - 4.09)	0.225	FAIL WATER GENOTYPING
TNC_E2	0.993	0.580	0.441	0.455	1.52 (0.91 - 2.53)	0.107	0.353	0.483	1.96 (0.82 - 4.7)	0.127	YES
TNC_E25	0.978	0.781	0.091	0.151	1.77 (0.82 - 3.83)	0.143	0.059	0.107	1.92 (0.36 - 10.11)	0.435	YES
TNC_I24	0.986	0.772	0.090	0.151	1.79 (0.83 - 3.88)	0.135	0.059	0.107	1.92 (0.36 - 10.11)	0.435	YES
TNC_I4	0.993	0.394	0.426	0.466	1.55 (0.93 - 2.57)	0.092	0.353	0.483	1.96 (0.82 - 4.7)	0.127	YES
TNC_P1	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
TNC_P2	0.964	0.000	0.413	0.321	0.67 (0.39 - 1.16)	0.153	0.412	0.352	0.78 (0.32 - 1.87)	0.572	FAIL WATER GENOTYPING
TNC_P4	0.986	0.248	0.255	0.337	1.48 (0.85 - 2.58)	0.162	0.206	0.321	1.83 (0.67 - 4.98)	0.235	YES
TNC_P5	0.993	0.430	0.436	0.455	1.55 (0.93 - 2.58)	0.090	0.353	0.483	1.96 (0.82 - 4.7)	0.127	YES
TNF10252	0.978	0.138	0.185	0.239	1.38 (0.75 - 2.56)	0.301	0.147	0.207	1.51 (0.48 - 4.74)	0.475	YES
TNF10411	0.396	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE
TNF10513	0.978	1.000	0.000	0.011	NA	NA	0.000	0.000	NA	NA	NO
TNF6547	0.957	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
TNF7178	0.986	0.418	0.441	0.466	1.11 (0.66 - 1.84)	0.697	0.324	0.466	1.82 (0.75 - 4.41)	0.182	YES
TNF8647	0.993	0.432	0.239	0.261	1.12 (0.63 - 2.01)	0.693	0.206	0.207	1.01 (0.35 - 2.86)	0.991	YES
TNF9367	1.000	0.676	0.197	0.233	1.24 (0.68 - 2.28)	0.483	0.147	0.200	1.45 (0.46 - 4.54)	0.522	YES
TNF9585	0.993	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TNFEXON1AB	0.942	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	NO
TNFEXON4AAB	0.986	0.917	0.011	0.000	NA	NA	0.000	0.000	NA	NA	NO
ZSWIM_E8A	0.986	0.380	0.399	0.453	1.25 (0.75 - 2.09)	0.395	0.471	0.393	1.74 (0.74 - 4.11)	0.206	YES
ZSWIM_E8B	0.986	0.917	0.011	0.011	1.06 (0.09 - 11.82)	0.964	0.029	0.017	0.58 (0.04 - 9.57)	0.699	YES
ZSWIM_I6	0.993	0.959	0.005	0.000	NA	NA	0.000	0.000	NA	NA	NO
ZSWIM_I6B	0.971	0.003	0.033	0.047	1.45 (0.4 - 5.27)	0.573	0.029	0.034	1.18 (0.1 - 13.5)	0.895	OUT OF HWE
ZSWIM_I6C	0.799	1.000	0.000	0.000	NA	NA	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE

Table 9

Comparisons of each SNP in the control and cranial cruciate ligament rupture (CCLR) Labrador Retriever populations. The SNP identity, genotyping rate (call rate), Hardy Weinberg equilibrium P value (HWE) (by Chi squared analysis), minor allele frequencies (MAF) for control (C) and cranial cruciate ligament disease (CCLD) populations, and the two populations stratified on the basis of neuter status (neutered only are denoted S) are listed. Case - control comparison by Chi squared analysis (P value), with Monte Carlo correction for significant associations (denoted CV) are also listed. SNPs are noted to be heterozygous if the MAF > 1% (YES), low (LOW) if MAF < 1%, and not (NO) if homozygous. SNPs were the genotyping test failed (FAIL), the reasons for failure (waters genotyping, low genotyping rate or no probe in the reaction are listed), and SNPs out of Hardy Weinberg Equilibrium are also indicated (OUT OF HWE).

SNP ID	CALL RATE	HWE	MAF C	MAF CCLR	P Value	OR (95% CI)	HWE S	MAF C S	MAF CCLR S	P Value S	OR S (95% CI)	Heterozygous
10_10S308	0.970	0.002	0.044	0.042	0.928	0.95 (0.33 - 2.77)	0.004	0.085	0.060	0.586	0.69 (0.18 - 2.66)	OUT OF HWE
10_11R124	0.895	0.724	0.215	0.297	0.109	1.54 (0.9 - 2.62)	0.751	0.255	0.350	0.253	1.58 (0.72 - 3.45)	YES
10_13Y85	0.970	0.855	0.249	0.266	0.715	1.1 (0.67 - 1.79)	0.580	0.311	0.340	0.720	1.14 (0.56 - 2.33)	YES
10_14R553	0.948	0.337	0.237	0.265	0.539	1.16 (0.72 - 1.89)	0.757	0.292	0.333	0.610	1.21 (0.58 - 2.51)	YES
10_1R105	0.972	0.807	0.245	0.270	0.585	1.14 (0.71 - 1.84)	0.580	0.311	0.340	0.720	1.14 (0.56 - 2.33)	YES
10_1R117	0.895	0.761	0.202	0.269	0.167	1.46 (0.85 - 2.49)	0.997	0.240	0.400	0.057	2.11 (0.97 - 4.58)	YES
10_1R218	0.958	0.426	0.202	0.214	0.775	1.08 (0.64 - 1.81)	0.889	0.245	0.280	0.643	1.2 (0.56 - 2.56)	YES
10_2R420	0.982	0.827	0.246	0.270	0.597	1.14 (0.71 - 1.83)	0.298	0.315	0.340	0.753	1.12 (0.55 - 2.29)	YES
10_3M171	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_4Y100	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_6R426	0.976	0.978	0.001	0.000	0.699	NA	1.000	0.000	0.000	NA	NA	LOW
10_6Y135	0.984	0.450	0.199	0.220	0.617	1.14 (0.68 - 1.9)	0.923	0.241	0.280	0.598	1.23 (0.57 - 2.62)	YES
10_9R210	0.970	0.978	0.001	0.000	0.710	NA	1.000	0.000	0.000	#N/A	NA	LOW
12B_01M115	0.974	0.736	0.162	0.266	0.013, CV 0.019	1.87 (1.13 - 3.09)	0.711	0.160	0.271	0.109	1.94 (0.86 - 4.42)	YES
12B_01Y90	0.944	0.777	0.163	0.244	0.059	1.66 (0.98 - 2.84)	0.624	0.167	0.273	0.136	1.88 (0.81 - 4.32)	YES
12B_02M407	0.968	0.379	0.315	0.300	0.769	0.93 (0.59 - 1.48)	0.793	0.259	0.340	0.296	1.47 (0.71 - 3.04)	YES
12B_02W232	0.970	0.405	0.147	0.120	0.480	0.79 (0.42 - 1.51)	0.066	0.111	0.140	0.604	1.3 (0.48 - 3.54)	YES
12B_02Y146	0.952	0.002	0.087	0.076	0.729	0.87 (0.38 - 1.96)	0.661	0.113	0.040	0.135	0.33 (0.07 - 1.52)	OUT OF HWE
12B_02Y190	0.970	0.603	0.144	0.120	0.521	0.81 (0.43 - 1.54)	0.066	0.111	0.140	0.604	1.3 (0.48 - 3.54)	YES
12B_03R196	0.964	0.489	0.323	0.294	0.558	0.87 (0.55 - 1.38)	0.714	0.287	0.327	0.606	1.21 (0.59 - 2.46)	YES
12B_03R462	0.972	0.913	0.167	0.180	0.749	1.09 (0.63 - 1.89)	0.201	0.148	0.200	0.414	1.44 (0.6 - 3.44)	YES
12B_03Y82	0.964	0.569	0.143	0.122	0.580	0.83 (0.44 - 1.59)	0.066	0.111	0.140	0.604	1.3 (0.48 - 3.54)	YES
12B_10R105	0.974	0.956	0.003	0.000	0.578	NA	1.000	0.000	0.000	#N/A	NA	LOW
12B_12Y142	0.972	0.956	0.003	0.000	0.582	NA	0.945	0.009	0.000	0.495	NA	LOW

SNP ID	CALL RATE	HWE	MAF C	MAF CCLR	P Value	OR (95% CI)	HWE S	MAF C S	MAF CCLR S	P Value S	OR S (95% CI)	Heterozygous
2_12Y206	0.976	0.934	0.004	0.000	0.502	NA	1.000	0.000	0.000	#N/A	NA	LOW
4_12M397	0.996	0.475	0.351	0.250	0.045, CV 0.05	0.62 (0.38 - 0.99)	0.979	0.306	0.200	0.166	0.57 (0.25 - 1.27)	YES
4_13S97	0.972	0.475	0.183	0.270	0.04, CV 0.035	1.65 (1.02 - 2.68)	0.405	0.194	0.340	0.046, CV 0.073	2.13 (1 - 4.54)	YES
4_1K110	0.966	0.562	0.354	0.245	0.036, CV 0.056	0.59 (0.36 - 0.97)	0.723	0.292	0.200	0.221	0.6 (0.27 - 1.36)	YES
4_22Y152	0.962	0.240	0.353	0.255	0.061	0.63 (0.38 - 1.03)	0.868	0.283	0.208	0.328	0.67 (0.3 - 1.51)	YES
4_25Y336	0.962	0.000	0.071	0.042	0.278	0.57 (0.2 - 1.61)	0.013	0.093	0.083	0.852	0.89 (0.26 - 3)	OUT OF HWE
4_2M351	0.954	0.989	0.346	0.250	0.062	0.63 (0.39 - 1.03)	0.681	0.298	0.200	0.197	0.59 (0.26 - 1.32)	YES
4_7S246	0.970	0.001	0.458	0.480	0.675	1.09 (0.72 - 1.67)	0.057	0.500	0.460	0.640	0.85 (0.44 - 1.67)	OUT OF HWE
4_8R458	0.972	0.529	0.185	0.270	0.046, CV 0.057	1.63 (1.01 - 2.64)	0.405	0.194	0.340	0.046, CV 0.073	2.13 (1 - 4.54)	YES
6_10Y257	0.976	0.846	0.011	0.000	0.303	NA	0.945	0.009	0.000	0.495	NA	YES
6_18R120	0.851	0.000	0.018	0.000	0.200	NA	1.000	0.000	0.000	#N/A	#N/A	FAIL WATER GENOTYPING
6_20R191	0.966	0.774	0.205	0.210	0.911	1.03 (0.61 - 1.73)	0.117	0.176	0.220	0.511	1.32 (0.57 - 3.04)	YES
6_20R240	0.962	0.400	0.156	0.160	0.921	1.03 (0.57 - 1.86)	0.230	0.142	0.180	0.534	1.33 (0.54 - 3.29)	YES
6_20R412	0.974	0.642	0.090	0.083	0.835	0.92 (0.43 - 1.99)	0.662	0.057	0.100	0.323	1.85 (0.54 - 6.39)	YES
6_6R431	0.970	0.176	0.105	0.117	0.726	1.13 (0.57 - 2.22)	0.353	0.113	0.125	0.833	1.12 (0.39 - 3.18)	YES
6_7R485	0.964	0.194	0.315	0.383	0.186	1.35 (0.86 - 2.11)	0.723	0.292	0.480	0.022, CV 0.052	2.23 (1.11 - 4.47)	YES
6_7S166	0.972	0.868	0.009	0.000	0.350	NA	1.000	0.000	0.000	#N/A	NA	LOW
6_8R289	0.974	0.912	0.006	0.000	0.452	NA	0.832	0.028	0.000	0.239	NA	FAIL WATER GENOTYPING
6_8W328	0.974	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
ANK_E5	0.964	0.589	0.029	0.041	0.524	1.43 (0.48 - 4.28)	0.834	0.028	0.040	0.683	1.46 (0.24 - 9.01)	YES
ANK_I1	0.966	0.909	0.342	0.300	0.408	0.82 (0.52 - 1.3)	0.540	0.333	0.280	0.503	0.78 (0.37 - 1.62)	YES
ANK_I2	0.968	0.124	0.354	0.394	0.453	1.18 (0.76 - 1.85)	0.651	0.380	0.326	0.527	0.79 (0.38 - 1.64)	YES
ANK_I3	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
ANK_I4	0.966	0.459	0.292	0.276	0.731	0.92 (0.57 - 1.48)	0.572	0.278	0.260	0.815	0.91 (0.43 - 1.95)	YES
ATP_I10	0.966	0.014	0.380	0.410	0.565	1.13 (0.74 - 1.74)	0.729	0.370	0.440	0.404	1.34 (0.68 - 2.64)	OUT OF HWE
ATP_I15	0.857	0.051	0.386	0.385	0.998	1 (0.64 - 1.56)	0.927	0.392	0.417	0.775	1.11 (0.55 - 2.23)	YES
ATP_I5	0.966	0.014	0.383	0.398	0.781	1.06 (0.69 - 1.64)	0.729	0.370	0.417	0.583	1.21 (0.61 - 2.43)	OUT OF HWE
ATP_I6	0.417	0.967	0.003	0.000	0.802	NA	0.880	0.042	0.000	0.557	NA	FAIL LOW GENOTYPING RATE
ATP_I7	0.956	0.011	0.380	0.410	0.571	1.13 (0.74 - 1.74)	0.729	0.370	0.440	0.404	1.34 (0.68 - 2.64)	OUT OF HWE
ATP_I8	0.962	0.001	0.395	0.439	0.407	1.2 (0.78 - 1.84)	0.294	0.389	0.480	0.280	1.45 (0.74 - 2.85)	OUT OF HWE
ATP_I8B	0.929	0.008	0.370	0.417	0.378	1.22 (0.79 - 1.88)	0.729	0.370	0.438	0.428	1.32 (0.66 - 2.64)	OUT OF HWE

SNP ID	*CALL RATE*	*HWE*	*MAF C*	*MAF CCLR*	*P Value*	*OR (95% CI)*	*HWE S*	*MAF C S*	*MAF CCLR S*	*P Value S*	*OR S (95% CI)*	*Heterozygous*
ATP_P1	0.805	0.018	0.126	0.207	0.048	1.81 (1 - 3.27)	0.166	0.174	0.229	0.442	1.41 (0.59 - 3.37)	FAIL LOW GENOTYPING RATE
IL1A10084	0.980	0.825	0.012	0.021	0.473	1.76 (0.37 - 8.42)	1.000	0.000	0.020	0.144	NA	YES
IL1A11235	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1A12227	0.974	0.805	0.200	0.200	0.994	1 (0.59 - 1.69)	0.298	0.204	0.220	0.815	1.1 (0.49 - 2.5)	YES
IL1AA	0.772	0.974	0.002	0.000	0.673	NA	1.000	0.000	0.000	#N/A	NA	LOW
IL1AB	0.843	0.903	0.007	0.011	0.739	1.45 (0.16 - 13.13)	1.000	0.000	0.021	0.184	NA	LOW
IL1AC	0.873	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AD	0.831	0.002	0.410	0.413	0.955	1.01 (0.65 - 1.59)	0.363	0.417	0.375	0.624	0.84 (0.42 - 1.69)	OUT OF HWE
IL1AE	0.927	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X221	0.952	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X255	0.966	0.002	0.421	0.440	0.720	1.08 (0.71 - 1.65)	0.363	0.417	0.380	0.662	0.86 (0.43 - 1.71)	OUT OF HWE
LEPRA	0.875	0.929	0.005	0.000	0.494	NA	1.000	0.000	0.000	#N/A	NA	LOW
LEPRB	0.843	0.029	0.264	0.356	0.071	1.54 (0.96 - 2.47)	0.044	0.279	0.326	0.573	1.25 (0.58 - 2.72)	OUT OF HWE
LEPRC	0.859	0.028	0.214	0.191	0.628	0.87 (0.5 - 1.52)	0.647	0.233	0.208	0.738	0.86 (0.37 - 2.02)	OUT OF HWE
MMP13A	0.869	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP13B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP13D	0.847	0.491	0.044	0.064	0.391	1.49 (0.59 - 3.76)	0.817	0.033	0.083	0.202	2.64 (0.57 - 12.3)	YES
MMP13E	0.897	0.977	0.002	0.000	0.696	NA	1.000	0.000	0.000	#N/A	NA	LOW
MMP13F	0.873	0.976	0.002	0.000	0.688	NA	1.000	0.000	0.000	#N/A	NA	LOW
MMP13G	0.863	0.175	0.032	0.061	0.161	1.95 (0.75 - 5.04)	0.817	0.033	0.083	0.202	2.64 (0.57 - 12.3)	YES
MMP13H	0.887	0.977	0.002	0.000	0.684	NA	1.000	0.000	0.000	#N/A	NA	LOW
MMP13J	0.875	0.977	0.002	0.000	0.688	NA	1.000	0.000	0.000	#N/A	NA	LOW
MMP13K	0.895	0.000	0.036	0.020	0.436	0.56 (0.13 - 2.44)	0.000	0.044	0.042	0.939	0.93 (0.16 - 5.3)	OUT OF HWE
MMP2A	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP3A	0.885	0.476	0.081	0.052	0.324	0.62 (0.24 - 1.61)	0.693	0.056	0.083	0.529	1.55 (0.39 - 6.05)	YES
MMP3B	0.996	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP9_D	0.962	0.516	0.200	0.260	0.171	1.4 (0.86 - 2.28)	0.239	0.213	0.220	0.920	1.04 (0.46 - 2.35)	YES
MMP9_I8	0.960	0.000	0.271	0.367	0.049	1.56 (1 - 2.43)	0.000	0.333	0.333	1.000	1 (0.49 - 2.06)	OUT OF HWE
MMP9_P1	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9_P4	0.829	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
MMP9_U1	0.948	0.038	0.158	0.234	0.065	1.63 (0.97 - 2.75)	0.142	0.167	0.239	0.292	1.57 (0.67 - 3.66)	OUT OF HWE
MMP9_U2	0.972	0.527	0.289	0.260	0.544	0.86 (0.54 - 1.39)	0.979	0.306	0.240	0.396	0.72 (0.33 - 1.55)	YES
MMP9A	0.440	1.000	0.000	0.031	0.003	NA	1.000	0.000	0.071	0.057	NA	FAIL LOW GENOTYPING RATE
MMP9C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9D	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
SPARC_I2	0.970	0.799	0.062	0.080	0.492	1.32 (0.6 - 2.9)	0.721	0.046	0.080	0.395	1.79 (0.46 - 6.98)	YES
SPARC_I2B	0.958	0.367	0.048	0.082	0.159	1.77 (0.79 - 3.98)	0.834	0.028	0.080	0.138	3.04 (0.65 - 14.15)	YES

SNP ID	CALL RATE	HWE	MAF C	MAF CCLR	P Value	OR (95% CI)	HWE S	MAF C S	MAF CCLR S	P Value S	OR S (95% CI)	Heterozygous
SPARC_I3	0.887	0.003	0.047	0.000	0.028	NA	0.890	0.019	0.000	0.333	NA	OUT OF HWE
SPARC_P2	0.599	0.000	0.307	0.370	0.386	1.32 (0.7 - 2.49)	0.034	0.286	0.375	0.430	1.5 (0.55 - 4.12)	FAIL LOW GENOTYPING RATE
SPARC_P3	0.919	0.007	0.341	0.375	0.519	1.16 (0.74 - 1.81)	0.539	0.361	0.354	0.933	0.97 (0.48 - 1.97)	OUT OF HWE
SPARC_P4	0.962	0.020	0.398	0.375	0.669	0.91 (0.58 - 1.41)	0.218	0.435	0.438	0.979	1.01 (0.51 - 2)	OUT OF HWE
SPARC_P5	0.968	0.026	0.260	0.250	0.834	0.95 (0.58 - 1.54)	0.840	0.204	0.200	0.957	0.98 (0.42 - 2.26)	OUT OF HWE
SPARC_P7	0.966	0.033	0.346	0.370	0.640	1.11 (0.72 - 1.72)	0.539	0.361	0.360	0.989	1 (0.5 - 2)	OUT OF HWE
TIMP1A	0.885	0.000	0.033	0.031	0.943	0.96 (0.28 - 3.29)	0.817	0.033	0.063	0.424	1.93 (0.37 - 9.97)	OUT OF HWE
TIMP1B	0.970	0.000	0.020	0.031	0.474	1.59 (0.44 - 5.67)	0.834	0.028	0.060	0.324	2.23 (0.43 - 11.48)	OUT OF HWE
TIMP2_I1	0.960	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I1C	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I2	0.944	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I3	0.950	0.622	0.420	0.479	0.270	1.27 (0.83 - 1.96)	0.251	0.407	0.438	0.725	1.13 (0.57 - 2.25)	YES
TIMP2A	0.407	0.583	0.046	0.083	0.334	1.89 (0.51 - 7.06)	0.333	0.200	0.063	0.216	0.27 (0.03 - 2.44)	FAIL LOW GENOTYPING RATE
TIMP3A	0.869	0.430	0.366	0.362	0.933	0.98 (0.62 - 1.54)	0.823	0.344	0.396	0.550	1.25 (0.6 - 2.57)	YES
TIMP3B	0.879	0.380	0.376	0.330	0.390	0.82 (0.52 - 1.3)	0.651	0.380	0.354	0.761	0.9 (0.44 - 1.82)	YES
TIMP3C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
TIMP4_I1	0.980	0.000	0.025	0.020	0.750	0.79 (0.18 - 3.46)	1.000	0.000	0.038	NA	NA	FAIL WATER GENOTYPING
TIMP4_U	0.950	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TIMP4A	0.859	0.019	0.096	0.067	0.370	0.67 (0.28 - 1.61)	0.632	0.067	0.063	0.925	0.93 (0.22 - 3.91)	OUT OF HWE
TIMP4D	0.877	0.448	0.084	0.082	0.940	0.97 (0.44 - 2.12)	0.689	0.057	0.063	0.893	1.11 (0.25 - 4.85)	YES
TIMP4_D	0.877	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNC_D	0.960	0.015	0.037	0.020	0.392	0.53 (0.12 - 2.3)	0.006	0.046	0.020	0.421	0.42 (0.05 - 3.7)	OUT OF HWE
TNC_E1	0.982	0.000	0.127	0.088	0.262	0.66 (0.32 - 1.37)	0.405	0.102	0.096	0.910	0.94 (0.31 - 2.86)	FAIL WATER GENOTYPING
TNC_E2	0.968	0.117	0.102	0.060	0.189	0.56 (0.24 - 1.34)	0.405	0.102	0.060	0.389	0.56 (0.15 - 2.11)	YES
TNC_E25	0.962	0.584	0.137	0.090	0.193	0.62 (0.3 - 1.28)	0.201	0.148	0.080	0.231	0.5 (0.16 - 1.58)	YES
TNC_I24	0.960	0.656	0.135	0.090	0.212	0.63 (0.31 - 1.3)	0.201	0.148	0.080	0.231	0.5 (0.16 - 1.58)	YES
TNC_I4	0.968	0.039	0.333	0.327	0.894	0.97 (0.62 - 1.52)	0.572	0.278	0.380	0.196	1.59 (0.78 - 3.24)	OUT OF HWE
TNC_P1	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNC_P2	0.925	0.292	0.056	0.052	0.862	0.92 (0.35 - 2.41)	0.543	0.078	0.042	0.400	0.51 (0.1 - 2.5)	FAIL WATER GENOTYPING
TNC_P4	0.962	0.092	0.144	0.150	0.868	1.05 (0.58 - 1.9)	0.274	0.130	0.120	0.866	0.92 (0.33 - 2.54)	YES
TNC_P5	0.964	0.102	0.101	0.070	0.334	0.67 (0.3 - 1.51)	0.405	0.102	0.080	0.663	0.77 (0.23 - 2.54)	YES
TNF10252	0.960	0.120	0.047	0.031	0.481	0.65 (0.19 - 2.17)	0.777	0.037	0.060	0.514	1.66 (0.36 - 7.71)	YES
TNF10411	0.814	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF10513	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF6547	0.980	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO

SNP ID	*CALL RATE*	*HWE*	*MAF C*	*MAF CCLR*	*P Value*	*OR (95% CI)*	*HWE S*	*MAF C S*	*MAF CCLR S*	*P Value S*	*OR S (95% CI)*	*Heterozygous*
TNF7178	0.970	0.426	0.038	0.030	0.697	0.79 (0.23 - 2.65)	0.777	0.037	0.040	0.928	1.08 (0.19 - 6.12)	YES
TNF8647	0.976	0.004	0.230	0.200	0.507	0.84 (0.5 - 1.41)	0.050	0.148	0.280	0.049	2.24 (0.99 - 5.05)	OUT OF HWE
TNF9367	0.978	0.004	0.058	0.040	0.456	0.67 (0.23 - 1.92)	0.721	0.046	0.060	0.715	1.31 (0.3 - 5.73)	OUT OF HWE
TNF9585	0.994	0.000	0.024	0.020	0.823	0.84 (0.19 - 3.73)	1.000	0.000	0.038	0.040	NA	FAIL WATER GENOTYPING
TNFEXON1AB	0.972	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNFEXON4AAB	0.972	0.978	0.002	0.000	0.697	NA	1.000	0.000	0.000	#N/A	NA	LOW
ZSWIM_E8A	0.958	0.007	0.482	0.430	0.335	0.81 (0.53 - 1.24)	0.713	0.444	0.440	0.176	1.59 (0.81 - 3.13)	OUT OF HWE
ZSWIM_E8B	0.964	0.022	0.259	0.367	0.025	1.66 (1.06 - 2.59)	0.095	0.185	0.340	0.033, CV 0.046	2.27 (1.06 - 4.85)	OUT OF HWE
ZSWIM_I6	0.962	0.454	0.040	0.050	0.631	1.27 (0.48 - 3.39)	0.666	0.056	0.060	0.911	1.09 (0.26 - 4.53)	YES
ZSWIM_I6B	0.913	0.008	0.097	0.043	0.092	0.42 (0.15 - 1.19)	0.304	0.090	0.065	0.613	0.71 (0.18 - 2.74)	OUT OF HWE
ZSWIM_I6C	0.343	0.942	0.010	0.000	0.379	NA	1.000	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE

Table 10

Comparisons of each SNP in the control and hip dysplasia Labrador Retriever populations. The SNP identity, genotyping rate (call rate), Hardy Weinberg equilibrium P value (HWE) (by Chi squared analysis), minor allele frequencies (MAF) for control (C) and hip dysplasia (HD) populations, and the two populations stratified on the basis of sex status (males only are denoted S) are listed. Case - control comparison by Chi squared analysis (P value), with Monte Carlo correction for significant associations (denoted CV) are also listed. SNPs are noted to be heterozygous if the MAF > 1% (YES), low (LOW) if MAF < 1%, and not (NO) if homozygous. SNPs were the genotyping test failed (FAIL), the reasons for failure (waters genotyping, low genotyping rate or no probe in the reaction are listed), and SNPs out of Hardy Weinberg Equilibrium are also indicated (OUT OF HWE).

SNP ID	*Call Rate*	*HWE*	*MAF C*	*MAF HD*	*P Value*	*OR (95% CI))*	*HWE S*	*MAF C S*	*MAF HD S*	*P Value S*	*OR S (95% CI)*	*Heterozygous*
10_10S308	0.970	0.002	0.044	0.016	0.281	0.35 (0.05 - 2.59)	0.209	0.050	0.029	0.602	0.58 (0.07 - 4.58)	OUT OF HWE
10_11R124	0.895	0.724	0.215	0.141	0.160	0.6 (0.29 - 1.24)	0.504	0.213	0.176	0.626	0.79 (0.31 - 2.01)	YES
10_13Y85	0.970	0.855	0.249	0.156	0.099	0.56 (0.28 - 1.12)	0.417	0.250	0.206	0.573	0.78 (0.32 - 1.87)	YES
10_14R553	0.948	0.337	0.237	0.156	0.144	0.6 (0.3 - 1.2)	0.988	0.234	0.206	0.711	0.85 (0.35 - 2.04)	YES
10_1R105	0.972	0.807	0.245	0.156	0.112	0.57 (0.28 - 1.15)	0.417	0.250	0.206	0.573	0.78 (0.32 - 1.87)	YES
10_1R117	0.895	0.761	0.202	0.133	0.201	0.61 (0.28 - 1.31)	0.422	0.192	0.188	0.951	0.97 (0.38 - 2.49)	YES
10_1R218	0.958	0.426	0.202	0.141	0.239	0.65 (0.31 - 1.34)	0.965	0.199	0.176	0.754	0.86 (0.34 - 2.19)	YES
10_2R420	0.982	0.827	0.246	0.156	0.109	0.57 (0.28 - 1.14)	0.238	0.252	0.206	0.558	0.77 (0.32 - 1.85)	YES
10_3M171	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_4Y100	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_6R426	0.976	0.978	0.001	0.000	0.757	NA	0.965	0.004	0.000	0.720	NA	LOW
10_6Y135	0.984	0.450	0.199	0.141	0.263	0.66 (0.32 - 1.37)	0.964	0.195	0.176	0.791	0.88 (0.35 - 2.24)	YES
10_9R210	0.970	0.978	0.001	0.000	0.757	NA	0.965	0.004	0.000	0.719	NA	LOW
12B_01M115	0.974	0.736	0.162	0.141	0.653	0.85 (0.41 - 1.76)	0.919	0.170	0.206	0.608	1.26 (0.52 - 3.08)	YES
12B_01Y90	0.944	0.777	0.163	0.141	0.647	0.84 (0.4 - 1.76)	0.863	0.169	0.206	0.596	1.27 (0.52 - 3.11)	YES
12B_02M407	0.968	0.379	0.315	0.297	0.770	0.92 (0.52 - 1.61)	0.529	0.323	0.353	0.727	1.14 (0.54 - 2.42)	YES
12B_02W232	0.970	0.405	0.147	0.065	0.075	0.4 (0.14 - 1.13)	0.449	0.149	0.094	0.400	0.59 (0.17 - 2.04)	YES
12B_02Y146	0.952	0.002	0.087	0.177	0.020	2.27 (1.12 - 4.59)	0.992	0.088	0.063	0.628	0.69 (0.16 - 3.09)	OUT OF HWE
12B_02Y190	0.970	0.603	0.144	0.078	0.145	0.5 (0.2 - 1.29)	0.863	0.145	0.118	0.667	0.79 (0.26 - 2.36)	YES
12B_03R196	0.964	0.489	0.323	0.313	0.862	0.95 (0.55 - 1.66)	0.511	0.337	0.353	0.855	1.07 (0.51 - 2.27)	YES
12B_03R462	0.972	0.913	0.167	0.219	0.296	1.39 (0.75 - 2.61)	0.225	0.174	0.235	0.384	1.46 (0.62 - 3.42)	YES
12B_03Y82	0.964	0.569	0.143	0.078	0.148	0.51 (0.2 - 1.3)	0.804	0.143	0.118	0.684	0.8 (0.27 - 2.39)	YES
12B_10R105	0.974	0.956	0.003	0.000	0.659	NA	0.930	0.008	0.000	0.609	NA	LOW
12B_12Y142	0.972	0.956	0.003	0.016	0.133	5.24 (0.47 - 58.58)	1.000	0.000	0.000	NA	NA	LOW
2_12Y206	0.976	0.934	0.004	0.000	0.591	NA	1.000	0.000	0.000	NA	NA	LOW

SNP ID	Call Rate	HWE	MAF C	MAF HD	P Value	OR (95% CI))	HWE S	MAF C S	MAF HD S	P Value S	OR S (95% CI)	Heterozygous
4_12M397	0.996	0.475	0.351	0.313	0.532	0.84 (0.48 - 1.46)	0.023	0.327	0.294	0.699	0.86 (0.39 - 1.87)	YES
4_13S97	0.972	0.475	0.183	0.328	0.005, CV 0.008	2.18 (1.25 - 3.81)	0.228	0.174	0.412	0.001, CV 0.004	3.32 (1.56 - 7.05)	YES
4_1K110	0.966	0.562	0.354	0.306	0.454	0.81 (0.46 - 1.42)	0.029	0.323	0.281	0.632	0.82 (0.36 - 1.85)	YES
4_22Y152	0.962	0.240	0.353	0.313	0.516	0.83 (0.48 - 1.45)	0.018	0.326	0.294	0.710	0.86 (0.39 - 1.88)	YES
4_25Y336	0.962	0.000	0.071	0.016	0.087	0.21 (0.03 - 1.52)	0.000	0.058	0.029	0.494	0.49 (0.06 - 3.87)	OUT OF HWE
4_2M351	0.954	0.989	0.346	0.333	0.842	0.94 (0.54 - 1.65)	0.029	0.323	0.313	0.904	0.95 (0.43 - 2.1)	YES
4_7S246	0.970	0.001	0.458	0.359	0.132	0.66 (0.39 - 1.13)	0.000	0.496	0.294	0.026	0.42 (0.19 - 0.92)	OUT OF HWE
4_8R458	0.972	0.529	0.185	0.328	0.006, CV 0.008	2.15 (1.23 - 3.76)	0.190	0.172	0.412	0.001, CV 0.005	3.38 (1.59 - 7.18)	YES
6_10Y257	0.976	0.846	0.011	0.016	0.707	1.49 (0.18 - 12.34)	0.895	0.011	0.029	0.389	2.64 (0.27 - 26.09)	YES
6_18R120	0.851	0.000	0.018	0.000	0.287	NA	0.000	0.023	0.000	0.371	NA	FAIL WATER GENOTYPING
6_20R191	0.966	0.774	0.205	0.094	0.032, CV 0.05	0.4 (0.17 - 0.95)	0.228	0.200	0.088	0.116	0.39 (0.11 - 1.32)	YES
6_20R240	0.962	0.400	0.156	0.094	0.186	0.56 (0.24 - 1.34)	0.890	0.156	0.088	0.293	0.52 (0.15 - 1.79)	YES
6_20R412	0.974	0.642	0.090	0.063	0.460	0.68 (0.24 - 1.92)	0.924	0.083	0.059	0.621	0.69 (0.15 - 3.06)	YES
6_6R431	0.970	0.176	0.105	0.031	0.058	0.27 (0.07 - 1.15)	0.656	0.106	0.029	0.156	0.26 (0.03 - 1.94)	YES
6_7R485	0.964	0.194	0.315	0.406	0.135	1.49 (0.88 - 2.52)	0.617	0.305	0.324	0.829	1.09 (0.51 - 2.34)	YES
6_7S166	0.972	0.868	0.009	0.000	0.445	NA	0.895	0.011	0.000	0.531	NA	LOW
6_8R289	0.974	0.912	0.006	0.000	0.535	NA	0.930	0.008	0.000	0.611	NA	FAIL WATER GENOTYPING
6_8W328	0.974	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
ANK_E5	0.964	0.589	0.029	0.032	0.883	1.12 (0.25 - 4.91)	0.717	0.031	0.000	0.300	NA	YES
ANK_I1	0.966	0.909	0.342	0.344	0.977	1.01 (0.59 - 1.73)	0.966	0.327	0.382	0.519	1.27 (0.61 - 2.67)	YES
ANK_I2	0.968	0.124	0.354	0.387	0.602	1.15 (0.68 - 1.97)	0.249	0.317	0.353	0.671	1.18 (0.56 - 2.49)	YES
ANK_I3	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
ANK_I4	0.966	0.459	0.292	0.290	0.972	0.99 (0.56 - 1.76)	0.914	0.281	0.294	0.871	1.07 (0.49 - 2.34)	YES
ATP_I10	0.966	0.014	0.380	0.484	0.102	1.53 (0.92 - 2.57)	0.972	0.373	0.471	0.079	1.89 (0.92 - 3.88)	OUT OF HWE
ATP_I15	0.857	0.051	0.386	0.483	0.051	1.71 (0.99 - 2.94)	0.933	0.387	0.469	0.117	1.8 (0.86 - 3.78)	YES
ATP_I5	0.966	0.014	0.383	0.468	0.193	1.41 (0.84 - 2.38)	0.972	0.373	0.471	0.079	1.89 (0.92 - 3.88)	OUT OF HWE
ATP_I6	0.417	0.967	0.003	0.000	0.906	NA	0.934	0.013	0.000	0.817	NA	FAIL LOW GENOTYPING RATE
ATP_I7	0.956	0.011	0.380	0.500	0.065	1.63 (0.97 - 2.75)	0.972	0.373	0.438	0.039, CV 0.066	2.16 (1.03 - 4.54)	OUT OF HWE
ATP_I8	0.962	0.001	0.395	0.468	0.263	1.35 (0.8 - 2.27)	0.464	0.392	0.471	0.126	1.74 (0.85 - 3.57)	OUT OF HWE
ATP_I8B	0.929	0.008	0.370	0.500	0.044	1.7 (1.01 - 2.88)	1.000	0.375	0.438	0.041, CV 0.061	2.14 (1.02 - 4.5)	OUT OF HWE

SNP ID	Call Rate	HWE	MAF C	MAF HD	P Value	OR (95% CI))	HWE S	MAF C S	MAF HD S	P Value S	OR S (95% CI)	Heterozygous
ATP_P1	0.805	0.018	0.126	0.167	0.427	1.38 (0.62 - 3.08)	0.052	0.157	0.115	0.574	0.7 (0.2 - 2.46)	FAIL LOW GENOTYPING RATE
IL1A10084	0.980	0.825	0.012	0.000	0.379	NA	0.860	0.015	0.000	0.470	NA	YES
IL1A11235	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1A12227	0.974	0.805	0.200	0.219	0.717	1.12 (0.6 - 2.09)	0.091	0.218	0.294	0.318	1.49 (0.68 - 3.3)	YES
IL1AA	0.772	0.974	0.002	0.000	0.768	NA	1.000	0.000	0.000	NA	NA	LOW
IL1AB	0.843	0.903	0.007	0.000	0.551	NA	0.961	0.005	0.000	0.728	NA	LOW
IL1AC	0.873	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AD	0.831	0.002	0.410	0.400	0.882	0.96 (0.56 - 1.65)	0.112	0.453	0.344	0.238	0.63 (0.29 - 1.36)	OUT OF HWE
IL1AE	0.927	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X221	0.952	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X255	0.966	0.002	0.421	0.406	0.820	0.94 (0.56 - 1.59)	0.075	0.462	0.324	0.127	0.56 (0.26 - 1.19)	OUT OF HWE
LEPRA	0.875	0.929	0.005	0.000	0.596	NA	1.000	0.000	0.000	NA	NA	LOW
LEPRB	0.843	0.029	0.264	0.220	0.500	0.79 (0.39 - 1.58)	0.364	0.292	0.269	0.805	0.89 (0.36 - 2.23)	OUT OF HWE
LEPRC	0.859	0.028	0.214	0.167	0.419	0.74 (0.35 - 1.55)	0.115	0.198	0.179	0.805	0.88 (0.32 - 2.44)	OUT OF HWE
MMP13A	0.869	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP13B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP13D	0.847	0.491	0.044	0.058	0.640	1.34 (0.39 - 4.62)	0.725	0.033	0.077	0.268	2.44 (0.48 - 12.42)	YES
MMP13E	0.897	0.977	0.002	0.000	0.752	NA	0.965	0.004	0.000	0.726	NA	LOW
MMP13F	0.873	0.976	0.002	0.000	0.758	NA	1.000	0.000	0.000	NA	NA	LOW
MMP13G	0.863	0.175	0.032	0.040	0.772	1.25 (0.28 - 5.53)	0.765	0.028	0.071	0.227	2.67 (0.51 - 13.91)	YES
MMP13H	0.887	0.977	0.002	0.000	0.760	NA	1.000	0.000	0.000	NA	NA	LOW
MMP13J	0.875	0.977	0.002	0.000	0.744	NA	0.965	0.004	0.000	0.726	NA	LOW
MMP13K	0.895	0.000	0.036	0.000	0.158	NA	0.000	0.040	0.000	0.280	NA	OUT OF HWE
MMP2A	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP3A	0.885	0.476	0.081	0.111	0.445	1.42 (0.58 - 3.49)	0.355	0.080	0.000	0.120	NA	YES
MMP3B	0.996	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP9_D	0.962	0.516	0.200	0.234	0.518	1.22 (0.66 - 2.25)	0.913	0.200	0.265	0.382	1.44 (0.63 - 3.27)	YES
MMP9_I8	0.960	0.000	0.271	0.375	0.078	1.61 (0.94 - 2.75)	0.000	0.269	0.412	0.084	1.9 (0.91 - 3.97)	OUT OF HWE
MMP9_P1	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9_P4	0.829	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
MMP9_U1	0.948	0.038	0.158	0.226	0.168	1.56 (0.83 - 2.93)	0.159	0.151	0.235	0.210	1.73 (0.73 - 4.09)	OUT OF HWE
MMP9_U2	0.972	0.527	0.289	0.290	0.988	1 (0.57 - 1.78)	0.819	0.323	0.294	0.733	0.87 (0.4 - 1.91)	YES
MMP9A	0.440	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE
MMP9C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9D	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
SPARC_I2	0.970	0.799	0.062	0.063	0.986	1.01 (0.35 - 2.92)	0.455	0.062	0.059	0.950	0.95 (0.21 - 4.34)	YES
SPARC_I2B	0.958	0.367	0.048	0.048	0.980	1.02 (0.3 - 3.42)	0.581	0.046	0.029	0.655	0.63 (0.08 - 4.97)	YES
SPARC_I3	0.887	0.003	0.047	0.097	0.097	2.15 (0.85 - 5.42)	0.446	0.062	0.125	0.185	2.16 (0.68 - 6.92)	OUT OF HWE
SPARC_P2	0.599	0.000	0.307	0.278	0.713	0.87 (0.41 - 1.85)	0.000	0.333	0.313	0.866	0.91 (0.3 - 2.74)	FAIL LOW GENOTYPING RATE

SNP ID	Call Rate	HWE	MAF C	MAF HD	P Value	OR (95% CI))	HWE S	MAF C S	MAF HD S	P Value S	OR S (95% CI)	Heterozygous
SPARC_P3	0.919	0.007	0.341	0.403	0.329	1.3 (0.76 - 2.22)	0.659	0.327	0.281	0.602	0.81 (0.36 - 1.82)	OUT OF HWE
SPARC_P4	0.962	0.020	0.398	0.323	0.246	0.72 (0.41 - 1.26)	0.837	0.415	0.441	0.774	1.11 (0.54 - 2.28)	OUT OF HWE
SPARC_P5	0.968	0.026	0.260	0.290	0.602	1.17 (0.66 - 2.07)	0.944	0.265	0.265	0.993	1 (0.44 - 2.24)	OUT OF HWE
SPARC_P7	0.966	0.033	0.346	0.406	0.335	1.29 (0.77 - 2.18)	0.659	0.327	0.294	0.701	0.86 (0.39 - 1.87)	OUT OF HWE
TIMP1A	0.885	0.000	0.033	0.038	0.822	1.19 (0.27 - 5.24)	0.000	0.027	0.077	0.173	3 (0.57 - 15.7)	OUT OF HWE
TIMP1B	0.970	0.000	0.020	0.031	0.527	1.62 (0.36 - 7.34)	0.000	0.008	0.059	0.015	8.19 (1.11 - 60.14)	OUT OF HWE
TIMP2_I1	0.960	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I1C	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I2	0.944	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I3	0.950	0.622	0.420	0.468	0.087	1.57 (0.93 - 2.66)	0.230	0.446	0.441	0.960	0.98 (0.48 - 2.02)	YES
TIMP2A	0.407	0.583	0.046	0.029	0.661	0.63 (0.08 - 5.01)	0.780	0.036	0.071	0.530	2.04 (0.21 - 19.65)	FAIL LOW GENOTYPING RATE
TIMP3A	0.869	0.430	0.366	0.259	0.117	0.61 (0.32 - 1.14)	0.761	0.356	0.357	0.995	1 (0.44 - 2.28)	YES
TIMP3B	0.879	0.380	0.376	0.323	0.409	0.79 (0.45 - 1.38)	0.620	0.374	0.375	0.992	1 (0.47 - 2.14)	YES
TIMP3C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
TIMP4_U	0.980	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TIMP4_D	0.950	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TIMP4A	0.859	0.019	0.096	0.096	0.999	1 (0.38 - 2.62)	0.289	0.093	0.143	0.401	1.63 (0.52 - 5.18)	OUT OF HWE
TIMP4D	0.877	0.448	0.084	0.096	0.762	1.16 (0.44 - 3.06)	0.316	0.088	0.115	0.646	1.35 (0.37 - 4.92)	YES
TIMP4_I1	0.877	0.000	0.025	0.000	NA	NA	0.000	0.015	0.000	NA	NA	FAIL WATER GENOTYPING
TNC_D	0.960	0.015	0.037	0.031	0.821	0.84 (0.19 - 3.66)	0.683	0.035	0.059	0.484	1.74 (0.36 - 8.43)	OUT OF HWE
TNC_E1	0.982	0.000	0.127	0.065	0.149	0.47 (0.17 - 1.34)	0.004	0.095	0.118	0.682	1.26 (0.41 - 3.88)	FAIL WATER GENOTYPING
TNC_E2	0.968	0.117	0.102	0.063	0.316	0.59 (0.21 - 1.67)	0.048	0.092	0.118	0.636	1.31 (0.43 - 4.04)	YES
TNC_E25	0.962	0.584	0.137	0.156	0.674	1.16 (0.57 - 2.37)	0.336	0.119	0.176	0.344	1.58 (0.61 - 4.13)	YES
TNC_I24	0.960	0.656	0.135	0.156	0.636	1.19 (0.58 - 2.42)	0.336	0.119	0.176	0.344	1.58 (0.61 - 4.13)	YES
TNC_I4	0.968	0.039	0.333	0.290	0.491	0.82 (0.46 - 1.45)	0.003	0.312	0.412	0.240	1.55 (0.74 - 3.22)	OUT OF HWE
TNC_P1	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNC_P2	0.925	0.292	0.056	0.065	0.794	1.15 (0.4 - 3.36)	0.447	0.063	0.118	0.245	1.97 (0.62 - 6.27)	FAIL WATER GENOTYPING
TNC_P4	0.962	0.092	0.144	0.188	0.346	1.37 (0.71 - 2.67)	0.345	0.131	0.118	0.830	0.89 (0.29 - 2.67)	YES
TNC_P5	0.964	0.102	0.101	0.063	0.326	0.6 (0.21 - 1.69)	0.048	0.092	0.118	0.636	1.31 (0.43 - 4.04)	YES
TNF10252	0.960	0.120	0.047	0.031	0.559	0.65 (0.15 - 2.78)	0.360	0.058	0.029	0.494	0.49 (0.06 - 3.87)	YES
TNF10411	0.814	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF10513	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF6547	0.980	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF7178	0.970	0.426	0.038	0.016	0.361	0.4 (0.05 - 3.03)	0.718	0.031	0.000	0.302	NA	YES
TNF8647	0.976	0.004	0.230	0.234	0.933	1.03 (0.56 - 1.88)	0.007	0.223	0.235	0.877	1.07 (0.46 - 2.49)	OUT OF HWE

SNP ID	*Call Rate*	*HWE*	*MAF C*	*MAF HD*	*P Value*	*OR (95% CI))*	*HWE S*	*MAF C S*	*MAF HD S*	*P Value S*	*OR S (95% CI)*	*Heterozygous*
TNF9367	0.978	0.004	0.058	0.031	0.367	0.52 (0.12 - 2.21)	0.426	0.060	0.029	0.465	0.47 (0.06 - 3.69)	OUT OF HWE
TNF9585	0.994	0.000	0.024	0.000	0.214	NA	0.930	0.008	0.000	0.612	NA	FAIL WATER GENOTYPING
TNFEXON1AB	0.972	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNFEXON4AAB	0.972	0.978	0.002	0.000	0.756	NA	1.000	0.000	0.000	NA	NA	LOW
ZSWIM_E8A	0.958	0.007	0.482	0.484	0.973	1.01 (0.6 - 1.7)	0.055	0.485	0.469	0.865	0.94 (0.45 - 1.96)	OUT OF HWE
ZSWIM_E8B	0.964	0.022	0.259	0.226	0.565	0.83 (0.45 - 1.55)	0.084	0.265	0.118	0.061	0.37 (0.13 - 1.09)	OUT OF HWE
ZSWIM_I6	0.962	0.454	0.040	0.031	0.737	0.78 (0.18 - 3.36)	0.581	0.046	0.059	0.744	1.29 (0.28 - 6.03)	YES
ZSWIM_I6B	0.913	0.008	0.097	0.100	0.948	1.03 (0.43 - 2.49)	0.148	0.081	0.147	0.201	1.97 (0.69 - 5.64)	OUT OF HWE
ZSWIM_I6C	0.343	0.942	0.010	0.000	0.701	NA	0.919	0.020	0.000	0.652	NA	FAIL LOW GENOTYPING RATE

Table 11

Comparisons of each SNP in the control and elbow dysplasia Labrador Retriever populations. The SNP identity, genotyping rate (call rate), Hardy Weinberg equilibrium P value (HWE) (by Chi squared analysis), minor allele frequencies (MAF) for control (C) and elbow dysplasia (ED) populations, and the two populations stratified on the basis of sex status (males only are denoted S) are listed. Case - control comparison by Chi squared analysis (P value), with Monte Carlo correction for significant associations (denoted CV) are also listed. SNPs are noted to be heterozygous if the MAF > 1% (YES), low (LOW) if MAF < 1%, and not (NO) if homozygous. SNPs were the genotyping test failed (FAIL), the reasons for failure (waters genotyping, low genotyping rate or no probe in the reaction are listed), and SNPs out of Hardy Weinberg Equilibrium are also indicated (OUT OF HWE).

SNP ID	Call Rate	HWE	MAF C	MAF ED	P Value	OR (95% CI)	HWE S	MAF C S	MAF ED S	P Value S	OR S (95% CI)	Heterozygous
10_10S308	0.970	0.002	0.044	0.058	0.454	1.34 (0.62 - 2.89)	0.209	0.050	0.042	0.758	0.85 (0.3 - 2.43)	OUT OF HWE
10_11R124	0.895	0.724	0.215	0.233	0.676	1.11 (0.69 - 1.77)	0.504	0.213	0.200	0.801	0.93 (0.51 - 1.68)	YES
10_13Y85	0.970	0.855	0.249	0.227	0.581	0.89 (0.59 - 1.35)	0.417	0.250	0.190	0.199	0.7 (0.41 - 1.21)	YES
10_14R553	0.948	0.337	0.237	0.207	0.430	0.84 (0.54 - 1.3)	0.988	0.234	0.161	0.112	0.63 (0.35 - 1.12)	YES
10_1R105	0.972	0.807	0.245	0.237	0.844	0.96 (0.64 - 1.45)	0.417	0.250	0.203	0.321	0.77 (0.45 - 1.3)	YES
10_1R117	0.895	0.761	0.202	0.186	0.700	0.91 (0.55 - 1.5)	0.422	0.192	0.140	0.273	0.68 (0.34 - 1.36)	YES
10_1R218	0.958	0.426	0.202	0.164	0.295	0.78 (0.49 - 1.25)	0.965	0.199	0.140	0.175	0.66 (0.36 - 1.21)	YES
10_2R420	0.982	0.827	0.246	0.231	0.698	0.92 (0.61 - 1.39)	0.238	0.252	0.195	0.224	0.72 (0.42 - 1.23)	YES
10_3M171	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_4Y100	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
10_6R426	0.976	0.978	0.001	0.000	0.631	NA	0.965	0.004	0.000	0.508	NA	LOW
10_6Y135	0.984	0.450	0.199	0.188	0.753	0.93 (0.6 - 1.45)	0.964	0.195	0.172	0.585	0.86 (0.49 - 1.5)	YES
10_9R210	0.970	0.978	0.001	0.000	0.629	NA	0.965	0.004	0.000	0.503	NA	LOW
12B_01M115	0.974	0.736	0.162	0.234	0.035, CV 0.049	1.58 (1.03 - 2.41)	0.919	0.170	0.207	0.396	1.27 (0.73 - 2.21)	YES
12B_01Y90	0.944	0.777	0.163	0.226	0.068	1.5 (0.97 - 2.33)	0.863	0.169	0.211	0.341	1.31 (0.75 - 2.28)	YES
12B_02M407	0.968	0.379	0.315	0.314	0.991	1 (0.69 - 1.45)	0.529	0.323	0.322	0.984	1 (0.63 - 1.59)	YES
12B_02W232	0.970	0.405	0.147	0.078	0.024, CV 0.021	0.49 (0.26 - 0.92)	0.449	0.149	0.095	0.153	0.6 (0.3 - 1.22)	YES
12B_02Y146	0.952	0.002	0.087	0.127	0.133	1.52 (0.88 - 2.65)	0.992	0.088	0.134	0.176	1.61 (0.8 - 3.21)	OUT OF HWE
12B_02Y190	0.970	0.603	0.144	0.077	0.026, CV 0.035	0.5 (0.27 - 0.93)	0.863	0.145	0.093	0.163	0.61 (0.3 - 1.23)	YES
12B_03R196	0.964	0.489	0.323	0.309	0.740	0.94 (0.64 - 1.37)	0.511	0.337	0.316	0.686	0.91 (0.57 - 1.45)	YES
12B_03R462	0.972	0.913	0.167	0.234	0.052	1.52 (0.99 - 2.33)	0.225	0.174	0.224	0.253	1.37 (0.8 - 2.35)	YES
12B_03Y82	0.964	0.569	0.143	0.083	0.047, CV 0.048	0.54 (0.3 - 1)	0.804	0.143	0.102	0.265	0.68 (0.34 - 1.35)	YES
12B_10R105	0.974	0.956	0.003	0.000	0.488	NA	0.930	0.008	0.000	0.337	NA	LOW
12B_12Y142	0.972	0.956	0.003	0.000	0.492	NA	1.000	0.000	0.000	NA	NA	LOW

SNP ID	Call Rate	HWE	MAF C	MAF ED	P Value	OR (95% CI)	HWE S	MAF C S	MAF ED S	P Value S	OR S (95% CI)	Heterozygous
2_12Y206	0.976	0.934	0.004	0.000	0.405	NA	1.000	0.000	0.000	NA	NA	LOW
4_12M397	0.996	0.475	0.351	0.302	0.237	0.8 (0.55 - 1.16)	0.023	0.327	0.298	0.571	0.88 (0.55 - 1.39)	YES
4_13S97	0.972	0.475	0.183	0.212	0.408	1.2 (0.78 - 1.85)	0.228	0.174	0.220	0.287	1.34 (0.78 - 2.3)	YES
4_1K110	0.966	0.562	0.354	0.305	0.252	0.8 (0.55 - 1.17)	0.029	0.323	0.302	0.681	0.91 (0.56 - 1.45)	YES
4_22Y152	0.962	0.240	0.353	0.305	0.260	0.81 (0.55 - 1.18)	0.018	0.326	0.302	0.643	0.89 (0.56 - 1.44)	YES
4_25Y336	0.962	0.000	0.071	0.032	0.075	0.44 (0.17 - 1.12)	0.000	0.058	0.009	0.029	0.14 (0.02 - 1.09)	OUT OF HWE
4_2M351	0.954	0.989	0.346	0.305	0.334	0.83 (0.57 - 1.21)	0.029	0.323	0.302	0.681	0.91 (0.56 - 1.45)	YES
4_7S246	0.970	0.001	0.458	0.481	0.601	1.1 (0.77 - 1.56)	0.000	0.496	0.475	0.697	0.92 (0.59 - 1.42)	OUT OF HWE
4_8R458	0.972	0.529	0.185	0.209	0.489	1.16 (0.76 - 1.79)	0.190	0.172	0.217	0.295	1.33 (0.78 - 2.29)	YES
6_10Y257	0.976	0.846	0.011	0.000	0.198	NA	0.895	0.011	0.000	0.245	NA	YES
6_18R120	0.851	0.000	0.018	0.013	0.696	0.74 (0.16 - 3.4)	0.000	0.023	0.018	0.750	0.77 (0.15 - 3.87)	FAIL WATER GENOTYPING
6_20R191	0.966	0.774	0.205	0.201	0.915	0.98 (0.63 - 1.51)	0.228	0.200	0.172	0.530	0.83 (0.47 - 1.47)	YES
6_20R240	0.962	0.400	0.156	0.164	0.786	1.07 (0.66 - 1.72)	0.890	0.156	0.132	0.533	0.82 (0.43 - 1.54)	YES
6_20R412	0.974	0.642	0.090	0.109	0.459	1.24 (0.7 - 2.19)	0.924	0.083	0.076	0.815	0.91 (0.41 - 2.04)	YES
6_6R431	0.970	0.176	0.105	0.083	0.417	0.77 (0.42 - 1.44)	0.656	0.106	0.085	0.520	0.78 (0.37 - 1.66)	YES
6_7R485	0.964	0.194	0.315	0.329	0.735	1.07 (0.73 - 1.55)	0.617	0.305	0.368	0.230	1.33 (0.84 - 2.11)	YES
6_7S166	0.972	0.868	0.009	0.006	0.739	0.7 (0.08 - 5.84)	0.895	0.011	0.008	0.781	0.73 (0.08 - 7.05)	YES
6_8R289	0.974	0.912	0.006	0.006	0.960	1.06 (0.12 - 9.53)	0.930	0.008	0.008	0.938	1.1 (0.1 - 12.26)	FAIL WATER GENOTYPING
6_8W328	0.974	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
ANK_E4B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
ANK_E5	0.964	0.589	0.029	0.032	0.838	1.11 (0.41 - 3.02)	0.717	0.031	0.034	0.872	1.11 (0.33 - 3.75)	YES
ANK_I1	0.966	0.909	0.342	0.344	0.958	1.01 (0.7 - 1.46)	0.966	0.327	0.362	0.506	1.17 (0.74 - 1.85)	YES
ANK_I2	0.968	0.124	0.354	0.301	0.213	0.79 (0.54 - 1.15)	0.249	0.317	0.271	0.371	0.8 (0.5 - 1.3)	YES
ANK_I3	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
ANK_I4	0.966	0.459	0.292	0.292	0.996	1 (0.68 - 1.47)	0.914	0.281	0.310	0.559	1.15 (0.72 - 1.86)	YES
ATP_I10	0.966	0.014	0.380	0.331	0.259	0.81 (0.56 - 1.17)	0.972	0.373	0.328	0.396	0.82 (0.52 - 1.3)	OUT OF HWE
ATP_I15	0.857	0.051	0.386	0.312	0.107	0.72 (0.49 - 1.07)	0.933	0.387	0.314	0.201	0.73 (0.44 - 1.19)	YES
ATP_I5	0.966	0.014	0.383	0.331	0.228	0.8 (0.55 - 1.15)	0.972	0.373	0.328	0.396	0.82 (0.52 - 1.3)	OUT OF HWE
ATP_I6	0.417	0.967	0.003	0.037	0.016, CV 0.072	10.96 (0.98 - 123.1)	0.934	0.013	0.026	0.614	2.03 (0.12 - 33.32)	FAIL LOW GENOTYPING RATE
ATP_I7	0.956	0.011	0.380	0.320	0.167	0.77 (0.53 - 1.12)	0.972	0.373	0.325	0.368	0.81 (0.51 - 1.29)	OUT OF HWE
ATP_I8	0.962	0.001	0.395	0.344	0.245	0.8 (0.56 - 1.16)	0.464	0.392	0.345	0.380	0.82 (0.52 - 1.29)	OUT OF HWE
ATP_I8B	0.929	0.008	0.370	0.311	0.178	0.77 (0.52 - 1.13)	1.000	0.375	0.313	0.249	0.76 (0.47 - 1.22)	OUT OF HWE
ATP_P1	0.805	0.018	0.126	0.217	0.007	1.92 (1.19 - 3.1)	0.052	0.157	0.194	0.403	1.29 (0.71 - 2.36)	FAIL LOW GENOTYPING RATE

SNP ID	Call Rate	HWE	MAF C	MAF ED	P Value	OR (95% CI)	HWE S	MAF C S	MAF ED S	P Value S	OR S (95% CI)	Heterozygous
IL1A10084	0.980	0.825	0.012	0.019	0.499	1.58 (0.42 - 6.03)	0.860	0.015	0.016	0.927	1.08 (0.2 - 6)	YES
IL1A11235	0.964	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1A12227	0.974	0.805	0.200	0.195	0.891	0.97 (0.62 - 1.51)	0.091	0.218	0.207	0.807	0.94 (0.55 - 1.6)	YES
IL1AA	0.772	0.974	0.002	0.000	0.583	NA	1.000	0.000	0.000	NA	NA	LOW
IL1AB	0.843	0.903	0.007	0.007	0.930	0.91 (0.1 - 8.17)	0.961	0.005	0.000	0.471	NA	LOW
IL1AC	0.873	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AD	0.831	0.002	0.410	0.407	0.952	0.99 (0.68 - 1.44)	0.112	0.453	0.433	0.719	0.92 (0.58 - 1.45)	OUT OF HWE
IL1AE	0.927	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X221	0.952	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
IL1AE7X255	0.966	0.002	0.421	0.396	0.573	0.9 (0.63 - 1.29)	0.075	0.462	0.414	0.386	0.82 (0.53 - 1.28)	OUT OF HWE
LEPRA	0.875	0.929	0.005	0.000	0.377	NA	1.000	0.000	0.000	NA	NA	LOW
LEPRB	0.843	0.029	0.264	0.197	0.094	0.69 (0.44 - 1.07)	0.364	0.292	0.175	0.02, CV 0.024	0.52 (0.29 - 0.91)	OUT OF HWE
LEPRC	0.859	0.028	0.214	0.203	0.774	0.94 (0.6 - 1.47)	0.115	0.198	0.191	0.875	0.96 (0.54 - 1.7)	OUT OF HWE
MMP13A	0.869	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP13B	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP13D	0.847	0.491	0.044	0.061	0.382	1.42 (0.65 - 3.12)	0.725	0.033	0.055	0.352	1.69 (0.55 - 5.16)	YES
MMP13E	0.897	0.977	0.002	0.000	0.623	NA	0.965	0.004	0.000	0.516	NA	LOW
MMP13F	0.873	0.976	0.002	0.000	0.604	NA	1.000	0.000	0.000	NA	NA	LOW
MMP13G	0.863	0.175	0.032	0.052	0.252	1.64 (0.7 - 3.84)	0.765	0.028	0.043	0.467	1.56 (0.47 - 5.23)	YES
MMP13H	0.887	0.977	0.002	0.000	0.607	NA	1.000	0.000	0.000	NA	NA	LOW
MMP13J	0.875	0.977	0.002	0.000	0.622	NA	0.965	0.004	0.000	0.524	NA	LOW
MMP13K	0.895	0.000	0.036	0.006	0.057	0.18 (0.02 - 1.32)	0.000	0.040	0.000	0.029	NA	OUT OF HWE
MMP2A	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP3A	0.885	0.476	0.081	0.071	0.694	0.87 (0.44 - 1.73)	0.355	0.080	0.078	0.929	0.96 (0.42 - 2.22)	YES
MMP3B	0.996	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
MMP9_D	0.962	0.516	0.200	0.221	0.571	1.13 (0.74 - 1.73)	0.913	0.200	0.241	0.365	1.27 (0.76 - 2.15)	YES
MMP9_I8	0.960	0.000	0.271	0.342	0.081	1.4 (0.96 - 2.04)	0.000	0.269	0.362	0.069	1.54 (0.97 - 2.46)	OUT OF HWE
MMP9_P1	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9_P4	0.829	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
MMP9_U1	0.948	0.038	0.158	0.214	0.093	1.46 (0.94 - 2.26)	0.159	0.151	0.237	0.043, CV 0.061	1.75 (1.01 - 3.01)	OUT OF HWE
MMP9_U2	0.972	0.527	0.289	0.285	0.909	0.98 (0.67 - 1.44)	0.819	0.323	0.292	0.539	0.86 (0.54 - 1.38)	YES
MMP9A	0.440	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL LOW GENOTYPING RATE
MMP9C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
MMP9D	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
SPARC_I2	0.970	0.799	0.062	0.045	0.434	0.72 (0.32 - 1.64)	0.455	0.062	0.043	0.472	0.69 (0.25 - 1.92)	YES
SPARC_I2B	0.958	0.367	0.048	0.045	0.881	0.94 (0.41 - 2.17)	0.581	0.046	0.042	0.869	0.91 (0.32 - 2.66)	YES

SNP ID	Call Rate	HWE	MAF C	MAF ED	P Value	OR (95% CI)	HWE S	MAF C S	MAF ED S	P Value S	OR S (95% CI)	Heterozygous
SPARC_I3	0.887	0.003	0.047	0.028	0.300	0.57 (0.2 - 1.66)	0.446	0.062	0.019	0.079	0.29 (0.06 - 1.26)	OUT OF HWE
SPARC_P2	0.599	0.000	0.307	0.224	0.142	0.65 (0.37 - 1.16)	0.000	0.333	0.190	0.039	0.47 (0.23 - 0.97)	FAIL LOW GENOTYPING RATE
SPARC_P3	0.919	0.007	0.341	0.403	0.164	1.3 (0.9 - 1.89)	0.659	0.327	0.426	0.071	1.53 (0.96 - 2.42)	OUT OF HWE
SPARC_P4	0.962	0.020	0.398	0.429	0.469	1.14 (0.8 - 1.62)	0.837	0.415	0.432	0.759	1.07 (0.69 - 1.66)	OUT OF HWE
SPARC_P5	0.968	0.026	0.260	0.173	0.023	0.6 (0.38 - 0.94)	0.944	0.265	0.144	0.009, CV 0.01	0.47 (0.26 - 0.84)	OUT OF HWE
SPARC_P7	0.966	0.033	0.346	0.404	0.176	1.28 (0.9 - 1.83)	0.659	0.327	0.432	0.048, CV 0.055	1.57 (1 - 2.45)	OUT OF HWE
TIMP1A	0.885	0.000	0.033	0.020	0.406	0.6 (0.17 - 2.04)	0.000	0.027	0.018	0.589	0.64 (0.13 - 3.24)	OUT OF HWE
TIMP1B	0.970	0.000	0.020	0.020	0.986	1.01 (0.29 - 3.59)	0.000	0.008	0.017	0.395	2.3 (0.32 - 16.52)	OUT OF HWE
TIMP2_I1	0.960	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I1C	0.968	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I2	0.944	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL WATER GENOTYPING
TIMP2_I3	0.950	0.622	0.420	0.377	0.331	0.84 (0.58 - 1.2)	0.230	0.446	0.388	0.296	0.79 (0.5 - 1.23)	YES
TIMP2A	0.407	0.583	0.046	0.040	0.856	0.87 (0.19 - 4)	0.780	0.036	0.053	0.661	1.47 (0.26 - 8.38)	FAIL LOW GENOTYPING RATE
TIMP3A	0.869	0.430	0.366	0.289	0.078	0.71 (0.48 - 1.04)	0.761	0.356	0.307	0.367	0.8 (0.49 - 1.3)	YES
TIMP3B	0.879	0.380	0.376	0.317	0.190	0.77 (0.52 - 1.14)	0.620	0.374	0.330	0.428	0.83 (0.51 - 1.33)	YES
TIMP3C	0.061	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	FAIL NO PROBE
TIMP4_I1	0.980	0.000	0.025	0.013	0.358	0.51 (0.12 - 2.22)	0.000	0.015	0.017	0.880	1.14 (0.21 - 6.32)	FAIL WATER GENOTYPING
TIMP4_U	0.950	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TIMP4A	0.859	0.019	0.096	0.086	0.692	0.88 (0.47 - 1.66)	0.289	0.093	0.096	0.908	1.05 (0.48 - 2.27)	OUT OF HWE
TIMP4D	0.877	0.448	0.084	0.086	0.949	1.02 (0.54 - 1.94)	0.316	0.088	0.096	0.798	1.11 (0.51 - 2.42)	YES
TIMP4_D	0.877	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNC_D	0.960	0.015	0.037	0.032	0.795	0.88 (0.33 - 2.34)	0.683	0.035	0.026	0.656	0.74 (0.2 - 2.79)	OUT OF HWE
TNC_E1	0.982	0.000	0.127	0.127	0.982	0.99 (0.59 - 1.67)	0.004	0.095	0.158	0.074	1.78 (0.94 - 3.38)	FAIL WATER GENOTYPING
TNC_E2	0.968	0.117	0.102	0.091	0.692	0.89 (0.48 - 1.62)	0.048	0.092	0.112	0.552	1.24 (0.61 - 2.53)	YES
TNC_E25	0.962	0.584	0.137	0.158	0.509	1.18 (0.72 - 1.92)	0.336	0.119	0.164	0.240	1.45 (0.78 - 2.69)	YES
TNC_I24	0.960	0.656	0.135	0.162	0.379	1.24 (0.77 - 2.02)	0.336	0.119	0.164	0.240	1.45 (0.78 - 2.69)	YES
TNC_I4	0.968	0.039	0.333	0.321	0.760	0.94 (0.65 - 1.37)	0.003	0.312	0.347	0.489	1.18 (0.74 - 1.87)	OUT OF HWE
TNC_P1	0.970	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNC_P2	0.925	0.292	0.056	0.071	0.482	1.29 (0.64 - 2.59)	0.447	0.063	0.086	0.429	1.39 (0.61 - 3.17)	FAIL WATER GENOTYPING
TNC_P4	0.962	0.092	0.144	0.123	0.512	0.84 (0.5 - 1.42)	0.345	0.131	0.121	0.787	0.91 (0.47 - 1.77)	YES
TNC_P5	0.964	0.102	0.101	0.091	0.717	0.89 (0.49 - 1.64)	0.048	0.092	0.112	0.552	1.24 (0.61 - 2.53)	YES
TNF10252	0.960	0.120	0.047	0.006	0.019, CV 0.024	0.13 (0.02 - 0.97)	0.360	0.058	0.009	0.029, CV 0.036	0.14 (0.02 - 1.09)	YES
TNF10411	0.814	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO

SNP ID	*Call Rate*	*HWE*	*MAF C*	*MAF ED*	*P Value*	*OR (95% CI)*	*HWE S*	*MAF C S*	*MAF ED S*	*P Value S*	*OR S (95% CI)*	*Heterozygous*
TNF10513	0.968	1.000	0.000	0.006	NA	NA	1.000	0.000	0.009	NA	NA	NO
TNF6547	0.980	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNF7178	0.970	0.426	0.038	0.032	0.728	0.84 (0.32 - 2.23)	0.718	0.031	0.017	0.444	0.55 (0.11 - 2.62)	YES
TNF8647	0.976	0.004	0.230	0.263	0.381	1.2 (0.8 - 1.78)	0.007	0.223	0.280	0.235	1.35 (0.82 - 2.21)	OUT OF HWE
TNF9367	0.978	0.004	0.058	0.013	0.018	0.21 (0.05 - 0.88)	0.426	0.060	0.017	0.065	0.27 (0.06 - 1.19)	OUT OF HWE
TNF9585	0.994	0.000	0.024	0.025	0.935	1.05 (0.35 - 3.18)	0.930	0.008	0.032	0.065	4.4 (0.8 - 24.35)	FAIL WATER GENOTYPING
TNFEXON1AB	0.972	1.000	0.000	0.000	NA	NA	1.000	0.000	0.000	NA	NA	NO
TNFEXON4AAB	0.972	0.978	0.002	0.000	0.627	NA	1.000	0.000	0.000	NA	NA	LOW
ZSWIM_E8A	0.958	0.007	0.482	0.493	0.580	1.11 (0.78 - 1.58)	0.055	0.485	0.500	0.784	1.06 (0.69 - 1.65)	OUT OF HWE
ZSWIM_E8B	0.964	0.022	0.259	0.314	0.164	1.31 (0.9 - 1.92)	0.084	0.265	0.339	0.143	1.42 (0.89 - 2.27)	OUT OF HWE
ZSWIM_I6	0.962	0.454	0.040	0.039	0.964	0.98 (0.4 - 2.42)	0.581	0.046	0.043	0.895	0.93 (0.32 - 2.71)	YES
ZSWIM_I6B	0.913	0.008	0.097	0.040	0.025	0.39 (0.16 - 0.91)	0.148	0.081	0.044	0.200	0.52 (0.19 - 1.43)	OUT OF HWE
ZSWIM_I6C	0.343	0.942	0.010	0.009	0.903	0.84 (0.05 - 13.62)	0.919	0.020	0.012	0.722	0.61 (0.04 - 9.89)	FAIL LOW GENOTYPING RATE

Table 12

Minor allele frequencies (MAF) of SNPs in the Labrador Retriever (LR) and Golden Retriever (GR) control populations, stratified by sex (Male [M] and Female [F]) or neuter (Neutered [N] and Entire [E]) status and compared by Chi square analysis (P value, NA= Not applicable).

SNP	*MAF LR E*	*MAF LR N*	*P Value*	*MAF LR M*	*MAF LR F*	*P Value*	*MAF GR N*	*MAF GR E*	*P Value*
10_10S308	0.036	0.085	0.036	0.05	0.043	0.748	0.176	0.24	0.428
10_11R124	0.204	0.255	0.265	0.213	0.219	0.87	0.231	0.228	0.976
10_13Y85	0.235	0.311	0.106	0.25	0.252	0.956	0.353	0.425	0.444
10_14R553	0.226	0.292	0.157	0.234	0.248	0.733	0.353	0.403	0.593
10_1R105	0.229	0.311	0.083	0.25	0.243	0.867	0.375	0.428	0.583
10_1R117	0.184	0.24	0.202	0.192	0.202	0.789	0.136	0.157	0.803
10_1R218	0.187	0.245	0.187	0.199	0.201	0.964	0.176	0.174	0.968
10_2R420	0.232	0.315	0.079	0.252	0.248	0.917	0.344	0.428	0.381
10_3M171	0	0	NA	0	0	NA	0	0	NA
10_4Y100	0	0	NA	0	0	NA	0	0	NA
10_6R426	0.003	0	0.597	0.004	0	0.352	0	0	NA
10_6Y135	0.186	0.241	0.203	0.195	0.2	0.9	0.176	0.178	0.987
10_9R210	0.003	0	0.601	0.004	0	0.35	0	0	NA
12B_01M115	0.17	0.16	0.812	0.17	0.165	0.877	0.059	0.041	0.651
12B_01Y90	0.168	0.167	0.983	0.169	0.165	0.905	0.059	0.014	0.116
12B_02M407	0.327	0.259	0.18	0.323	0.299	0.57	0.382	0.283	0.253
12B_02W232	0.156	0.111	0.241	0.149	0.143	0.866	0.441	0.407	0.712
12B_02Y146	0.08	0.113	0.288	0.088	0.087	0.974	0.029	0.007	0.264
12B_02Y190	0.153	0.111	0.269	0.145	0.143	0.946	0.441	0.408	0.722
12B_03R196	0.332	0.287	0.383	0.337	0.304	0.43	0.382	0.28	0.239
12B_03R462	0.168	0.148	0.615	0.174	0.152	0.509	0.176	0.316	0.106
12B_03Y82	0.152	0.111	0.28	0.143	0.143	0.986	0.441	0.407	0.712
12B_10R105	0.005	0	0.449	0.008	0	0.19	0	0	NA
12B_12Y142	0	0.009	0.061	0	0.004	0.279	0	0	NA
2_12Y206	0.003	0	0.597	0	0.004	0.282	0	0	NA
4_12M397	0.34	0.306	0.499	0.327	0.339	0.776	0.353	0.434	0.385
4_13S97	0.192	0.194	0.949	0.174	0.213	0.275	0.029	0.066	0.416
4_1K110	0.341	0.292	0.345	0.323	0.339	0.706	0.353	0.438	0.364
4_22Y152	0.343	0.283	0.246	0.326	0.335	0.831	0.353	0.438	0.364
4_25Y336	0.048	0.093	0.079	0.058	0.058	0.987	0.438	0.46	0.292
4_2M351	0.338	0.298	0.446	0.323	0.336	0.757	0.353	0.438	0.364
4_7S246	0.466	0.5	0.528	0.496	0.446	0.273	0.382	0.5	0.214
4_8R458	0.193	0.194	0.976	0.172	0.219	0.191	0.029	0.068	0.4
6_10Y257	0.01	0.009	0.919	0.011	0.009	0.768	0	0	NA
6_18R120	0.021	0	0.132	0.023	0.009	0.219	0	0.027	0.344
6_20R191	0.194	0.176	0.671	0.2	0.179	0.549	0.412	0.396	0.865
6_20R240	0.145	0.142	0.926	0.156	0.13	0.412	0.353	0.342	0.908
6_20R412	0.08	0.057	0.419	0.083	0.065	0.445	0.353	0.329	0.788
6_6R431	0.106	0.113	0.824	0.106	0.109	0.925	0.059	0.076	0.723
6_7R485	0.324	0.292	0.539	0.305	0.33	0.551	0.029	0.068	0.392
6_7S166	0.01	0	0.293	0.011	0.004	0.381	0	0.014	0.493
6_8R289	0.003	0.028	0.009	0.008	0.009	0.89	0	0	NA
6_8W328	0	0	NA	0	0	NA	0	0	NA
ANK_E4	0	0	NA	0	0	NA	0	0	NA

SNP	MAF LR E	MAF LR N	P Value	MAF LR M	MAF LR F	P Value	MAF GR N	MAF GR E	P Value
ANK_E4B	0	0	NA	0	0	NA	0	0	NA
ANK_E5	0.037	0.028	0.656	0.031	0.039	0.614	0.125	0.105	0.744
ANK_I1	0.352	0.333	0.722	0.327	0.372	0.301	0.5	0.441	0.531
ANK_I2	0.32	0.38	0.248	0.317	0.352	0.406	0.471	0.447	0.386
ANK_I3	0	0	NA	0	0	NA	0	0	NA
ANK_I4	0.301	0.278	0.64	0.281	0.313	0.435	0.188	0.213	0.744
ATP_I10	0.394	0.37	0.662	0.373	0.406	0.455	0.412	0.283	0.141
ATP_I15	0.401	0.392	0.87	0.387	0.413	0.563	0.412	0.283	0.141
ATP_I5	0.395	0.37	0.639	0.373	0.409	0.42	0.375	0.283	0.301
ATP_I6	0	0.042	0.029	0.013	0	0.365	0	0	NA
ATP_I7	0.392	0.37	0.69	0.373	0.403	0.504	0.412	0.283	0.141
ATP_I8	0.411	0.389	0.68	0.392	0.422	0.508	0.375	0.293	0.364
ATP_I8B	0.393	0.37	0.67	0.375	0.403	0.534	0.412	0.287	0.154
ATP_P1	0.158	0.174	0.718	0.157	0.167	0.801	0.133	0.302	0.062
IL1A10084	0.015	0	0.198	0.015	0.009	0.513	0	0	NA
IL1A11235	0	0	NA	0	0	NA	0	0	NA
IL1A12227	0.198	0.204	0.904	0.218	0.178	0.269	0.294	0.243	0.539
IL1AA	0	0	NA	0	0	NA	0	0	NA
IL1AB	0.01	0	0.367	0.005	0.011	0.459	0	0	NA
IL1AC	0	0	NA	0	0	NA	0	0	NA
IL1AD	0.431	0.417	0.788	0.453	0.399	0.227	0.033	0.021	0.678
IL1AE	0	0	NA	0	0	NA	0	0	NA
IL1AE7X221	0	0	NA	0	0	NA	0	0	NA
IL1AE7X255	0.443	0.417	0.63	0.462	0.409	0.236	0.031	0.02	0.694
LEPRA	0.003	0	0.6	0	0.005	0.287	0	0	NA
LEPRB	0.288	0.279	0.865	0.292	0.28	0.777	0.029	0.013	0.504
LEPRC	0.189	0.233	0.351	0.198	0.199	0.984	0.235	0.24	0.954
MMP13A	0	0	NA	0	0	NA	0	0	NA
MMP13B	0	0	NA	0	0	NA	0	0	NA
MMP13D	0.051	0.033	0.493	0.033	0.062	0.169	0	0	NA
MMP13E	0.003	0	0.595	0.004	0	0.35	0	0	NA
MMP13F	0.003	0	0.612	0	0.005	0.289	0	0.007	0.633
MMP13G	0.039	0.033	0.805	0.028	0.049	0.276	0	0	NA
MMP13H	0.003	0	0.6	0	0.005	0.282	0	0	NA
MMP13J	0.003	0	0.593	0.004	0	0.355	0	0	NA
MMP13K	0.015	0.044	0.089	0.04	0	0.005	0.235	0.291	0.518
MMP2A	0	0	NA	0	0	NA	0	0	NA
MMP3A	0.085	0.056	0.36	0.08	0.077	0.884	0.088	0.16	0.286
MMP3B	0	0	NA	0	0	NA	0	0	NA
MMP9_D	0.184	0.213	0.492	0	0.179	0.549	0.147	0.237	0.254
MMP9_I8	0.259	0.333	0.129	0.151	0.283	0.731	0.059	0.018	0.208
MMP9_P1	0	0	NA	0	0	NA	0	0	NA
MMP9_P4	0	0	NA	0	0	NA	0	0	NA
MMP9_U1	0.146	0.167	0.587	0	0.149	0.95	0.029	0.019	0.712
MMP9_U2	0.301	0.306	0.928	0	0.278	0.281	0.412	0.461	0.178
MMP9A	0	0	NA	0	0	NA	0	0	NA
MMP9C	0	0	NA	0	0	NA	0	0	NA
MMP9D	0	0	NA	0	0	NA	0	0	NA
SPARC_I2	0.077	0.046	0.274	0.062	0.08	0.435	0.147	0.184	0.608
SPARC_I2B	0.055	0.028	0.248	0.046	0.052	0.758	0.147	0.153	0.927
SPARC_I3	0.051	0.019	0.15	0.062	0.022	0.032	0.235	0.226	0.908

SNP	*MAF LR E*	*MAF LR N*	*P Value*	*MAF LR M*	*MAF LR F*	*P Value*	*MAF GR N*	*MAF GR E*	*P Value*
SPARC_P2	0.344	0.286	0.402	0.333	0.333	1	0.231	0.177	0.535
SPARC_P3	0.327	0.361	0.509	0.327	0.344	0.696	0.353	0.393	0.662
SPARC_P4	0.401	0.435	0.518	0.415	0.4	0.729	0.382	0.34	0.64
SPARC_P5	0.277	0.204	0.123	0.265	0.257	0.824	0.265	0.263	0.985
SPARC_P7	0.328	0.361	0.521	0.327	0.345	0.671	0.471	0.461	0.467
TIMP1A	0.034	0.033	0.992	0.027	0.041	0.434	0.059	0.041	0.639
TIMP1B	0.016	0.028	0.405	0.008	0.031	0.056	0.063	0.041	0.585
TIMP2_I1	0	0	NA	0	0	NA	0	0	NA
TIMP2_I1C	0	0	NA	0	0	NA	0	0	NA
TIMP2_I2	0	0	NA	0	0	NA	0	0	NA
TIMP2_I3	0.434	0.407	0.619	0.036	0.409	0.409	0.382	0.187	0.013
TIMP2A	0.03	0.2	0	0	0.081	0.174	0	0	NA
TIMP3A	0.372	0.344	0.633	0.356	0.376	0.678	0.313	0.397	0.371
TIMP3B	0.371	0.38	0.871	0.374	0.372	0.957	0.233	0.384	0.118
TIMP3C	0	0	NA	0	0	NA	0	0	NA
TIMP4_D	0	0	NA	0	0	NA	0	0	NA
TIMP4_I1	0.016	0	0.191	0.093	0.009	0.52	0	0	NA
TIMP4_U	0	0	NA	0	0	NA	0	0	NA
TIMP4A	0.114	0.067	0.194	0	0.116	0.444	0.059	0.02	0.209
TIMP4D	0.111	0.057	0.131	0	0.112	0.411	0	0.021	0.752
TNC_D	0.034	0.046	0.563	0.015	0.04	0.762	0.059	0.066	0.881
TNC_E1	0.112	0.102	0.766	0.035	0.126	0.278	0.353	0.46	0.256
TNC_E2	0.106	0.102	0.906	0.095	0.119	0.33	0.353	0.467	0.226
TNC_E25	0.143	0.148	0.89	0.092	0.173	0.095	0.059	0.1	0.454
TNC_I24	0.141	0.148	0.851	0.119	0.17	0.114	0.059	0.099	0.466
TNC_I4	0.346	0.278	0.186	0.119	0.352	0.34	0.353	0.447	0.315
TNC_P1	0	0	NA	0	0	NA	0	0	NA
TNC_P2	0.071	0.078	0.81	0	0.084	0.394	0.412	0.412	0.997
TNC_P4	0.128	0.13	0.957	0.063	0.125	0.85	0.206	0.27	0.442
TNC_P5	0.106	0.102	0.906	0.131	0.119	0.33	0.353	0.461	0.253
TNF10252	0.043	0.037	0.792	0.092	0.023	0.054	0.147	0.189	0.565
TNF10411	0	0	NA	0	0	NA	0	0	NA
TNF10513	0	0	NA	0	0	NA	0	0	NA
TNF6547	0	0	NA	0	0	NA	0	0	NA
TNF7178	0.024	0.037	0.452	0	0.022	0.576	0.324	0.467	0.129
TNF8647	0.238	0.148	0.045	0.031	0.213	0.779	0.206	0.243	0.641
TNF9367	0.052	0.046	0.825	0.223	0.039	0.286	0.147	0.204	0.448
TNF9585	0.008	0	0.359	0.06	0.004	0.65	0	0	NA
TNFEXON1AB	0	0	NA	0	0	NA	0	0	NA
TNFEXON4AAB	0	0	NA	0	0	NA	0	0.013	0.498
ZSWIM_E8A	0.476	0.444	0.146	0	0.496	0.663	0.471	0.368	0.083
ZSWIM_E8B	0.246	0.185	0.186	0.485	0.196	0.068	0.029	0.007	0.248
ZSWIM_I6	0.045	0.056	0.656	0.265	0.049	0.879	0	0.007	0.635
ZSWIM_I6B	0.109	0.09	0.588	0.046	0.132	0.071	0.029	0.034	0.898
ZSWIM_I6C	0.014	0	0.562	0.081	0	0.335	0	0	NA

Table 13

Haplotype frequencies estimates (F) of the Labrador Retriever control population (Control) and their comparison with haplotype frequency estimates of the cohorts of Labrador Retrievers with cranial cruciate ligament rupture (CCLR), elbow dysplasia (ED) and hip dysplasia (HD) by corrected Chi squared or Fishers exact tests.

Gene	*SNPs*	*Ordered Haplotype*	*F Controls*	*F CCLR*	*P Value*	*F ED*	*P Value*	*F HD*	*P Value*
ANKRD10	ANK_I1 ANK_I2 ANK_I4 ANK_E5	A,G,G,G	34.7	38.1	0.263	29.3	0.150	38.4	0.841
		A,A,G,G	26.2	29.8		34.4		22.1	
		G,A,A,G	21.7	21.7		22.8		20.9	
		G,A,G,G	9.7	4.2		7.4		10.2	
		A,A,A,G	4.5	2.1		2.1		5.1	
		Other	3.1	4.1		4.0		41.7	
IL4	4_22Y152 4_13S97 4_12M397 4_8R458 4_2M351 4_1K110	C,G,A,A,A,G	46.3	48.3	**0.029**	48.8	0.493	35.9	**0.033**
		T,G,C,A,C,T	34.9	23.5		30.2		31.2	
		C,C,A,G,A,G	18.4	27.2		20.9		32.8	
		Other	0.5	1.0		0.0		0.0	
IL6	6_6R431 6_7R485 6_20R191 6_20R240 6_20R412	A,G,A,G,G	48.9	41.1	0.465	48.2	0.736	53.1	0.117
		A,A,A,G,G	30.5	37.9		31.9		37.5	
		A,G,G,A,A	9.1	8.8		10.9		6.3	
		G,G,G,A,G	6.8	8.5		5.1		0.0	
		Other	4.7	3.6		3.9		3.1	
IL10	10_1R105 10_1R117 10_1R218 10_2R420 10_6Y135 10_11R124 10_13Y85 10_14R553	A,A,G,G,C,A,T,G	74.3	66.5	0.559	73.3	0.571	84.4	0.275
		G,G,A,A,T,G,C,A	19.9	23.1		17.4		14.1	
		G,A,G,A,C,A,C,A	3.8	3.7		5.7		1.6	
		A,G,G,G,C,A,T,G	0.3	5.4		1.7		0.0	
		Other	1.7	1.3		2.0		0.0	
IL12B	12B_01M115 12B_01Y90 12B_02Y190 12B_02W232 12B_02M407 12B_03Y82 12B_03R196 12B_03R462	A,T,A,A,T,G,G,C	51.8	43.4	**0.023**	45.8	**0.010**	56.3	0.344
		A,T,C,A,T,A,A,C	16.9	17.6		22.4		21.9	
		C,C,A,A,T,G,G,C	15.8	27.2		22.9		12.5	
		A,T,C,T,C,A,G,T	14.5	10.7		7.4		7.8	
		Other	1.1	1.1		1.6		1.6	
MMP9	MMP9_U2 MMP9_D	G,C	56.2	51.2	0.148	52.5	0.410	47.2	0.124
		T,C	23.8	22.8		25.4		29.3	
		G,T	14.9	22.8		19.0		22.7	
		Other	5.1	3.2		3.1		0.7	
SPARC	SPARC_I2 SPARC_I2B	G,G	93.7	92.0	0.222	95.5	0.871	93.7	1.000
		C,A	4.6	8.0		4.5		4.7	
		Other	1.7	0.0		0.0		1.6	
TIMP3	TIMP3B TIMP3A	C,C	62.1	64.3	0.405	67.1	0.405	68.5	0.276
		A,T	36.5	35.7		30.0		25.7	
		Other	1.4	0.0		3.0		5.8	
TNC	TNC_P4 TNC_P5 TNC_E2 TNC_I24 TNC_E25	C,G,G,C,G	68.0	74.6	0.251	69.7	0.620	72.4	0.253
		T,G,G,C,G	12.7	13.3		10.3		10.2	
		C,G,G,T,A	8.2	3.7		9.6		3.5	
		C,T,A,C,G	5.2	2.1		3.5		1.8	
		Other	5.9	6.3		6.9		12.1	
TNFa	TNF7178 TNF10252	T,A	92.1	93.9	0.856	96.1	0.116	95.3	0.710
		T,T	4.1	3.2		3.1		3.1	
		Other	3.8	2.9		0.8		1.6	

Table 14

Haplotype frequencies estimates (F) of the Labrador Retriever control population (control), following stratification on the basis of sex (males or female) or neuter (neutered or entire) status, and their comparison by corrected Chi squared or Fishers exact tests (P Value).

Gene	*SNPs*	*Ordered Haplotype*	*F Controls*	*F Male*	*F Female*	*P Value*	*F Neutered*	*F Entire*	*P Value*
ANKRD10	ANK_I1 ANK_I2 ANK_I4 ANK_E5	A,G,G,G	34.7	31.7	34.2	0.333	36.9	31.8	0.377
		A,A,G,G	26.2	31.3	24.3		23.5	29.3	
		G,A,A,G	21.7	20.0	23.9		19.8	22.4	
		G,A,G,G	9.7	9.1	10.2		11.8	9.0	
		A,A,A,G	4.5	4.7	3.5		5.2	3.8	
		Other	3.1	3.0	3.9		2.8	3.6	
IL4	4_22Y152 4_13S97 4_12M397 4_8R458 4_2M351 4_1K110	C,G,A,A,A,G	46.3	49.6	45.2	0.519	50.9	46.7	0.654
		T,G,C,A,C,T	34.9	32.7	33.5		29.6	34.0	
		C,C,A,G,A,G	18.4	17.3	20.9		18.5	19.1	
		Other	0.5	0.4	0.4		0.9	0.3	
IL6	6_6R431 6_7R485 6_20R191 6_20R240 6_20R412	A,G,A,G,G	48.9	50.7	49.6	0.783	53.4	49.0	0.635
		A,A,A,G,G	30.5	29.0	32.6		29.0	31.4	
		A,G,G,A,A	9.1	8.4	6.5		5.9	8.0	
		G,G,G,A,G	6.8	7.3	6.5		8.8	6.4	
		Other	4.7	4.6	4.8		2.9	5.2	
IL10	10_1R105 10_1R117 10_1R218 10_2R420 10_6Y135 10_11R124 10_13Y85 10_14R553	A,A,G,G,C,A,T,G	74.3	74.0	74.3	0.955	67.6	76.0	0.146
		G,G,A,A,T,G,C,A	19.9	19.5	19.6		24.1	18.3	
		G,A,G,A,C,A,C,A	3.8	4.5	3.9		6.5	3.6	
		A,G,G,G,C,A,T,G	0.3	0.4	0.5		0.0	0.5	
		Other	1.7	1.5	1.8		1.9	1.6	
IL12B	12B_01M115 12B_01Y90 12B_02Y190 12B_02W232 12B_02M407 12B_03Y82 12B_03R196 12B_03R462	A,T,A,A,T,G,G,C	51.8	49.1	53.9	0.604	55.6	50.3	0.619
		A,T,C,A,T,A,A,C	16.9	17.7	15.2		14.8	17.0	
		C,C,A,A,T,G,G,C	15.8	16.6	16.1		15.7	16.5	
		A,T,C,T,C,A,G,T	14.5	14.6	14.3		11.1	15.5	
		Other	1.1	2.0	0.4		2.8	0.8	
MMP9	MMP9_U2 MMP9_D	G,C	56.2	52.6	60.0	0.346	56.4	55.6	0.273
		T,C	23.8	27.4	22.1		22.3	26.1	
		G,T	14.9	15.1	12.2		13.0	14.3	
		Other	5.1	4.9	5.7		8.3	4.0	
SPARC	SPARC_I2 SPARC_I2B	G,G	93.7	93.9	91.7	0.506	95.4	92.2	0.000
		C,A	4.6	4.5	5.2		2.8	5.4	
		Other	1.7	1.5	3.1		1.9	2.4	
TIMP3	TIMP3B TIMP3A	C,C	62.1	62.6	62.2	0.932	62.0	62.5	1.000
		A,T	36.5	36.5	36.8		35.7	36.9	
		Other	1.4	0.9	1.0		2.3	0.6	
TNC	TNC_P4 TNC_P5 TNC_E2 TNC_I24 TNC_E25	C,G,G,C,G	68.0	71.4	68.0	0.777	71.8	69.6	0.866
		T,G,G,C,G	12.7	12.2	9.6		8.1	11.5	
		C,G,G,T,A	8.2	6.7	8.8		6.7	7.9	
		C,T,A,C,G	5.2	4.5	5.1		4.8	4.6	
		Other	5.9	5.2	8.5		8.6	6.4	

Gene	*SNPs*	*Ordered Haplotype*	*F Controls*	*F Male*	*F Female*	*P Value*	*F Neutered*	*F Entire*	*P Value*
TNFα	TNF7178 TNF10252	T,A	92.1	91.4	95.9	0.416	93.4	93.5	0.595
		T,T	4.1	5.6	1.8		2.9	4.0	
		Other	3.8	3.1	2.2		3.7	2.5	

Table 15

Haplotype frequencies estimates (F) of the general Golden Retriever control population (Control), and their comparison with haplotype frequency estimates population of Golden Retrievers with cranial cruciate ligament rupture (CCLR) by corrected Chi squared or Fishers exact tests (P Value). Further comparison of the haplotype frequency estimates of the control population stratified on neuter status (neutered or entire) is also listed.

Gene	*SNPs*	*Ordered Haplotype*	*F Controls*	*F CCLR*	*P Value*	*F Entire*	*F Neuter*	*P Value*
ANKRD10	ANK_I1 ANK_I2 ANK_I4 ANK_E5	A,G,G,G	53.7	67.0	**0.029**	54.6	47.1	0.248
		G,A,G,G	24.4	13.6		22.5	33.3	
		G,A,A,C	10.9	4.5		10.5	13.3	
		G,A,A,G	8.8	13.6		10.4	3.3	
		Other	2.1	1.1		2.0	2.9	
ATP11B	ATP_I5 ATP_I7 ATP_I8 ATP_I8B ATP_I10 ATP_I15	A,T,A,G,A,T	68.6	71.1	0.563	71.0	58.8	0.257
		C,G,C,A,C,G	30.9	26.3		28.3	41.2	
		Other	0.5	2.6		0.7	0.0	
IL1A	IL1A12227 IL1AE7X255	T,G	73.4	81.1	0.128	73.8	70.6	0.817
		C,G	24.4	15.6		24.3	26.5	
		Other	2.2	3.3		2.0	2.9	
IL4	4_25Y336 4_22Y152 4_13S97 4_12M397 4_8R458 4_7S246 4_2M351 4_1K110	C,C,G,A,A,C,A,G	47.2	54.1	0.210	45.3	55.6	0.422
		T,T,G,C,A,G,C,T	42.0	33.3		43.4	35.3	
		T,C,C,A,G,G,A,G	5.3	5.6		5.9	2.9	
		T,C,G,A,A,C,A,G	4.9	7.1		4.7	0.0	
		Other	0.6	0.0		0.7	6.2	
IL6	6_6R431 6_7R485 6_20R191 6_20R240 6_20R412	A,G,A,G,G	54.0	43.0	0.221	53.0	55.9	1.000
		A,G,G,A,A	33.7	43.2		33.7	35.3	
		A,A,A,G,G	4.5	5.4		4.7	2.9	
		G,G,A,G,G	2.8	5.0		3.5	0.0	
		Other	5.1	3.4		5.0	5.9	
IL10	10_1R105 10_1R218 10_2R420 10_6Y135 10_10S308 10_11R124 10_13Y85 10_14R553	A,G,G,C,C,A,T,G	58.5	73.3	**0.021**	57.2	64.7	0.677
		G,G,A,C,G,A,C,A	23.9	12.4		25.0	17.6	
		G,A,A,T,C,G,C,A	16.5	10.0		16.4	17.6	
		Other	1.1	4.2		1.4	0.0	
IL12B	12B_02Y190 12B_02Y146 12B_02W232 12B_03R462 12B_03Y82	C,C,T,G,T	41.5	30.9	0.240	40.8	44.1	0.238
		C,T,A,A,C	29.3	32.8		31.6	17.6	
		C,T,A,G,C	28.1	35.0		26.9	35.3	
		Other	1.1	1.3		0.7	2.9	

Gene	*SNPs*	*Ordered Haplotype*	*F Controls*	*F CCLR*	*P Value*	*F Entire*	*F Neuter*	*P Value*
LEPR	LEPRC LEPRB	A,A	75.8	71.3	0.381	76.0	76.5	1.000
		T,A	22.6	27.5		22.7	20.6	
		Other	1.6	1.2		1.3	2.9	
SPARC	SPARC_P3 SPARC_P4 SPARC_P5 SPARC_P7 SPARC_I2 SPARC_I2B SPARC_I3	T,G,G,T,G,G,G	37.8	36.4	0.907	38.8	35.3	0.776
		A,G,A,A,G,G,A	23.4	22.6		23.0	23.5	
		A,A,G,A,G,G,G	21.3	22.8		19.7	26.5	
		A,A,G,T,C,A,G	14.4	11.2		15.1	11.8	
		Other	3.2	6.9		3.3	2.9	
TIMP3	TIMP3B TIMP3A	A,T	63.4	68.6	0.427	61.3	70.6	0.431
		C,C	36.6	31.4		38.7	29.4	
		Other	0.0	0.0		0.0	0.0	
TIMP4	TIMP4D TIMP4A	C,G	97.3	96.5	1.000	98.0	94.1	0.519
		Other	2.7	3.5		2.0	5.9	
TNC	TNC_P4 TNC_P5 TNC_E2 TNC_I4 TNC_I24 TNC_E25	C,T,A,G,C,G	54.7	45.5	0.209	52.6	64.7	0.221
		T,G,G,T,C,G	25.5	32.4		27.0	17.6	
		C,G,G,T,C,G	9.2	5.9		8.6	11.8	
		C,G,G,T,T,A	7.8	13.8		9.1	2.9	
		Other	2.8	2.4		2.7	2.9	
TNFa	TNF7178 TNF8647 TNF9367 TNF10252	A,A,T,A	44.2	46.9	0.609	46.9	32.4	0.111
		T,A,T,A	31.8	27.5		28.8	47.1	
		T,C,C,T	18.6	23.3		19.0	14.7	
		Other	5.3	2.3		5.3	5.9	
ZSWIM2	ZSWIM_E8B ZSWIM_E8A	G,C	60.1	54.7	0.439	63.2	44.1	0.051
		G,T	38.8	44.2		36.2	52.9	
		Other	1.1	1.1		0.7	2.9	

Publications

List of Papers Published

Clements, D.N., Carter, S.D, Innes, J.F. and Ollier, W.E. (2006) The genetic basis of canine osteoarthritis secondary to joint dysplasia: A comparative review. *American Journal of Veterinary Research* **67**, 909-918

Clements, D.N., Vaughan-Thomas, A., Peansukmanee, S., Carter, S.D., Innes, J.F., Ollier, W.E.R. and Clegg, P.D. (2006) Assessment of the use of RNA quality metrics for the screening of normal and pathological canine articular cartilage samples. *American Journal of Veterinary Research* **67**, 1438-1444

Clements, D.N., Carter, S.D., Innes, J.F.I., Ollier, W.E.R. and Day, P.J.R. (2006) Analysis of normal and osteoarthritic canine cartilage mRNA expression by quantitative real-time PCR. *Arthritis Research and Therapy* **8**, R158

Clements, D.N., Carter, S.D., Innes, J.F., Ollier, W.E.R., Day, P.J. (2007) Gene expression profiling of normal and ruptured canine anterior cruciate ligaments. *Osteoarthritis and Cartilage , In Press*

Clements, D.N., Fitzpatrick, N., Carter, S.D., Day, P.J. (2007) Cartilage gene expression correlates with radiographic severity of canine elbow osteoarthritis. *The Veterinary Journal. In Press.*

Clements, D.N., Wood, S., Carter, S.D., Ollier, W.E.R. (2007) Assessment of the quality and quantity of genomic DNA recovered from canine blood samples by three different extraction methods. *Research in Veterinary Science. In Press.*

Ayers, D., **Clements, D.N.**, Salway, F., Day, P.D. (2007) Expression stability of commonly used reference genes in canine articular connective tissues. *BMC Veterinary Research* **3**, 7

Maccoux, L., **Clements, D.N.**, Salway, F., Day, P.J. (2007) Identification of Novel Reference Genes from Microarray Data. *BMC Molecular Biology* **8**, 62

Maccoux, L.J., Salway, F., Day, P.J.R., **Clements, D.N.** (2007) Expression profiling of select cytokines in canine osteoarthritis tissues. *Veterinary Immunology and Immunopathology* **118**, 59-67

Abstracts Presented

Abstract 1

Presented at the British Society for Matrix Biology Autumn Conference, Manchester, 12-13th September 2005

Abstract published the International Journal of Experimental Pathology (2006) **87** (1), A11-A12

Evaluation of the quality of RNA extracted from healthy and osteoarthritic canine articular cartilage

Dylan N. Clements[1], Siriporn Peansukmanee[1], William E.R. Ollier[2], Stuart Carter[1], John F. Innes[1], Anne Vaughan-Thomas[2].

[1]Connective Tissue Research Group, Faculty of Veterinary Science, University of Liverpool, Liverpool, L69 3BX
[2]Centre for Integrated Genomic Medical Research, University of Manchester, Oxford Rd, Manchester, M13 9PT.

Introduction

RNA extraction from canine articular cartilage is problematic, and usually yields low volumes of RNA of low concentration. RNA quality is defined by the sum of RNA purity and integrity. Ultimately, RNA quality may affect the results of downstream expression profiling, such as mRNA qualification. A number of metric assessments using small sample volumes (1μl) are now available to determine the quality of RNA extractions, although they may add significantly to the cost and time of the extraction procedure. Metric assessments were used to evaluate the quality of healthy and osteoarthritic canine articular cartilage samples submitted to a collection program, to determine their usefulness for performing expression analysis.

Methods

Chondrocyte cultures (n = 12), and normal (n = 12) and osteoarthritic (n = 37) canine articular cartilage samples were extracted using three different methods of phenol-chloroform extraction (with or without isopropanol precipitation and silica membrane clean up). The quality metrics of each sample were determined using; a UV spectrophotometer (Ultrospec 2000), a Nanodrop spectrophotometer and microfluidic capillary electrophoresis (Agilent 2100 Bioanalyser). The spectrophotometer absorbance (A) A_{260} : A_{280} ratio, Nanodrop A_{260} : A_{280} ratio and A_{260} : A_{230} ratios were recorded. The electrophorectic trace was used to calculate the ribosomal peak ratio (RR), the degradation factor (DF), the RNA integrity number (RIN). The metrics were

compared to a conventional assessment of RNA integrity (visual analysis of the electropherogram) to determine the most useful quality metrics.

Results

No differences in methods of extraction were determined using the metric assessments. RNA extracted from cultured chondrocytes was of higher quality than that extracted from normal canine cartilage, and osteoarthritic canine cartilage was of the lowest quality. The RIN and the RR were the most sensitive metrics for determining high RNA integrity, whereas the DF was most specific. The RIN and DF could not be calculated for all samples. Moderate correlations were found between the purity (as determined by absorbance ratios) and integrity (as determined by the electrophoretic trace) of extracted RNA. A significant proportion (35%) of osteoarthritic articular cartilage samples were determined as being of low quality.

Discussion

No single metric assessment provided a sensitive or specific assessment of RNA quality. Although spectrophotometer determination of absorbance ratios provides a broad estimate of quality, metrics using the results of microfluidic capillary assessment provide more accurate assessments of quality. We recommend that the DF, RR and RIN are assessed for all RNA extraction procedures to audit and maintain the quality of sample preparation, and minimise variation in downstream profiling.

Abstract 2

Presented at the British Society for Matrix Biology Spring Conference, Cambridge, 10th-11th April 2006

Abstract published the Proceedings of the British Society for Matrix Biology Spring Conference (2006), p6

Matrix-associated gene expression in naturally-occurring canine osteoarthritis

Dylan N. Clements[1], Stuart D. Carter[1], John, F. Innes[1], William, E.R. Ollier[2], Philip J.R. Day[2]

[1]Connective Tissue Research Group, Faculty of Veterinary Science, University of Liverpool, L69 3BX, UK.
[2]Centre for Integrated Genomic Medical Research, University of Manchester, M13 9PT, UK.

Introduction:
Osteoarthritis (OA) is the most common debilitating disease of mammalian joints. The dog is commonly used as an experimental model of OA for investigating both the pathogenesis and treatment of the disease. This study hypothesises that if the expression of selected proteases, matrix molecules and collagens is modulated in naturally-occurring canine OA, then this expression data would provide a means to compare pathological changes found with those reported in naturally-occurring human OA and experimentally induced canine OA.

Methods
Articular cartilage was harvested from the femoral heads of dogs with and without OA. RNA was extracted and real-time RT-PCR assays designed for structural matrix molecules; COL1A2, COL2A1, COL3A1, COL5A1, COL9A3, AGC1, BGN, CSPG2, DCN, LUM, TNC, VIM, proteases; ADAMTS5, CTSB, CTSD, MMP9, MMP13, protease inhibitors; TIMPs -1, -2 and -4 and genomic DNA and reference genes; GAPDH, TBP, RPL13A, and SDHA.

Results
Three genes (SDHA, TIMP2 and TIMP4) were determined to be significantly down-regulated in canine OA cartilage. Thirteen genes (AGC1, BGN, COL1A2, COL2A1, COL3A1, COL5A1, CSPG2, CTSB, CSTD, LUM, MMP13, TIMP1 and TNC) were determined to be significantly up-regulated in the OA samples. SDHA was unsuitable for use as a reference gene in canine osteoarthritic articular cartilage.

Discussion
The altered cartilage expression of genes for selected matrix molecules and key mediators of the proteolytic degradation in naturally-occurring canine hip OA are broadly similar to the changes reported in experimental canine stifle osteoarthritis and

naturally occurring human osteoarthritis (hip and knee). This suggests that the pathological mechanisms underlying the development and progression of canine osteoarthritis are likely to resemble those associated with the development of human osteoarthritis.

Abstract 3

Presented at the British Small Animal Veterinary Association Annual Congress, Birmingham 20th April 2006

Abstract Published in the Proceedings of the British Small Animal Veterinary Association Congress (2006), p496

Gene expression profiling of normal and ruptured cranial cruciate ligament

Clements DN[1], Innes JF[1], Carter SD[1], Ollier WER[2].

[1]Connective Tissue Research Group, Faculty of Veterinary Science, The University of Liverpool, L69 7ZJ.
[2]Centre for Integrated Genomic Medical Research, The University of Manchester, M13 9PT

Introduction
Biochemical differences exist between the composition of normal cranial cruciate ligaments (CCL) from breeds of dogs predisposed to CCL rupture (e.g. Labrador Retriever [LR]) when compared to those at low risk (e.g. Greyhounds [GH]) from CCL rupture. We evaluated differences in gene expression in normal and ruptured CCL tissue from these breeds using a canine whole genome oligonucleotide microarray.

Methods
Grossly normal CCL was harvested from the healthy stifles of 5 LRs and 5 GHs euthanatized for welfare reasons, with natural occurring, non-orthopaedic disease. CCL remnants were harvested from the stifles of 5 LRs with CCL rupture undergoing routine surgical treatment of the condition. Messenger RNA was extracted and amplified. The aRNA samples were labelled and hybridized to custom designed 44,000 gene canine whole genome oligonucleotide microarray chips. Array scan data were normalised by locally weighted linear regression, and the comparison of expression data between the groups performed using corrected t-tests.

Results
779 transcripts were significantly up-regulated (to different degrees) in the normal GH CCL, compared to the normal LR CCL, including BMP3, COL6A3, PDGFC, GAPDG and IGF1. 892 transcripts were significantly down-regulated in the normal GH CCL compared to the normal LR CCL, including IL-4, SAA1, CHST11, BMP1 and TNC. 4038 transcripts were up-regulated in pathological LR CCL when compared to normal CCL, including IL-1α, ADAM4, COL1A2, GAPDH and NOS. 5419 transcripts were down-regulated in pathological LR CCL when compared to normal LR CCL, including ADAM5, IGF1, BMP1 and HAS1.

Conclusions

A larger number of genes are up- or down- (dys-) regulated in the pathological LR CCL when compared to the normal LR CCL. 44% of genes dys-regulated in normal LR CCL, when compared to the normal GH CCL, were also dys-regulated in pathological LR CCL suggesting that some of these genes may have a role in the development of CCL rupture. The gene expression profiles indicated that there is increased transcriptional activity in the normal GH CCL when compared to the normal LR CCL. GAPDH is unsuitable as a control gene for quantitative PCR of genes dys-regulated in CCL rupture. Understanding the pathological mechanism of this disabling disease will may ultimately allow prevention or better treatment in the future.

Abstract 4

Presented at the European College of Veterinary Surgeons Annual Congress, Seville, 29th June 2006

Abstract Published in the Veterinary Surgery (2006) **35** (4), E2

Gene expression Profiles of Normal and Osteoarthritic Coxofemoral Articular Cartilage

Clements DN*[1,2], Innes JF[1], Carter SD[1], Ollier WER[2].

[1]University of Liverpool, Liverpool, UK.
[2]Centre for Integrated Genomic Medical Research, Manchester, UK.

Introduction
Canine osteoarthritis (OA) is a complex disease involving the dys-regulation of multiple genes. The development of a canine whole genome oligonucleotide microarray allows for the global assessment of gene expression in normal and diseased tissues, such as articular cartilage.

Materials and Methods
Articular cartilage was harvested from a normal hip joint of 5 Labrador Retrievers euthanatized for welfare reasons, with natural occurring, non-orthopaedic disease, and from a diseased hip joint of 5 Labrador retrievers with OA (secondary to hip dysplasia), undergoing total hip replacement surgery. The extracted mRNA was amplified using a double amplification procedure. The aRNA samples were labelled and hybridized to custom designed 44,000 canine whole genome oligonucleotide microarray chips. Array data were normalised by locally weighted linear regression, and the comparison of expression data between the groups performed using corrected t-tests.

Results
In the OA samples 2286 transcripts were dys-regulated, of which 1920 were annotated. 1399 transcripts were up-regulated (1192 annotated), and 887 transcripts were down-regulated (728 annotated). Up regulated genes of interest included; TIMP2, MMP7 and IL12. Down regulated genes of interest included TNFSF11, BMP2, IGFBP2, CSTD and HAS1.

Discussion
Expression profiling of end stage OA articular cartilage identified dys-regulation of a large number of genes not previously reported to be associated with the development or progression of OA. A number of genes previously reported as being differentially expressed in human OA articular cartilage were not dys-regulated, which may reflect the sample size, the method of evaluation, the nature of the disease, or species differences between dogs and humans. Understanding the pathological mechanism of

this disabling disease may ultimately allow prevention or better treatment in the future.

Abstract 5

Presented at the Osteoarthritis Research Society International, Seville, 10-11th December 2006

Abstract Published in the Osteoarthritis and Cartilage (2006) **14** (S2), S53

Gene Expression Profiles of Normal and Ruptured Cranial Cruciate Ligament

DN Clements[1,2], JF Innes[1], SD Carter[1], WER Ollier[2], PJ Day[2]

[1] Faculty of Veterinary Science, The University of Liverpool, UK;
[2] Centre for Integrated Genomic Medical Research, The University of Manchester, UK.

Introduction
Anterior cruciate ligament (ACL) rupture is associated with the development of knee osteoarthritis (OA). In the dog, a similar association between cranial cruciate ligament (CCL) rupture and stifle OA is also reported. The aim of the study was to investigate if there were differences in gene expression between normal CCL and ruptured CCL. Gene expression in the normal CCL was also compared between breeds of dog predisposed to CCL rupture and breeds of dog at low risk of CCL rupture.

Methods
Grossly normal CCL was harvested from the healthy stifles of 10 dogs euthanatized for welfare reasons, due to naturally occurring, non-orthopaedic disease (5 from breeds predisposed to CCL rupture and 5 from breeds protected from rupture). CCL remnants were harvested from the stifles of 5 dogs with naturally-occurring CCL rupture undergoing surgical treatment of the condition. mRNA was extracted, amplified, labelled and hybridized to a 44,000 gene canine whole genome oligonucleotide microarray chips. Array scan data were normalised and compared between the different groups. Differential expression of selected genes was confirmed by quantitative PCR (qPCR) using a larger number of CCL samples (21 ruptured, 13 normal [predisposed to CCL rupture], 7 normal [low-risk of CCL rupture]).

Results
When comparing the ruptured CCL to normal CCL, 99 transcripts were significantly up-regulated (including CTSK, COL3, CASP8), and 16 transcripts were significantly down-regulated. qPCR confirmed the increased expression of CASP8, COL3 in the ruptured CCL, and increased expression of additional genes including COL1, MMP2 and IGF1. No significant differences were identified the gene expression profiles of the normal CCL of breeds predisposed to CCL rupture when compared to the normal CCL of breeds at low risk of CCL rupture, either by microarray or qPCR.

Conclusions

Altered transcriptional activity in was identified in the ruptured CCL when compared to the normal CCL. A general pattern of up-regulation of expression of selected proteases and matrix-associated genes characterises the transcriptome of the ruptured CCL. Differences in the risk of CCL rupture between breeds of dog predisposed to CCL rupture when compared to breeds of dog at low risk of CCL rupture could not be related to changes in gene expression in the normal CCL.

Abstract 6

Presented at the British Small Animal Veterinary Association Annual Congress, Birmingham 12th April 2007

Abstract Published in the Proceedings of the British Small Animal Veterinary Association Congress (2007) p438

A Candidate Gene Approach to Canine Osteoarthritis

Clements DN[1,2], Barnes A[1], Short AD[2], Kennedy, LK[2], Ferguson JF[3], Butterworth SJ[4], Fitzpatrick N[5], Pead M[6], Bennett D[7], Innes JF[1], Carter SD[1], Ollier WER[2].

[1]The Musculoskeletal Research Group, Faculty of Veterinary Science, The University of Liverpool, L69 7ZJ.
[2]Centre for Integrated Genomic Medical Research, The University of Manchester, M13 9PT
[3]East Neuk Veterinary Clinic, St Monans, Fife, KY10 2DW
[4]Weighbridge Referral Centre, Swansea, SA6 8QF
[5]Fitzpartick Referrals, Farnham, Surrey, GU10 2DZ
[6]Queen Mother Hospital for Small Animals, The Royal Veterinary College, Hatfield, AL9 7TA
[7]The University of Glasgow Veterinary School, Glasgow, G61 1QH

Introduction

Canine osteoarthritis commonly occurs in association with a primary disease, such as hip dysplasia, elbow dysplasia or cranial cruciate ligament disease. Dog breeds may differ in their susceptibility to both the primary disease and their relative risk for developing osteoarthritis. We evaluated the allele frequencies of single nucleotide polymorphisms (SNPs) in a number of candidate genes, in cohorts of breed-matched dogs with osteoarthritis and a general population of dogs.

Methods

DNA was extracted from whole blood samples from Labrador Retrievers (LR) and Golden Retrievers (GR) surgically treated for elbow dysplasia (LR n = 81), hip dysplasia (LR n = 32), cruciate ligament disease (LR n = 51, GR 45), and a general population vaccinated against Rabies (LR n = 341, GR n = 94). All diseased samples were obtained from the UK DNA Archive for Companion Animals (http://pcwww.liv.ac.uk/DNA_Archive_for_Companion_Animals/). SNPs in six cytokine genes (IL-1, IL-4, IL-6, IL-10, IL-12, TNF-α), two matrix metalloproteinase genes (MMP-3, MMP13), three inhibitors of metalloproteinase genes (TIMP1, TIMP-3, TIMP-4) and Leptin receptor (LEPR) were genotyped using the Sequenom Mass Array platform. Minor allele and genotype frequencies were calculated and compared by Chi square analysis and corrected by Monte Carlo simulation tests.

Results

Significant associations were identified for SNPs in IL-10, MMP13 and TIMP1 and

cruciate ligament disease in Golden Retrievers. Significant associations were identified for SNPs in IL-12 and TNF-α and elbow dysplasia, SNPs in two genes IL-4 and IL-12 and cruciate ligament disease, SNPs in IL-4 and IL-12 and hip dysplasia in Labrador Retrievers. When all component conditions were considered together, significant associations were identified for SNPs in IL-4, IL-12, MMP13 and TNF-α and osteoarthritis in Labrador Retrievers.

Conclusions

SNPs in a number of candidate genes were identified to have significant associations with common canine orthopaedic disorders in two dog breeds. The control population was not phenotyped for disease, thus true associations between SNP allele frequencies and disease may have been considerably stronger. The benefits of multi-centre collaboration and sample collection into a national archive are highlighted by this study.