Genetic Markers of a Predisposition to Lumbar Disc Degeneration in Young Adults

by

Travis Fulton

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved June 2016 by the
Graduate Supervisory Committee:

Juergen Liebig, Chair
Nicholas Theodore
Janet Neisewander

ARIZONA STATE UNIVERSITY

August 2016

ABSTRACT

Intervertebral Disc Degeneration (IVDD) is a complex phenomenon characterizing the desiccation and structural compromise of the primary joint in the human spine. The intervertebral disc (IVD) serves to connect vertebral bodies, cushion shock, and allows for flexion and extension of the vertebral column. Often presenting in the 4th and 5th decade of life as low back pain, this disease was originally believed to be the result of natural "wear and tear" coupled with repetitive mechanical insult, and as such many studies focus on patients between 40 and 50 years of age. Research over the past two decades, however, has demonstrated that environmental factors have only a modest effect on disc degeneration, with genetic influences playing a much more substantial role. Extensive research has focused on this process, though definitive risk factors and a clear pathophysiology have proved elusive. The aim of this study was to assemble a cohort of patients exhibiting definitive signs of degeneration that were well below the average age of presentation with minimal to no exposure with suspected environmental risk factors and conduct a targeted genome analysis in an attempt to elucidate a common genetic component. Through whole genome sequencing and analysis the results corroborated findings in a previous study, as well as demonstrated a potential connection and influence between mutations found in IVD structural or functional genes, and the provocation of IVDD. Though the sample size was limited in scale and age, these findings suggest further IVDD research into the association of variants in collagen, aggrecan and the insulin-like growth factor receptor genes of young patients with an early onset of disc degeneration and minimal exposure to suspected risk factors is merited.

DEDICATION

This thesis is dedicated Liz, without your constant support, patience, and love, I would never have made it through the last two years, much less this project.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Nicholas Theodore, for his constant support and guidance over the past three years, and for the opportunity to work on this project. I would also like to thank Dr. Yashar Kalani for his assistance and input, as well as Dr. Juergen Liebig and Dr. Janet Neisewander for their guidance and support.

Special thanks to the research and administrative staff of BNI, Jill, Julie, and Daytrel for going beyond the call of duty on all too many Saturday mornings.

TABLE OF CONTENTS

Page

LIST OF TABLES……………………………………………………………………vii

CHAPTER

1 INTRODUCTION

1.1 Overview 1

1.2 Structure of the Intervertebral Disc 1

1.2.1 Annulus Fibrosus 2

1.2.2 Nucleus Pulposus 3

1.2.3 Cartilaginous End Plate 5

1.3 Studying Intervertebral Disc Degeneration 6

1.3.1 Healthy vs. Degenerate Discs 7

1.3.2 Clinical Presentation and Identification 13

1.3.3 Previous Studies and Genetic Influence 17

1.4 Hypothesis 21

2 MATERIALS AND METHODS

2.1 Institutional Review Board Amendments 22

2.2 Gene Selection 23

2.3 Patient Selection 23

2.3.1 Inclusion Criteria 24

2.3.2 Exclusion Criteria 24

2.3.3 Final Patient Selection 25

2.4 Collection Preparation 25

CHAPTER Page

2.5 Sample Collection 27

2.6 Annotation and Data Storage 28

2.7 Data Analysis 28

3 RESULTS AND INTERPRETATION

3.1 Collagen 30

3.1.1 Collagen 1 Alpha Subunit 1 30

3.1.2 Collagen 9 Alpha Subunit 2 31

3.1.3 Collagen 9 Alpha Subunit 3 31

3.1.4 Collagen 11 Alpha Subunit 1 32

3.1.5 Collagen 11 Alpha Subunit 2 33

3.2 Interleukins 33

3.2.1 Interleukin 1 33

3.2.2 Interleukin 6 Receptor 33

3.3 Vitamin D Receptor 34

3.4 Insulin-like Growth Factor 1 Receptor 34

3.5 Aggrecan 36

3.6 Matrix Remodeling Enzymes 39

3.6.1 Matrix Metalloproteinase 2 39

3.6.2 Matrix Metalloproteinase 3 39

3.6.3 Matrix Metalloproteinase 9 40

3.6.4 Thrombospondin 40

CHAPTER Page

4 DISCUSSION, FUTURE STUDIES, AND CONCLUSION

4.1 Collagen Findings 41

4.2 Interleukin Findings 42

4.3 Vitamin D Receptor Findings 43

4.4 Insulin-like Growth Factor Receptor Findings 43

4.5 Aggrecan Findings 44

4.6 Matrix Remodeling Enzyme Findings 45

4.7 Limitations and Conclusion 45

BIBLIOGRAPHY 47

LIST OF TABLES

Table Page

1. Modified Pfirrmann Grading Scale....................16
2. Abbreviations Used in Result Tables....................30
3. COL1A1 Variants Found in Patients....................30
4. COL9A2 Variants Found in Patients....................31
5. COL9A3 Variants Found in Patients....................31
6. COL11A1 Variants Found in Patients....................32
7. COL11A2 Variants Found in Patients....................33
8. IL6R Variants Found in Patients....................33
9. VDR Variants Found in Patients....................34
10. IGF1R Variants Found in Patients....................35
11. ACAN Variants Found in Patients....................37
12. MMP3 Variants Found in Patients....................39
13. THBS2 Variants Found in Patients....................40

Chapter 1

INTRODUCTION

1.1 Overview

Intervertebral Disc Disease (IVDD) is a diverse affliction that characterizes a structural failure of the intervertebral discs that provide mechanical support to the vertebral column. Intervertebral discs play a significant role in the support, durability, and flexibility of the spine (Adams et al., 2006), and disorders in relation to disc degeneration are common causes of morbidity and life quality deterioration (Frymoyer et al., 1992). As these discs become mechanically incompetent the spine becomes unstable and loses the ability to properly distribute weight. When this instability occurs in the lumbar discs of the lower back, it is one of the most common causes of low back pain (Benneker et al.), affecting as much as 80% of the population (Luo et al., 2004) with around 10% of patients becoming chronically disabled (Maniadakis et al., 2000). The etiology of IVDD was traditionally viewed as a result of "wear and tear" from mechanical injuries and other factors such as gender, occupation, cigarette smoking and exposure to vehicular vibration (Ala-kokko et al., 2002). However studies over the past 15 years have begun to suggest that degeneration is determined to a large extent, instead, by genetic influences (Sambrook et al., 1999).

1.2 Structure of the Intervertebral Disc

Intervertebral discs (IVDs) are the largest avascular tissues in the body and are positioned between the vertebral bodies acting as the main joints in the spinal column. There are a total of 23 discs throughout the spine, constituting approximately one-

third of the total length. Each disc is 8 to 10mm thick and 4cm in diameter. The discs lie between each vertebrae starting at the third cervical body (C-3) and extending down through the thoracic and lumbar region with the last disc between the final lumbar body (L-5) and the first sacral vertebrae (S-1). As greater pressure is exerted on the base of the spine, the thickness and surface area of the discs is greatest in the lumbar region and thinnest in the cervical levels, they are otherwise remarkably similar in composition and function (Coventry et al., 1945). The primary role of the disc is to evenly transmit loads generated from the body's weight and muscle movements through the vertebral column. The discs also function to create space between each vertebrae allowing for passage of the neural roots of the spinal cord and proper orientation of the facet joints (Pye et al., 2007).The IVD contain a multitude of chondrocyte-like and fibroblast-like cells, each of which is difficult to categorize into a distinct phenotype (Roberts et al., 2002). The discs are composed of three major components, the cartilaginous endplate (CEP), the annulus fibrosus (AF), and the nucleus pulposus (NP).

1.2.1 The Annulus Fibrosus

The outermost morphologically distinct region of the IVD is a series of 15-25 concentric rings, or lamellae, termed the annulus fibrosus. The rigid ring structure of the AF encompasses the NP allowing for the even distribution of pressure to neighboring vertebrae. The primary function of the AF is to withstand the tension created during deformation of the NP (Whatley et al. 2012). The exterior of the AF is firmly affixed to the CEP and contains the few sensory nerves of the IVD (Adams et al., 2006). The AF is composed of type I and II collagen and can be separated into

inner and outer sections. The inner AF must resist high hydrostatic pressure from the shifting of the NP under pressure and is composed of widely spaced lamellae (Antoniou et al., 1996). The outer lamella is subjected to more tensile stress and thus the layers are more closely packed and are highly organized (Shankar et al., 2009). While the two sections are composed of both type I and II collagen, the proportion of type I collagen increases from the inner part toward the outer annulus, while type II follows the opposite distribution (Kraemer et al., 1995).

The collagen fibers of the AF are oriented parallel to lamellae and at 60 degree angles to the vertical axis (Figure 2) (Raj et al., 2008). The orientation of the angles to the vertical axis alternates to the right and left with each successive layer, with the outermost layer highly organized at 60 degrees and gradually reducing to 45 degrees in the inner layers (Cassidy et al., 1989), providing optimal tensile strength for containment of the NP (Walker et al., 2004). Each lamella is bound to adjacent layers via radially extending elastin fibers that also contribute to the ability of the AF to return to its original arrangement following flexion or extension of the spine. The cells of the outer AF are fibroblast-like, elongated, thin, and aligned parallel to the collagen fibers, while those of the inner lamellae are also thin and elongated, but strictly chondrocyte-like (Urban et al., 2003). The dry weight of the AF consists of 50-70% collagen, 10-20% proteoglycans, and 10-15% noncollagenous proteins (Bruehlmann et al., 2002).

1.2.2 The Nucleus Pulposus

The nucleus pulposus is the gelatinous structure contained within the AF and sandwiched between the superior and inferior CEP. The NP resists and evenly

redistributes the compressive forces transmitted through the spine. The vast majority of the NP is composed of water, constituting 75-90% of the wet weight (Horner et al., 2002), this allows for the compressive loads to be transmitted hydrostatically throughout the interior of the AF and internal regions of the CEP, substantially reducing pressure on the AF and CEP interface (Figure 3). As the consistency of the NP is gel-like, it is much less dense than the AF. These two properties give the IVD its viscoelasticity and compressive strength which exceeds that of the neighboring osseous end plates (Walker et al., 2004).

The NP is composed of a large amount of proteoglycans, collagen, elastin, and chondrocyte-like cells (Shankar et al., 2009). Randomly organized collagen fibers and radially extending elastin fibers, comprising 25% and 2% of the dry weight respectively and hold the proteoglycan containing gel in place (Zhao et al., 2007). The chief proteoglycan of the NP is aggrecan, which aggregates to with hyaluronic acid, and contains the glycosaminoglycans (GAGs) keratin sulfate (KS) and chondroitin sulfate (CS) (McDevitt et al., 1988). Due to the large negative charge of the GAG chains in the proteoglycan molecules, water is bound within the NP allowing for a higher level of hydration than would otherwise be osmotically favorable (Yu et al., 2007). The cells of the NP are chondrocyte-like with larger cytoplasmic volume and more rounded than those found in the inner AF (Lee et al., 2007). The primary collagen found in the NP is type II, with the largest concentration at the center and being replaced with type I toward the periphery (Cassinelli et al., 2000), in a similar manner to that of the AF. While aggrecan makes up the majority of the proteoglycans found throughout the NP and makes the largest contribution to

hydration, other proteoglycans are still essential as they are involved in matrix repair (Shankar et al., 2009).

1.2.3 The Cartilaginous Endplate

The most superior and inferior regions of the disc are a thin 1mm thick layer of horizontal hyaline cartilage known as the cartilaginous endplates. The CEP is the interface between the osseous vertebral endplate of the vertebral bodies, the caudal and cephalic aspects of the AF, and the NP, serving as the gateway for nutrient and waste exchange with the disc (Moon et al., 2013). Early in life the CEP is highly vascularized with significant reductions being seen one year after birth, and almost no blood vessels being present by age 30 (Raj, 2008). As these vascular channels decrease the NP begins to rely on simple diffusion across the CEP to supply nutrients and export waste, leading to increasingly inefficient exchange. Nutrient exchange must then occur through capillary buds of the neighboring vertebrae which pass through small marrow contact channels that penetrate the calcified layer of the vertebral osseous endplate, but do not enter the CEP (Nachemson et al., 1970). This process results in poor nutrient exchange and is highly implicated in disc degeneration (Martin et al., 2002).

Another important function of the CEP is to distribute intradiscal pressures to adjacent vertebrae, thus preventing the NP from bulging into the underlying trabecular bone (Brinckmann et al., 1983). As the spine carries substantial forces and the discs are not vascularized, the CEP must balance the opposing biophysical requirements of being strong enough to prevent vertebral fracture, yet porous enough to allow diffusion from vertebral capillaries to the IVD cells (Lotz et al., 2013). The

CEP is attached to the AF by collagen fibers that run horizontal and parallel to the vertebral body and extend into the disc, fixing the structure in place. Like most articular cartilage, the CEP is composed of chondrocytes interspersed with proteoglycans and type I and II collagen fibers (Roughley et al., 2006).

1.3 Studying Intervertebral Disc Degeneration

There is no clear and widely accepted definition for disc degeneration (Freemont et al., 2009). This arises, in part, from the various ways in which disc degeneration is observed. To a radiologist loss of signal intensity or osteophyte proliferation can be diagnostic, for a biochemist changes in hydration or aggrecan content is indicative, and then to a pathologist the presence of fissures or a prolapse demonstrate degeneration. As degenerated disc and aging disc share many similar features, and it is difficult to differentiate changes that occur exclusively due to aging and those that are pathological (Urban et al., 2003) Adams et al. proposed a definition, suggesting that disc degeneration is “an aberrant, cell-mediated response to progressive structural failure” and defining a degenerate disc as “one with structural failure combined with accelerated or advanced signs of aging” (Adams et al., 2006). A universally accepted etiology of IVDD has proved equally difficult to establish. Studies implicate several risk factors for developing IVDD including; a genetic predisposition, impaired metabolite transport, mechanical strain, nutrition, tobacco use, alterations of matrix degradation / synthesis, inflammation, cell senescence, an inactive lifestyle, weightlifting, and aging (Raj et al., 2008; McNally et al., 1992; Battie et al., 1995; Battie el at., 1991; Pope et al., 1998).

In addition to a lack of a fully accepted definition of IVDD, studies aiming to elucidate a disease pathway are further confounded by the lack of an appropriate animal model organism (Urban et al., 2003). The primary difference between animal and human spines is the horizontal orientation of most animal spines, resulting in substantially different weight distribution. In rodents and most mammals, the NP is populated by notochordal cells throughout their life, while these cells disappear in humans after infancy. It has been demonstrated that in these animals the cells produce and maintain a high degree of sulphated GAGs and low collagen content, distinguishing them from IVDs found in humans (Butler et al., 1989). Furthermore in humans the CEP acts as a growth plate for the vertebral body, where in most animals a separate growth plate exists and the CEP is substantially thinner (Urban et al., 2003). Consequently, most information on the etiology of IVDD to date has come from human studies.

1.3.1 Healthy vs. Degenerate Discs

There are substantial changes in the morphology of the human IVD with both aging and degeneration. Disc degeneration involves both cell-mediated changes as well as structural disruptions, but which precedes the other is not clear (Modic et al., 2007). The three factors believed to play a role in IVDD are mechanical, nutritional, and genetic, each to a various extent (Modic et al., 2007). During degeneration the NP experiences changes in cell composition, hydration, and volume, becoming more fibrotic, and the boundary between the NP and AF becoming less distinct (Buckwalter et al., 1995). The structure and organization of the AF also experiences changes, with the lamellae becoming irregular, bifurcating and interdigitating, and the

collagen and elastin networks becoming increasingly disorganized (Urban et al., 2003). The CEP of the IVD becomes sclerotic, loses vascular contact and exhibits decreased permeability (Buckwalter et al., 1995).

As the largest mostly avascular tissue in the body, nutrient transport has long been suggested as a causative factor of IVDD (Nachemson et al., 1970). The healthy adult disc has few, if any, blood vessels, all of which are restricted to the periphery of the outer annulus. These vessels, in conjunction with the capillary beds of the surrounding soft tissue, provide the AF with sufficient nutrition (Horner et al., 2001). However, in the NP, metabolite transport relies on the capillaries of the vertebral body. These molecules must utilize simple diffusion to cross the CEP, and then penetrate the dense disc matrix, with cells in the center of the NP distanced up to 8mm from the closest vessel (Katz et al., 1986). Although transport across the CEP relies on the concentration gradient generated from disc cell metabolism, the viscoelastic properties of the AF allow for some distribution throughout the disc by convective fluid flow as a result of mechanical disc compression and recovery (Lotz et al., 2013). Cells form only about 1% of the entire volume of the IVD (Le Maitre et al., 2007), thus when unimpeded, this supplies adequate exchange for the metabolic needs of the disc (Holm et al., 1981). Although constituting the vast minority of the disc, the roles of IVD cells are vital to proper function, as they maintain and produce the components of the extracellular matrix (ECM) and control matrix synthesis and degradation.

The ability of the IVD to tolerate mechanical loads and provide stability to the spine depends a great deal on the integrity of the ECM. This requires a homeostasis

that balances the degradation of compromised macromolecules with the replacement of newly synthesized products (Buckwalter et al., 2005). Efficient turnover and remodeling of the matrix is accomplished by the action of matrix metalloproteinase (MMPs), a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS), and their regulators, the tissue inhibitor of metalloproteinases-1,2, and 3 (TIMPs 1-3), produced by the cells of the IVD (Le Maitre et al., 2007). By affecting this balance, cells of the IVD are able to adaptively remodel matrix properties in response to mechanical demands (Adams et al., 1997). Perturbation of this equilibrium leads structural compromise and loss of disc hydration (Urban et al., 2003). Consequently, the health of the disc is dependent on the ability of the cell to produce a matrix that is appropriate for the mechanical loads routinely encountered (Urban et al., 2000)

Adequate nutrition and metabolite exchange of IVD cells has a profound effect on both the viability of the disc cells and their ability to produce ECM proteins (Horner et al., 2001). As oxygen supply is low in the avascular disc, the cells rely primarily on glycolysis and fermentation for energy, creating higher than physiologically optimal concentrations of lactate and thus a lower than optimal pH (Diamant et al., 1968). Although unfavorable, under normal conditions the discs are able to maintain a sufficient cell density that, while low in comparison to other tissues, is capable of meeting the minimum requirements to sustain viability (Stockwell et al., 1971). However, this mechanism is insufficient to support matrix upkeep when the disc is subjected additional stressors (Ohshima et al., 1992). This results in cell death in

response to either increased cell demand or a decrease in nutrient supply (Horner et al., 2001).

As discs age, the permeability of the CEP is reduced compromising an already precarious exchange between the vertebral vessels and the IVD. The decreased permeability is the result of sclerosis, endplate chondrocyte apoptosis and calcification, the latter of which making the CEP completely impermeable (Nachemson et al., 1970). This has several deleterious effects on disc viability. The inability to efficiently remove glycolytic and fermentative byproducts, such as lactate, causes oxidative stress which leads to stress-induced premature senescence (Martin 2004) and apoptosis (Sambrook et al., 1999), further compromising the IVD. As nutrient supply is reduced a state of low oxygen and pH leads to decreased proteoglycan synthesis by the remaining viable cells (Ishihara et al., 1999). Matrix degradation also results in the release of several soluble factors and inflammatory cytokines (Gruber et al., 1998). Inflammatory cytokines establish a positive feedback loop, inhibiting the production of ECM and stimulating the production of degradative enzymes (Moon et al., 2014). The failure to maintain the ECM is a significant contributor to disc desiccation (Walker et al., 2004) and consequently, IVDD.

Along with nutritional stressors, mechanical stress also influences ECM maintenance and adaptation. As previously mentioned compression of the IVD facilitate a mild degree fluid exchange, however this pressure also plays an important role in matrix turnover. Mild Compressive loads stimulate MMPs and nitric oxide to degrade the ECM in preparation for new synthesis, as well as stimulating matrix synthesis (Neidlinger-Wilke 2006), and appropriate long-term compressive loading

has been shown to increase PG content (Lotz et al., 1998) leading to improved disc hydration. Excessive prolonged over-loading however inhibits anabolic gene expression while over stimulating MMP production (Neidlinger-Wilke 2006). It is of interest to note that, while tenuous link between heavy physical work and occupational lifting and degeneration has been established (Heliovaara et al., 1989; Deyo et al., 1989; Pope et al., 2002), a study of elite athletes that spent a mean of 26 years weightlifting with “maximal loads”, did not find a strong correlation with IVDD (Matsui et al., 1992), suggesting that perhaps the manner in which the weight is loaded, plays a larger role than the loading of weight itself.

Torsional stress of the spine is primarily resisted by the apophyseal joint formed by the facets. The annulus can resist axial rotation of the individual vertebra up to 3 degrees before collagen fibers begin to tear (Grignon et al., 2000). The facets restrict movement to 1.5 – 3 degrees, consequently damage can only occur from axial rotation when the spine is flexed, separating the apophyseal joint, and aggressively rotated (Adams et al., 1983). Damage from torsional rotation can present as either tears of the outer AF or as radial fissures emerging from the inner annulus (Aaro, 1996).

As degeneration progresses the NP experiences a reduction in both the size and number of PGs, resulting in the loss of adequate hydration. The strict organization of the NP and AF fibers also begins to become disorganized (Adams et al., 2006). This change is accompanied by transition of the NP consistency from gelatinous to more fibrous as cell phenotype shifts from chondrocyte-like to fibrocyte-like (Pattappa et al., 2012) and type II collagen is replaced with the denser type I collagen, which cross

link and hamper exchange of nutrients and waste (Shankar et al., 2009). This shift substantially impedes the ability of the NP to hydrostatically distribute compressive loads across the CEP and the AF.

Inappropriately distributed pressure causes the AF to crack and annular fissures begin to form. An annular fissure is a separation between annular fibers from their vertebral body insertions, or breaks through the fibers that extend through the concentric lamellae (Modic et al., 2007). Annular tears present in one of three categories, radial fissures, circumferential tears, and periphery tears (Figure 4) (Adams et al., 2004). Following injury, the IVD has a very limited capacity heal and sufficiently large tears must be replaced with granulation tissue, which lacks the tensile strength of collagen (McNally et al., 1992).

Structural disruption of the disc or endplate always leads to cell-mediated degenerative changes (Adams et al., 2006). Small tears of the outer AF are adequately repaired by the mechanically inferior granulation tissue, the sparse cell population, however, is not sufficient to replace any structural compromise (Colombini et al., 2008). The inflammatory cascade that follows disc injury recruits factors that induce matrix degradation, angiogenesis, and innervation (Moon et al., 2014).Annular fissures and neurovascular ingrowth into degenerative discs are a constant finding of histologic studies (Hunter et al., 2003). In the healthy IVD aggrecan inhibits endothelial growth (Johnson et al., 2002), thus as PG content declines in the degenerative disc, the loss of hydration is coupled to the loss of inhibition. Upon compromise of the superior aspect of the CEP, the CEP/AF interface, or of the outer annulus, the invasion of nerves and blood vessels results in addition deterioration of

IVD structural integrity and functionality. When the annulus is violated or scarred, the ability of the AF to contain the NP is drastically reduced (Bron et al., 2009), causing inappropriate pressure distribution throughout the annulus. Once the annulus is sufficiently compromised the rate progressive damage from inappropriate loading exceeds the rate of annular repair (Mengoni et al., 2015). Although increased vasculature supplies the deteriorating disc with increased nutrition, the coinciding structural compromise far exceeds the nutritional benefit. Thus injuries to the inner annulus or CEP that decompress the NP or moderate damage to the outer annulus leads to compounding damage and healing processes are over taken by degenerative changes (Hadjipavlou et al., 2008).

1.3.2 Clinical Presentation and Identification

IVDD is most often observed in patients between the 4th and 5th decade of life, and may present clinically as lower back pain (LBP) or sciatica. These symptoms are commonly accompanied by a radiological finding of disc prolapse, spinal stenosis, facet joint arthropathy or a combination of each (Saleem et al., 2013). In the absence of imaging IVDD can be difficult to recognize, as there is a high prevalence of asymptomatic patients with radiological signs of degeneration (Jensen et al., 1994). When symptoms are present, they are often diverse and can range from the most common manifestation of pain, to debilitating neuropathy (Modic et al., 2007).

The cause of IVDD induced LBP is a subject of much debate, but there is evidence to suggest that, neural ingrowth, mechanical nerve compression, and inflammation are likely culprits. Nerves invading past the outer annulus form free nerve endings within the AF lamellae that are subsequently compressed with flexion

and extension of the spine (Brisby et al., 2006) and resulting in what is commonly referred to as discogenic pain. When the annulus is fully compromised the NP herniates into the spinal canal, termed a herniated nucleus pulposus (HNP) and may cause pain by directly compressing the spinal cord or dorsal root ganglia. This compression can result in unilateral or bilateral leg pain that radiates down the leg and into the feet or toes, and can cause numbness in the dermatomes corresponding to the compression, also known as sciatica (Saleem et al., 2013). However, while 90% of sciatica cases are associated with the presence of an HNP, as few as 30% of HNPs are symptomatic (Jordan et al., 2011). When nuclear material leaks out of the AF it has also been demonstrated to cause pain by inducing nerve degeneration, increasing nerve fiber discharge, and attracting inflammatory cells (Brisby et al., 2006). The inflammatory response has been suggested to sensitize nerve ending both in and around the IVD leading perception of pain with only mild stimulation (Goupille et al., 1998).

Neuropathy also occurs with compression of the spinal nerves, presenting as peripheral neuropathy, myelopathy, or radiculopathy. Peripheral neuropathy associated with lumbar IVDD will often present as a "pins and needles" sensation or as tremor (Kashyap et al., 2004). When nerve compression is severe enough, muscle weakness can be observed as gait abnormalities or "foot drop", the latter of which is characterized as an inhibition in dorsiflexion of the ankle or foot (Liu et al., 2013). Myelopathy and radiculopathy present in a similar fashion, both can result in, pain, muscle weakness, fasciculation and sensory deficits of the dermatomes innervated by the compressed nerves. However, myelopathy refers to the direct compression of the

spinal cord and tends to occur more gradually, where radiculopathy refers to the compression of the nerve root. As the spinal cord variably terminates between T12 and L3 into the cauda equina, a collection of nerve roots, and most lumbar disc herniations occur between L4 and S1, radiculopathy is substantially more common in lumber herniations.

Disc prolapse, is a common findings in IVDD however there are additional signs of degeneration, consisting of diminished disc height, narrowing of facet, spondylophytes in the upper and lower endplates, stenosis of spinal canal, narrowing of lateral recess, real or apparent desiccation, fibrosis, extensive fissuring, mucinous degradation of annulus, defects and sclerosis of the end plates, osteophytes at the vertebral apophyses, loss of T2 signal, calcification vacuum changes, misalignment, and marrow changes (Nachemson et al., 1970) . As there are no definitive serum tests for biochemical markers of IVDD, the only ethically acceptable method of evaluating disc degeneration in a live population with a high degree of efficacy, is imaging (An et al., 2004). While X-rays can demonstrate osteophyte formation and CEP fractures, the preferred diagnostic tool for imaging and diagnosis of IVDD is magnetic resonance imaging (MRI) (Taher et al., 2012).

The radiographically observable alterations that occur during disc degeneration have been extensively described, but these changes must still be quantified for research purposes (Resnick et al., 1985). As many as 42 distinct grading systems for disc degeneration are described in literature, characterizing morphologic, histologic, and radiographic findings, most of which are based on subjective terminology (Lavy et al., 2010). 22 of these grading systems are specific to lumbar disc degeneration, 5

of which use MRIs. Of these five systems, only one, the Pfirrmann grading system, demonstrated substantial to almost perfect interobserver agreement (Lavy et al., 2010). The Pfirrmann system relies on T2 signal intensity, distinction of NP and AF, and disc space narrowing to grade disc degeneration on a scale of I through V, with V being the most severely degenerated discs (Pfirrmann et al., 2001). T2 signal intensity is a function of proteoglycan content (Pearce et al., 1991) and thus correlates with the progression of IVDD (Modic et al., 1988). This system was associated with a small degree of ambiguity and furter modified in 2007 by Griffith et al. and presented as the modified Pfirrmann grading system (Table 1)

Table 1. Modified Pfirrmann Grading System (Griffith et al. 2007)

Grade	Signal from nucleus and inner fibers of annulus	Inner and outer annulus	Height of disc
1	Uniform	Distinct	Normal
2	> Presacral Fat	Distinct	Normal
3	< Presacral Fat	Distinct	Normal
4	< Outer Fiber of Annulus	Indistinct	Normal
5	= Outer Annulus	Indistinct	Normal
6	= Outer Annulus	Indistinct	<30% Reduction
7	= Outer Annulus	Indistinct	30% to 60% Reduction
8	= Outer Annulus	Indistinct	>60% Reduction

The most common disc disorder that presents to spinal surgeons is a lumbar HNP (Urban et al., 2003). Disc herniation is defined as a “localized or focal displacement of disc material beyond the limits of the intervertebral space” (Schroeder et al., 2016), with the disc space margins defined superiorly and inferiorly by the borders of each CEP and peripherally by the vertebral ring apophyses. HNPs can be sub divided into three categories, a protrusion which occurs when an intact AF bulges into the spinal

canal, an extrusion, when the contiguous NP violates the outer AF, and a sequestration, when the NP penetrates the AF and is no longer continuous, with a distinctly separate fragment (Spengler et al., 1990). The mechanical force necessary to induce an HNP in a healthy IVD is substantially greater than what is normally encountered, with most vertebral bodies failing before the IVD herniates in laboratory tests (Adams et al., 1982). Thus some degree of degeneration is necessary before a disc can herniate (Moore et al., 1996) in the absence of moderate trauma.

1.3.3 Previous Studies and Genetic Influence

There is a substantial lack of epidemiological and population based data on disc degeneration (Hassett et al., 2003). Studies attempting to elucidate the environmental factors of IVDD are summarized well in a 1992 review, as focusing on gender, age, occupation, smoking and vehicular vibration (Frymoyer et al., 1992), with degeneration occurring primarily due to the accumulation of these factors in conjunction with mechanical injuries imposed on normal aging processes (Battie et al., 2004). However two decades later it is by and large believed that although there are certainly environmental elements, these factors seem to exert only minor effects on the development and progression of the disease when compared to genetic influences (Ala-kokko et al., 2002). Non-genetic risk factors still believed to exacerbate IVDD seem to be, an increased body mass index (BMI), age, diabetes, hyperlipidemia, mechanical insults, and smoking (Schroeder et al., 2016). One of the first major initiatives suggesting a genetic component to IVDD was the "Finnish Twin Study" which began in 1991. A cohort of 147 monozygotic and 153 dizygotic twins was recruited based on discordance of environmental and behavioral risk

factors believed to play a role at the time, and participated in a series of investigations (Battie et al., 2009). The studies found those twins heavily discordant for, driving, physical activity, leisure time, and physical loading, demonstrated almost no appreciable difference in disc degeneration, though cigarette smoking did seem to play a minor role (Battie et al., 2009). Most notable was a finding by Videman et al, demonstrating that lumbar disc degeneration had an almost 50% heritability (Videman et al., 2004).

Genetic influences of IVDD have been of increasing interest since the publication of the Finnish twin findings; with recent studies suggest genetic factors can increase the risk of developing IVDD up to six times that of the general population (Chan et al., 2009).

In over 20 years a myriad of studies have been conducted to try to determine a genetic component to IVDD. There are two common methods of studying genetic influences in IVDD, whole genome association studies (WGAS) and research directed investigation. WGAS involves sequencing the entire genome of pathological patients and control patients then comparing the frequency of alleles, variant mutations of the same gene, common to the pathological group with those of the control group. When an allele is found to occur at a statistically different frequency (Allele Frequency) to that of the control group it is considered associated with the pathology. This method casts a wide net, requires extensive analysis, and is generally costly, but can elucidate genes and pathways that were not previously suspected. Research directed investigations vary in methodology but follow three basic steps; identify pathology specific-genes through a literature review, select pathological

patients and control patients, then sequence patients for suspected genes. This method is fairly more elegant, with less data being analyzed and cost effective as only genes of interest are sequenced and investigated.

Most studies previously focused on determining IVDD candidate genes by examining associations of patients exhibiting signs of disc degeneration and mutations of exons affecting a single nucleotide, known as a single nucleotide polymorphism (SNP) (Videman et al., 2009). SNPs are alterations, as the name suggests, of a single nucleotide in a gene resulting in an alteration of the original sequence. A transition occurs when a purine (adenine [A] or guanine [G]) or a pyrimidine (cytosine [C] or thymine [T]) is replaced by a nucleotide of the same class (i.e. A to G or C to T). Transversions occur when a base is substituted with the base of a different class (i.e. C to G or A to T). An insertion or deletion occurs when a nucleotide is added or deleted and consequently alters the reading frame of the DNA, known as a frameshift. Frameshift mutations generally destroy the functionality of the translated protein. An SNP can result in three types of mutations; silent mutations, nonsense mutations, or missense mutations. A Silent mutation occurs when an alteration of a codon, the three nucleotide sequence that encodes an amino acid (AA), results in the production of an AA synonymous to that of the original codon. A nonsense mutation results in the altering of an AA codon to that of stop codon, resulting in a premature termination of translation. Finally a missense mutation is an alteration of a codon such that translation produces an AA that is not synonymous to the original codon.

These mutations can occur in exons, the expressed regions of a gene, introns, the intergenic regions, or in intergenic DNA. Introns are regions of DNA spaced between the exons, are removed prior to translation, and vary widely across individuals of the same species. Exons are the region of a gene that is ultimately converted to a mature mRNA. Exons contain a 5' untranslated region (5'UTR), coding region (CDS), and a 3' untranslated region (3'UTR). The 5' UTR plays a role in translation regulation, the CDS is the region that is translated into a protein, and the 3'UTR contains regulatory regions that function in post translational expression modification as well as stability and localization of the mRNA (Szostak et al., 2012). MicroRNA (MiRNA) are small noncoding RNA molecules that play a role in gene silencing. Most animal MiRNA binds to a region found in the 3'UTR (Wang et al., 122) and represent an additional form of 3'UTR posttranslational regulation. When a novel mutant allele is discovered it is reported to the database of SNPs (dbSNP) and assigned a Reference SNP cluster ID (rsNumber).

Genes that have most often been implicated in IVDD are classified into one of three categories; genes involved with construction and structure of the IVD, genes that produce and regulate degrading enzymes, and genes associated with bone structure. Collagen, a major structural component of the IVD, has several SNPs commonly found in patients with IVDD (Plujim et al., 2004; Karppinen et al., 2002). Mutations in several inflammatory genes have also shown an association with IVDD (Solovieva et al., 2004; Noponen-Hietala et al., 2005) as well as the MMPs involved in matrix turnover and repair (Takahashi et al., 2001; Dong et al., 2007; Sun et al., 2009; Hirose et al., 2008). The chief hydrating component of the NP, aggrecan

(Kawaguchi et al., 2008; Li et al., 2013), and the insulin-like growth factor receptor (IGF1) (Urano et al., 2008) are also implicated. Finally the vitamin D receptor (VDR) has been of interest in several genetic association studies (Kawaguchi et al., 2002; Cheung et al., 2006).

1.4 Hypothesis

IVDD is a complex disease that is characterized by advanced signs of disc aging and is suggested to have a genetic component. An HNP in the absence of trauma represents a clear indication of disc degeneration. We hypothesize those young adults who have experienced an atraumatic lumbar HNP in the absence of common risk factors represent the ideal population for genetic analysis in an attempt to identify genetic components of IVDD. Further these patients with defects of structural or functional genes, when subjected to minor trauma or activation of an inflammatory cascade, will display early manifestation of IVDD.

Chapter 2

Materials and Methods

All work in this section represents the efforts conducted by the graduate student. Work was conducted originally autonomously by the graduate student, then periodic meetings with the supervisor were held to confirm validity and applicability of the work. The work executed by the graduate student then periodically reviewed is referred to as work completed by “the graduate student”. Contributions of other bodies are referred to as “administrative staff” and include any efforts that were related to logistic support that was not directly connected to the investigation but was necessary to fulfill logistic or legal requirements and was carried out by hospital professional staff, or “SII”, which refers to the collective efforts of all individuals employed by System Imaginations Inc. who were involved in the project.

2.1 Institutional Review Board Amendments

An outdated Institutional Review Board (IRB) for Human Research application had previously been submitted. The IRB was modified by the graduate student to reflect the current aims of the study and submitted by hospital administrative staff. After revision the IRB was accepted and the investigation commenced. Three additional IRB amendments were made through the course of the study to accommodate a change from exome sequencing to whole genome sequencing, an increase in the volume of blood to be drawn and the inclusion of patients under the age of 18 in the study.

2.2 Gene Selection

An exhaustive literature review was conducted be the graduate student to select genes for investigation (Urban et al., 2003; Hadjipavlou et al., 2008; Colobini et al., 2008; Videman et al., 2004; Battie et al., 2009; Chan et al., 2006; Videman et al., 2009; Plujim et al., 2004; Karppinen et al., 2002; Solovieva et al., 2004; Takahashi et al., 2001; Dong et al., 2007; Sun et al., 2009; Hirose et al., 2008; Kawaguchi et al., 1999; Urano et al., 2008; Noponen-Hietala et al., 2005; Kawaguchi et al., 2002; Cheung et al., 2002; Roughley et al., 2006; Aigner et al., 1998; Roberts et al., 1998; Tilkeridis et al., 2005; Paassilta et al., 2001; Futoshi et al., 2007, Zhang et al., 2008; Kalichman et al., 2008; Mayer et al., 2013; Eskola et al., 2012; Betram et al., 2006; Chen et al., 2013; Han et al., 2013; Gologorsky et al., 2014; Cheung et al., 2010; Nerlich et al., 2005). Genes were investigated based on potential association with IVD structural components and inflammatory pathways. Genes were selected based on representation in literature and level of previous investigation. After review COL1A1, COL9A2, COL9A3, COL11A1, COL11A2, IGF1R, ACAN, IL1, IL6R, MMP2, MMP3, MMP9, THBS2 and VDR were selected further investigation in the study.

2.3 Patient Selection

All subjects were selected by the graduate student from a list of patients who had undergone a microdiscectomy surgical resection at the Phoenix St. Joseph's branch of the Barrow Neurological Institute (BNI). The patients had a positive diagnosis of lumbar disc herniation both pre and post operatively and were graded above grade II on the Pfirrmann grading scale. The mean age of patients at time of surgery was 25

years, with a mean time of 2.3 months between radiological diagnosis and surgery. 10 Patients self-identified as Caucasian/White, 4 as Hispanic/Latino, and 1 as African American. 6 patients had a single herniation between L5-S1, 5 Patients had a single herniation between L4-L5, 3 patients had three herniations between L3-S1, and 1 patient had two herniations between L3-L5.

2.3.1 Inclusion Criteria

An extensive literature review was conducted by the graduate student for risk factors that have been associated with IVDD (Roughley et al., 2006; Hadjipavlou et al., 2008; Aigner et al., 1998; Hassett et al., 2003, Roberts et al., 1998; Videman et al., 2004; Shankar et al., 2009; Raj, 2008, Battie et al., 2009; Schroeder et al., 2016; Adams et al., 2006; Modic et al., 2007), and inclusion criteria were determined. Patients must have undergone a microdiscectomy at BNI and received both a radiological diagnosis and operative diagnosis of lumbar HNP, and have grade II degeneration upon BNI review of MRIs to be considered. Only patients diagnosed with a lumbar HNP under the age of 30 were considered. Only patients that experienced HNP with no or minor trauma, defined as "greater than a sneeze, but less than a fall" where were considered.

2.3.2 Exclusion criteria

Exclusion criteria were determined by the graduate student based on the literature review conducted in section 2.3.1. Any patient that presented with an HNP immediately following a trauma, defined as "Any physical insult that would exceed that experienced on a daily basis", or any patient that experienced back pain within 1 year of trauma was excluded. All patients who participated in club or school,

baseball, swimming, basketball, kendo/martial arts, soccer, or football after graduating high school were excluded. Patients that had a BMI of over 30 and were classified as endomorphic or diagnosed with hyperlipidemia were excluded. Patients that worked jobs that required unsupervised heavy physical lifting (warehouse workers, furniture movers, construction workers etc.) or described their employment as "requiring sporadic physical exertion" were excluded. Any patient that experienced more than 10 hours of driving a week was excluded. Patients who smoked more than 1 pack year, defined as the product of the number of packs of cigarettes smoked per day and the number of years the patient has smoked, were excluded.

2.3.3 Final Patient Selection

1,005 patients presented between 2006 and 2015 with lumbar HNPs that were treated with microdiscectomy. The graduate student determined that 66 patients fulfilled all inclusion requirements, 23 of which were removed for presence of exclusion criteria. The 43 remaining patients were contacted by the graduate student and asked to participate, 39 agreed. Phone interviews were then conducted by the graduate student with the patients to confirm accuracy of inclusion and exclusion criteria, at which point 2 patients were excluded. After a second thorough review of the patient information, all 37 patients were approved by administrative staff for the study. Of these 37, 17 patients were able to participate in the study due to various working or personal constraints.

2.4 Collection Preparation

A third party bioinformatics company, System Imaginations Inc. (SII) was recruited to assist in identification of logistic needs for the study and for analytic

support in future studies. The graduate student working with SII determined that DNA would be collected from blood and saliva samples, protein samples would be collected from blood, and RNA would be collected from blood. Due to a shift in sequencing cost the graduate student and SII determined that whole genome sequencing was fiscally feasible and would be superior to exome sequencing. Various collection vessels were researched and debated between SII and the graduate student. EDTA vacutainers were determined to be the best option for DNA collection and storage of blood specimens (Banfi et al., 2007) per the graduate student, and Oragene Discover tubes were selected for saliva samples per SII suggestion. PAX Gene tubes were selected for collection of RNA blood samples per agreement of efficacy between SII and the graduate student. The P100 Blood Collection System (P100BCS) was selected for blood protein sample collection per agreement of efficacy between SII and the graduate student. Subjective patient data was desired for future studies and an SF36, Visual Analog Scale (VAS), and Oswestry Disability Index questionnaire were selected to collect this information per the graduate student. Morphological data was implicated in future studies and a 34 point measurement system was developed through coordination of SII personnel and the graduate student to account for potential future needs. Due to conflicting schedules and time constraints 17 of the 37 approved patients were able to participate in the study and were scheduled by administrative staff for study collections, however 2 sample were not used due to study timeline incompatibilities.

2.5 Sample Collection

Approved patients participating in study arrived at BNI and provided informed consent via the graduate student and administrative staff. Patients were vetted a third time for any exclusion criteria, no patients were excluded at this stage. Morphological information was collect in accordance with the developed 34 point system on each patient. Patients were taken to the clinical laboratory in the St. Joseph Medical Center and lab personnel extracted samples that were processed by the graduate student. Salvia samples were collected in Oragene tubes through oral sputum throughout the duration of the visit, when a sufficient sample was collected the tube was sealed, homogenized though inversion for 5 second, placed in a freezer with a holding temperature of -20 degrees Celsius ("the -20") for 24 hours, then transferred to a freezer with a holding temperature of -80 degrees Celsius ("the -80"). 3ml of blood was collected in EDTA Vacutainers, homogenized via inversion ten times, and then placed in the -80 for storage. 8.5 ml of blood was collected in P100BCS tubes, homogenized through inversion 10 times, allowed to incubated at room temperature for 2 – 4 hours, centrifuged at 2500g for 20 minutes, plasma supernatant was extracted and placed in cryovials which were placed in the -20 for 24 hours, then moved to the -80 for storage. 2.5ml of blood was collected in PAXgene tubes, homogenized through inversion 10 times, incubated at room temperature for 2 hours, placed in the -20 for 24 hours, and then moved to the -80 for storage. Upon collection of final specimens SII took procession of EDTA samples and handled the subsequent shipping and transport to Macrogen Clinical Labs for sequencing. All other samples were stored in the -80 for future studies. Protein, RNA and morphological data were

not used in the current study but were determined by the graduate student and SII to be of potential value in future studies.

2.6 Annotation and Data Storage

SII was chosen to annotate the sequenced data and store the electronic information. SII developed a graphic user interface, the "Spondylo Viewer", for querying of patients vs. genes with mutations allowing for searches yielding patients with mutations of suspected genes. SII vetted and selected Macrogen Clinical Labs to perform DNA extraction from EDTA stored sample for DNA QC, library prep and DNA sequencing. Upon completion of DNA sequencing, raw data was uploaded to the Spondylo Viewer. DNA integrity and quality was assessed by SII and determined to be of "excellent quality" and no samples were excluded.

2.7 Data Analysis

The American and European subset of The 1000 Genomes Project was selected as the optimal reference genome, as a more diverse population was included than in alternate references such as the GR37 or HG19, and the 2010 pilot paper was one of the most cited articles in biology in 2011 (Science Watch, 2012). The graduate student used the Spondylo Viewer to search selected genes for mutations, defined as an alteration from the reference genomes. Transition, transversion, insertion, and deletion locations were noted and mutation effects were determined by comparing alterations against the relevant reference. Mutation effect, nucleotide alteration, resulting AA change, distance of mutation from beginning of gene, allele frequency, and novelty, were determined and recorded by the graduate student, haplotype was supplied by SII. Mutations of interest were investigated and resulting protein alterations were characterized. A Chi Square test

was applied to determine the statistical relevance of findings. Observed frequencies from the study and expected frequencies supplied by the American subset of the 1000 Genomes Project were used in analysis. Due to the notably small sample size it was determined that a p value of 0.05 would be inappropriate and the more stringent p value of 0.005 was applied (Kenny et al., 2002).

Chapter 3

Results and Interpretation

Table 2. Abbreviations used in result tables.

Mut Effect	Type of mutation
Ref to Alt	Nucleotide change from reference genome
AA Change	Amino acid change
Mut Pos	Position from first nucleotide where mutation occurs
AF	Allele frequency (UKN = Unknown)
RSN	Reference SNP cluster ID reported to dbSNP
Haplo	Haplotype (Homo = Homozygous Hetero = Heterozygous)

3.1 Collagen

3.1.1 Collagen 1 Alpha subunit 1

The variant rs112626388 was found in 86.66% of patients, which is above the expected frequency of 69% ($p < 0.005$). This variant is in the 3'UTR. One patient, 6.6%, expressed the rare variant rs72667032, the expected frequency is 0.43% ($p < 0.005$). This variant occurs in the triple helix repeat, and is a shift from one hydrophobic AA to another. A novel 3' UTR deletion variant was found at position 1040 in 6.6% of patients.

Table 3. COL1A1 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_192	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_196	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_262	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_262	Missense	C to G	P to A	613	0.004	rs72667032	Hetero
P_419	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_445	Deletion	GA to G	3' UTR	1040	UKN	Novel	Hetero
P_531	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_562	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_654	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_659	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_692	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_725	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero

P_752	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_832	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero
P_942	Insertion	G to GTG	3' UTR	681/682	0.69	rs112626388	Hetero

3.1.2 Collagen 9 Alpha Subunit 2

The rare variant rs140041506 appeared in 13.3% of patients, with an expected frequency of only 1% ($p < 0.005$). This variation occurs in the triple helix region and represents a shift from a hydrophobic residue to a hydrophilic residue.

Table 4 COL9A2 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_562	Missense	G to T	P to T	1981	0.01	rs140041506	Hetero
P_692	Missense	G to T	P to T	1981	0.01	rs140041506	Hetero

3.1.3 Collagen 9 Alpha Subunit 3

The rare variant rs61734651 was found in 20.0% of patients, with an expected frequency of only 2.7% ($p < 0.005$). This mutation occurs in the triple helix domain and is the shift from a basic residue to a nonpolar residue. A novel variant was found at position 1625 in 6.6% of patients; this mutation is in the triple helix domain and is the transition from one hydrophobic AA to another.

Table 5 COL9A3 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_262	Missense	C to G	A to G	1625	UKN	Novel	Hetero
P_419	Missense	C to T	R to W	307	0.027	rs61734651	Hetero
P_659	Missense	C to T	R to W	307	0.027	rs61734651	Hetero
P_832	Missense	C to T	R to W	307	0.027	rs61734651	Hetero

3.1.4 Collagen 11 Alpha Subunit 1

The variant rs3753841 was found in 80% of patients with an expected frequency of 75.5% ($p > 0.005$). This variant is in the triple helix region; all patients with this allele were homozygous. The rare variants rs140954784 and rs144562769 were expressed in one patient, 6.6%, with an expected frequency of 0.005% ($p < 0.005$) and 0.002% ($p < 0.005$) respectively; both reside in the triple helix domain with the former being a change between two hydrophobic residues, and the latter a shift from a hydrophobic residue to a hydrophilic one. The variant rs36076089 was found in 20% of patients with an expected frequency of 1% ($p < 0.005$). This variant is found in the splice region and is not part of the splice site palindrome. The extremely rare variant rs143875783 was found in 6.6% of patients, with an expected frequency of 0.0% ($p < 0.005$). This variant is found in the 3' UTR.

Table 6 COL11A1 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_192	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_196	Missense	C to G	P to A	4477	0.0005	rs140954784	Hetero
P_196	Missense	C to A	P to T	3112	0.0002	rs144562769	Hetero
P_262	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_419	Splice	GA to G	SRV	652	0.01	rs36076089	Hetero
P_419	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_531	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_562	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_654	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_659	Transition	G to A	3' UTR	809	0.000	rs143875783	Hetero
P_659	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_692	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_725	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_752	Splice	GA to G	SRV	652	0.01	rs36076089	Hetero
P_752	Missense	C to T	P to L	3851	0.755	rs3753841	Homo

P_832	Missense	C to T	P to L	3851	0.755	rs3753841	Homo
P_942	Splice	GA to G	SRV	652	0.01	rs36076089	Hetero
P_942	Missense	C to T	P to L	3851	0.755	rs3753841	Homo

3.1.5 Collagen 11 Alpha Subunit 2

The rare variant rs2229784 was present in 13.3% of patients, with an expected frequency of 4% ($p < 0.005$). This mutation is in the triple helix region and is a shift from a hydrophobic residue to a hydrophilic residue.

Table 7 COL11A2 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_562	Missense	G to T	P to T	3625	0.05	rs2229784	Hetero
P_659	Missense	G to T	P to T	3625	0.05	rs2229784	Hetero

3.2 Interleukins

3.2.1 Interleukin 1 Alpha

No IL1A variants were found in the patients.

3.2.2 Interleukin 6 Receptor

The common variant rs2228145 was found in 46.6% of patients with an expected frequency of 53% ($p > 0.005$). This variant occurs in the fibronectin type III domain. All patients were homozygous for this allele.

Table 8 IL6R Variants

Patient	Mut Affect	Ref to Alt	AA Change	AA Pos	AF	RSN	Haplo
P_196	Missense	A to C	D to A	1073	0.53	rs2228145	Homo
P_445	Missense	A to C	D to A	1073	0.53	rs2228145	Homo
P_562	Missense	A to C	D to A	1073	0.53	rs2228145	Homo
P_654	Missense	A to C	D to A	1073	0.53	rs2228145	Homo
P_659	Missense	A to C	D to A	1073	0.53	rs2228145	Homo
P_725	Missense	A to C	D to A	1073	0.53	rs2228145	Homo

P_832	Missense	A to C	D to A	1073	0.53	rs2228145	Homo

3.3 Vitamin D Receptor

The Variant rs371312471 and a novel variant at 2102 were both found in two patients. These mutations were 1 nucleotide apart in the 3’ UTR. The rare variant rs9729 was found in 1 patient constituting a frequency of 6.6% with an expected value of 0.06% ($p < 0.005$). This mutation was found in the 3’ UTR.

Table 9 VDR Variants

Patient	Mut Effect	Ref to Alt	NT Change	Mut Pos	AF	RSN	Haplo
P_419	Transversion	A to T	2100	3’UTR	UKN	Novel	Hetero
P_419	Transversion	G to T	2102	3’UTR	UKN	rs371312471	Hetero
P_445	Transversion	A to T	2100	3’UTR	UKN	Novel	Hetero
P_445	Transversion	G to T	2102	3’UTR	UKN	rs371312471	Hetero
P_692	Transversion	C to G	1906	3’UTR	0.006	rs9729	Hetero

3.4 Insulin-like Growth Factor 1 Receptor

All IGF1R variants were found in the 3’ UTR. The common insertion rs3051367 was found in 66% of patients with an expected frequency of 44% ($p < 0.005$). The fairly common variant rs35715134 was found in 93% of patients and expected in 49% ($p < 0.005$). The deletion rs58523117 common to the American population was found in 93% of patients and expected in 85% ($p > 0.005$). Variant rs375097523 was expressed in 6.6% of patients and has previously been reported, but no population data is available. Mutation rs3051365 was found in 13% of patients with a population frequency of 1% ($p < 0.005$). 1 Patient expressed rs3833014 representing 6.6%% to an expected 8% ($p > 0.005$) of the cohort. 4 novel variants were discovered; a transversion at position 2291, a

deletion at 5771, a transition at 3603, and an insertion at 5741. All patients were heterozygous for variants.

Table 10 IGF1R Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_192	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_192	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_192	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_192	Insertion	C to CAT	3' UTR	1753/1754	0.01	3051365	Hetero
P_196	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_196	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_196	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_262	Transition	G to A	3' UTR	3443	UKN	375097523	Hetero
P_262	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_419	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_419	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_445	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_445	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_531	Transvers	C to A	3' UTR	2291	UKN	Novel	Hetero
P_531	Insertion	G to GC	3' UTR	3465	0.08	3833014	Hetero
P_531	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_531	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_531	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_531	Deletion	CA to C	3' UTR	5771	UKN	Novel	Hetero
P_562	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_562	Insertion	T to TC	3' UTR	5741	UKN	Novel	Hetero
P_624	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_624	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_624	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_654	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_654	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_654	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_659	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_659	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_659	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_692	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_692	Insertion	C to CAT	3' UTR	1753/1754	0.01	3051365	Hetero
P_692	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero

P_692	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_725	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_725	Transition	G to A	3' UTR	3603	UKN	Novel	Hetero
P_725	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_725	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_752	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_752	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_752	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_832	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_832	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero
P_832	Insertion	C to CTT	3' UTR	2851/2852	0.44	3051367	Hetero
P_942	Deletion	GGT to G	3' UTR	2529/2530	0.86	58523117	Hetero
P_942	Deletion	GA to G	3' UTR	1259	0.49	35715134	Hetero

3.5 Aggrecan

The variant rs35430524 was found in 1 patient making up 7% of an expected 15% ($p > 0.005$) of the study group. This variant effects the chondroitin sulfate I (CS-I) domain, replacing a hydrophobic AA with a polar residue. The variant rs12899191 was previously recorded but no population data is available, this allele was found in 67% of patients. The variant represent a shift from a polar to a non-polar AA and occurring in the CS-I domain. The uncommon allele rs938608 was found in 73% of patients with an expected frequency of 40% ($p < 0.005$). This mutation is a change from a polar to a non-polar residue and occurs in the CS-I region. The common allele rs3817428 was present in 47% of patients while expected in 87% ($p < 0.005$). This is a switch between two acidic AAs and is present in globular structural domain 3 (G3). The uncommon allele rs2882676 was displayed in 67% of patients with an expected frequency of 38% ($p < 0.005$). This mutation is a change from an acidic polar AA to a hydrophobic AA. Rs2882676 affects the chondroitin sulfate II (CS-II) domain. The uncommon allele rs938609 was present in 73% of patients and has a frequency of 38% ($p < 0.005$). This variant interchanges

between two polar residues, and resides in the CS-I domain. The rare allele rs117116488 was present in 13% of patients with an expected frequency of 0.5% ($p < 0.005$). This variant changes a polar residue to a non-polar residue. This variant is in the globular structural domain II B (G2B). The rare allele rs28559926 was found in 20% of patients with a population frequency of 5% ($p < 0.005$). This mutation results in an exchange between two acidic AAs, and is part of the CS-I domain. The uncommon variant rs3743398 was represented in 33% of patients and expected in 11% ($p < 0.005$). This variant changes between two nonpolar residues, and effects the CS-I region. The rare variant rs181736584 was found in 13% of patients and expected in 0.1% ($p < 0.005$). This is an alteration between two nonpolar bases and is part of the aggrecan core protein. The uncommon allele rs35546357 was present in 1 patient constituting 6.6%% of patients and was expect to account for 2% ($p > 0.005$). This variant is an alteration between a polar and nonpolar residue, and is in the CS-I domain. The common variant rs11638262 was expressed in 6.6%% of the cohort, and expected in 49% ($p < 0.005$). This mutation is a shift from a hydrophobic to a hydrophilic AA, and is in the CS-I domain.

Table 11 ACAN Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_192	Missense	C to A	P to T	2737	0.15	rs35430524	Hetero
P_192	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_196	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_196	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_196	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_196	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_196	Missense	C to T	S to L	1469	0.005	rs117116488	Hetero
P_196	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_196	Missense	G to C	E to D	4227	0.05	rs28559926	Hetero
P_196	Missense	C to T	P to L	2591	0.11	rs3743398	Hetero

P_262	Missense	C to T	P to L	2591	0.11	rs3743398	Hetero
P_262	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_262	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_262	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_262	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_262	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_262	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_419	Missense	C to T	P to L	2591	0.11	rs3743398	Hetero
P_419	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_419	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_419	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_419	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_419	Missense	T toC	F to L	1366	0.001	rs181736584	Hetero
P_419	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_445	Missense	G to C	E to D	4227	0.05	rs28559926	Hetero
P_531	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_531	Missense	A to G	T to A	4264	0.02	rs35546357	Hetero
P_531	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_531	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_562	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_624	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_624	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_624	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_624	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_654	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_654	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_654	Missense	T toC	F to L	1366	0.001	rs181736584	Hetero
P_654	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_654	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_659	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_659	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_659	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_659	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_692	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_692	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_692	Missense	G to A	A to T	3637	0.49	rs11638262	Hetero
P_692	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_725	Missense	C to T	P to L	2591	0.11	rs3743398	Hetero
P_725	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_725	Missense	T to A	S to T	2815	0.38	rs938609	Hetero

P_725	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_725	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_752	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_752	Missense	T to A	S to T	2815	0.38	rs938609	Hetero
P_752	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_752	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_752	Missense	A to G	T to A	4207	UKN	rs12899191	Hetero
P_832	Missense	G to C	E to D	4227	0.05	rs28559926	Hetero
P_832	Missense	C to T	S to L	1469	0.005	rs117116488	Hetero
P_942	Missense	A to C	E to A	4523	0.38	rs2882676	Hetero
P_942	Missense	G to T	S to I	2789	0.40	rs938608	Hetero
P_942	Missense	C to G	D to E	7005	0.87	rs3817428	Hetero
P_942	Missense	C to T	P to L	2591	0.11	rs3743398	Hetero
P_942	Missense	T to A	S to T	2815	0.38	rs938609	Hetero

3.6 Matrix Remodeling Enzymes

3.6.1 Matrix Metalloproteinase -2

No patients presented variants of MMP2.

3.6.2 Matrix Metalloproteinase- 3

73% of patients expressed the common allele variant rs679620, with an expected frequency of 68% ($p > 0.005$). This mutation was in the peptidoglycan binding region and resulted in a basic residue being replaced by an acidic residue.

Table 12 MMP3 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Zygosity
P_192	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_196	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_262	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_419	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_445	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_531	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_624	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_659	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_692	Missense	A to G	K to E	133	0.68	rs679620	Hetero

P_832	Missense	A to G	K to E	133	0.68	rs679620	Hetero
P_942	Missense	A to G	K to E	133	0.68	rs679620	Hetero

3.6.3 Matrix Metalloproteinase- 9

No patients were found to have variants in the MMP9 gene.

3.6.4 Thrombospondin 2

The rare variant rs151019608 was found in 13.3% of patients against an expected frequency of 3% ($p < 0.005$). This mutation is found in the 3' UTR. A novel mutation was found in 1 patient who also had the rs151019608 variant. The mutation was at nucleotide 3355 and is in the Concanavalin A-like lectin/glucanase domain. A second novel variant was found in 1 patient at nucleotide 761, in the 3' UTR.

Table 13 THBS2 Variants

Patient	Mut Effect	Ref to Alt	AA Change	Mut Pos	AF	RSN	Haplo
P_752	Missense	T to C	K to E	3355	ukn	Novel	Hetero
P_262	Transvers	T to G	3' UTR	761	ukn	Novel	Hetero
P_659	Deletion	CTAA to C	3' UTR	1139/1141	0.03	151019608	Hetero
P_752	Deletion	CTAA to C	3' UTR	1139/1141	0.03	151019608	Hetero

Chapter 4

Discussion, Future Studies, and Conclusion

4.1 Collagen Findings

Collagen is the chief component of the AF and is an obvious target for research into IVDD. COL1A1 is primarily found in the outer annulus and was the first gene to be investigated, and no immediate structural mutations were found. A finding of note was the over representation of the rs112626388 mutation that occurred 17% more often in our patients than in the general population. The 3' UTR plays a crucial role in gene expression and have been shown to play an irreplaceable role in post translational regulation of gene expression (Matoulkova et al., 2012). Compromise of this region could have a powerfully adverse effect on the viability and efficacy of the mRNA and in turn the downstream processes. Collagen turn over in the IVD is suspected to take over 100 years, so it is unlikely that a 3'UTR in COL1A1 variant would heavily impact remodeling, however, if the fibers were initial improperly trafficked, this could affect the structural integrity of the disc. Based on the over representation of this suspected gene in a highly susceptible population seems to merit further research into the COL1A1 3' UTR. COL9A2 and COL9A3 are heterotrimeric proteins found in small quantities in the NP and AF that are thought to function as an intermediary between collagens and noncollagenous proteins (Aspberg et al., 2012). The rare COL9A2 variant rs140041506 was overrepresented in the patient population. In this variant an AA in the triple helical region is changed from nonpolar to polar. Collagen is composed of three subunits that are woven together to form the characteristic triple helix. The fibers must remain highly structured to maintain viability in the IVD, as previously

discussed; damage to the IVD is synergistic. A shift from one AA class to another could have serious consequences on structural integrity due to alterations in the shape of the coil as the previously sequestered residue is exposed to the hydrophilic exterior. This would cause structural compromise as the adhesion between collagen and the tissue broke down, resulting in failure of the AF. This process is under continued investigation and future studies should continue this course. Speaking to the efficacy of the study the rs61734651 allele was highly over represented as it was found in 20% of patients, while expected in less than 3%. This variant has been described as the Trp3 allele and previously associated with IVDD (Paassilta et al., 2001), though a conclusive link has proven difficult to establish. A likely pathology of this variant would be similar to the compromise of induced by rs140041506. COL11A1 and COL11A2 are heterotrimeric fibers that are secreted as procollagen into the ECM and participate in matrix fiber formation (Aspberg et al., 2012). The COL11A1 allele rs36076089 was found to be over represented and affects the splice region. This variant may represent a less stable alternative splice variants that leads to ECM fiber formation that is less structurally sound. The rs2229784 COL11A2 allele was over represented and is characterized a shift from a hydrophobic residue to a hydrophilic residue in the triple helix domain. This would affect the stability of the helix in a manner similar to that found in the COL9A2 rs140041506 variant.

4.2 Interleukin Findings

There were no significant findings for IL-1 or IL-6 variants. Increased expression levels are seen with IVDD and a plausible pathology exists, as previously discussed. Several Variants have been described and associated in previous studies (Solovieva et

al., 2004; Noponen-Hietala et al., 2005) further investigation with a larger cohort is merited.

4.3 Vitamin D Receptor Findings

The VDR has been associated with IVDD in previous studies (Kawaguchi et al., 2002; Cheung et al., 2006), and is known to play a role in bone mineralization and remodeling. The VDR is a steroid nuclear receptor and thus accurate localization and trafficking is crucial for proper functioning. Two patients presented with two identical 3’ UTR variants, one variant has not been previously characterized and the other lacks proper population data, so statistical significance is difficult to ascertain. Of particular note however, is these mutations are 1 nucleotide apart. MiRNA recognize a 6 nucleotide seed region (Matoulkova et al., 2012) and compromising a third of this sequence would interfere with binding and consequently regulation. Further research into the role of the VDR 3’UTR and IVDD is required before any conclusions can be drawn.

4.4 Insulin-like Growth Factor Receptor Findings

IGF1R can regulate IVD ECM degradation and plays an important role in the normal functioning of the disc tissues and has previously been associated with IVDD (Urano et al., 2008). Several IGF1Rvariants were found to be over represented particularly rs35715134 and rs3051367. These are both variants of the 3’UTR and could play a role in altering proper IGR mediated MMP regulation leading to disc degeneration as discussed earlier. It is noteworthy that rs35715134 has an allele frequency of 49% in the American population, but was found in 14 of the 15 patients in

this study. This study has produced evidence that further genetic studies into IVDD etiology should certainly consider IGF1R variants.

4.5 Aggrecan Findings

Aggrecan is the primary hydrating agent of the IVD and as previously discussed, is central to IVDD onset and progression. ACAN variations were found in substantially higher abundance than variant alleles in any other investigated genes. Although absolute variant number can be accounted for due to the existence of more ACAN variants than in the other genes, several of these genes were found in much higher concentration than the established allele frequencies would predict. The rs938608 and rs938609 alleles were both found to be overrepresented and both found in the same 11 patients, these variations are located 26 nucleotides apart on the same chromosome thus recombination is unlikely and a coupled inheritance is expected. The variant rs2882676 however was found in 10 of the 11 individuals possessing rs938609 and rs938608 but is located almost 2000 base pairs away and is also over represented in the cohort. These variants are all found in the CS domains which code for one of the hydroscopic GAG components of aggrecan; chondroitin. Compromise of this protein inhibits the ability of the IVD to imbibe water and has been shown in studies to induce disc degeneration (Hadjipavlou et al., 2008). The alleles rs28559926 and rs3743398, also reside in the CS domain and were found to be over represented in our study. Two variants were found to be under represented rs11638262 and rs35430524. These alterations would be found in the CS and G3 structural domain respectively. The rs35430524 allele is a discrete alteration between aspartate and glutamate and is unlikely to affect tertiary or quaternary structure (Uitterlinden et al., 2004) and perhaps represents a more stable

conformation with superior hydroscopic characteristics. The rs11638262 allele affects the G3 globular domain which is involved in of the binding of GAGs, the shift from a hydrophobic alanine to a hydrophilic threonine could allow for improved GAG binding, leading to increased NP hydration. Further investigation of CS domain variants in certainly merited with a special emphasis on rs938609, rs938608, and rs2882676.

4.6 Matrix Remodeling Enzyme Findings

There were no statistically relevant findings for MMP variants. Although upregulating of the MMPs is highly implicated in IVD ECM degradation, as previously discussed, and has been characterized in literature (Chen et al., 2006), a definitive genetic link has yet to be established, and perhaps upstream elements and indirect regulators should be investigated. THBS2 is an important modulator of MMP2 and MMP9 levels, and thus the compromise of THBS2 expression could result in the observed MMP mediated ECM degradation found in IVDD. The variant rs151019608 is an alteration of 3'UTR and was found to be over represented in the study and consequently could play a role in aberrant MMP activation leading to IVDD.

4.7 Limitations and Conclusion

Identification of the genetic components of IVDD is a crucial element to elucidating a pathophysiology and for early detection of disc degeneration. A literature review of IVDD was conducted and 5 genes that have probable roles in IVDD and have previously been associated with the disease were selected for investigation. Due to wide variation in radiographic diagnosis of IVDD, a lumbar HNP in the absence of trauma or other risk factors was determined to be a definitive indication of IVDD. Due to the rarity of lumbar HNP without risk factors the study was heavily limited to only 15 patients. The

validity of the inclusion criteria was heavily supported by the over representation of the COL9A3 rs61734651 allele which has previously been indicated in IVDD. The IGF1R variant rs35715134 was also heavily over represented in our patient population and could contribute to IVDD progression through improper mediation of ECM turnover. Finally the aggrecan variants rs938609, rs938608, and rs2882676 were found substantially over represented and likely influence the ability of the disc to synthesize viable chondroitin monomers contributing to poor hydration of the NP and ultimately structural failure of the IVD. These aberrations are in pathways strongly associated IVDD. These findings are, however limited in scope of applicability due to the small sample size employed and the age of the patients. As previously noted, IVDD is poorly understood and characterized, thus it is conceivable that this term actually encompasses several manifestation of degeneration, each with discrete pathologies. If this were the case, the above variants may only be applicable as genetic risk factors to an atraumatic early manifestation of IVDD. This supposition would help to explain the presence of previously reported rare variants such as rs61734651 but the lack of consistently reported inflammatory variants. Baring these limitations in mind, future studies on the genetic elements of an early presentation of IVDD should be conducted on a large sample of young adults who have had minimal to no exposure to suspected risk factors but have suffered a lumbar HNP and should focus on alterations of collagen, the insulin-like growth factor receptor, and aggrecan.

REFERENCES

Aaro, Stig. "The sciatic syndrome." (1996): 641-642.

Adams, M. A., and W. C. Hutton. "Prolapsed Intervertebral Disc: A Hyperflexion Injury." *Spine* 7.3 (1982): 184-191.

Adams, M. A., and W. C. Hutton. "The mechanical function of the lumbar apophyseal joints." *Spine* 8.3 (1983): 327-330.

Adams, Michael A. "Biomechanics of back pain." *Acupuncture in medicine*22.4 (2004): 178-188.

Adams, Michael A., and Patricia Dolan. "Could sudden increases in physical activity cause degeneration of intervertebral discs?." *The Lancet* 350.9079 (1997): 734-735.

Adams, Michael A., and Peter J. Roughley. "What is intervertebral disc degeneration, and what causes it?." *Spine* 31.18 (2006): 2151-2161.

Aigner, T., et al. "Variation with age in the pattern of type X collagen expression in normal and scoliotic human intervertebral discs." *Calcified tissue international* 63.3 (1998): 263-268.

Ala-Kokko, Leena. "Genetic risk factors for lumbar disc disease." *Annals of medicine* 34.1 (2002): 42-47.

An, Howard S., et al. "Introduction: disc degeneration: summary." *Spine*29.23 (2004): 2677-2678.

Antoniou, John, et al. "The human lumbar intervertebral disc: evidence for changes in the biosynthesis and denaturation of the extracellular matrix with growth, maturation, ageing, and degeneration." *Journal of Clinical Investigation* 98.4 (1996): 996.

Aspberg, Anders. "The different roles of aggrecan interaction domains."*Journal of Histochemistry & Cytochemistry* 60.12 (2012): 987-996.

Banfi, Giuseppe, Gian Luca Salvagno, and Giuseppe Lippi. "The role of ethylenediamine tetraacetic acid (EDTA) as in vitro anticoagulant for diagnostic purposes." *Clinical Chemical Laboratory Medicine* 45.5 (2007): 565-576.

Battie, M. C., et al. "1995 Volvo Award in clinical sciences. Determinants of lumbar disc degeneration. A study relating lifetime exposures and magnetic resonance imaging findings in identical twins." *Spine* 20.24 (1995): 2601-2612.

Battié, Michele C., et al. "1991 Volvo Award in Clinical Sciences: Smoking and Lumbar Intervertebral Disc Degeneration: An MRI Study of Identical Twins." *Spine* 16.9 (1991): 1015-1021.

Battié, Michele C., et al. "The Twin Spine Study: contributions to a changing view of disc degeneration." *The Spine Journal* 9.1 (2009): 47-59.

Benneker, Lorin M., et al. "2004 Young Investigator Award Winner: vertebral endplate marrow contact channel occlusions and intervertebral disc degeneration." *Spine* 30.2 (2005): 167-173.

Bertram, Helge, et al. "Accelerated intervertebral disc degeneration in scoliosis versus physiological ageing develops against a background of enhanced anabolic gene expression." *Biochemical and biophysical research communications* 342.3 (2006): 963-972.

Brinckmann, Paul, et al. "Deformation of the vertebral end-plate under axial loading of the spine." *Spine* 8.8 (1983): 851-856.

Brisby, Helena. "Pathology and possible mechanisms of nervous system response to disc degeneration." *The Journal of Bone & Joint Surgery*88.suppl 2 (2006): 68-71.

Bron, Johannes Leendert, et al. "Repair, regenerative and supportive therapies of the annulus fibrosus: achievements and challenges." *European spine journal* 18.3 (2009): 301-313.

Bruehlmann, Sabina B., et al. "Regional variations in the cellular matrix of the annulus fibrosus of the intervertebral disc." *Journal of anatomy* 201.2 (2002): 159-171.

Buckwalter, Josef A., Henry J. Mankin, and Alan J. Grodzinsky. "Articular cartilage and osteoarthritis." *Instructional Course Lectures-American Academy of Orthopaedic Surgeons* 54 (2005): 465.

Buckwalter, Joseph A. "Aging and degeneration of the human intervertebral disc." *Spine* 20.11 (1995): 1307-1314.

Butler, Wallace F. "Comparative anatomy and development of the mammalian disc." *The biology of the intervertebral disc* 1 (1989): 83-108.

Cassidy, J. J., A. Hiltner, and E. Baer. "Hierarchical structure of the intervertebral disc." *Connective tissue research* 23.1 (1989): 75-88.

Cassinelli, Ezequiel H., and James D. Kang. "Current understanding of lumbar disc degeneration." *Operative techniques in orthopaedics* 10.4 (2000): 254-262.

Chan, Danny, et al. "Genetics of disc degeneration." *European spine journal*15.3 (2006): 317-325.

Chen, Jian-Min, Claude Férec, and David N. Cooper. "A systematic analysis of disease-associated variants in the 3′ regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3′ UTR variants." *Human genetics* 120.3 (2006): 301-333.

Chen, Kai, et al. "Gene expression profile analysis of human intervertebral disc degeneration." *Genetics and molecular biology* 36.3 (2013): 448-454.

Cheung, Kenneth MC, et al. "Association of the Taq I allele in vitamin D receptor with degenerative disc disease and disc bulge in a Chinese population." *Spine* 31.10 (2006): 1143-1148.

Cheung, Kenneth MC. "The relationship between disc degeneration, low back pain, and human pain genetics." *The spine journal* 10.11 (2010): 958-960.

Colombini, Alessandra, et al. "Pathophysiology of the human intervertebral disc." *The international journal of biochemistry & cell biology* 40.5 (2008): 837-842.

Coventry, Mark B., Ralph K. Ghormley, and James W. Kernohan. "The intervertebral disc: its microscopic anatomy and pathology." *The Journal of Bone & Joint Surgery* 27.3 (1945): 460-474.

Deyo, Richard A., and J. Edward Bass. "Lifestyle and low-back pain: the influence of smoking and obesity." *Spine* 14.5 (1989): 501-506.

Diamant, B., J. Karlsson, and A. Nachemson. "Correlation between lactate levels and pH in discs of patients with lumbar rhizopathies." *Experientia*24.12 (1968): 1195-1196.

Dong, D. M., et al. "Association between the-1306C/T polymorphism of matrix metalloproteinase-2 gene and lumbar disc disease in Chinese young adults." *European Spine Journal* 16.11 (2007): 1958-1961.

Eskola, Pasi J., et al. "Genetic association studies in lumbar disc degeneration: a systematic review." *PLoS One* 7.11 (2012): e49995.

Freemont, A. J. "The cellular pathobiology of the degenerate intervertebral disc and discogenic back pain." *Rheumatology* 48.1 (2009): 5-10.
Frymoyer, John W. "Predicting disability from low back pain." *Clinical orthopaedics and related research* 279 (1992): 101-109.

Gologorsky, Yakov, and John Chi. "Genetic predisposition to lumbar disc degeneration." *Neurosurgery* 74.2 (2014): N10-N11.

Goupille, Philippe, et al. "The role of inflammation in disk herniation-associated radiculopathy." *Seminars in arthritis and rheumatism*. Vol. 28. No. 1. WB Saunders, 1998.

Griffith, James F., et al. "Modified Pfirrmann grading system for lumbar intervertebral disc degeneration." *Spine* 32.24 (2007): E708-E712.

Grignon, B., et al. "The structure of the cartilaginous end-plates in elder people." *Surgical and Radiologic Anatomy* 22.1 (2000): 13-19.

Gruber, Helen E., and Edward N. Hanley Jr. "Analysis of aging and degeneration of the human intervertebral disc: comparison of surgical specimens with normal controls." *Spine* 23.7 (1998): 751-757.

Hadjipavlou, A. G., et al. "The pathophysiology of disc degeneration." *Bone & Joint Journal* 90.10 (2008): 1261-1270.

Han, I. B., et al. "Association between VEGF and eNOS gene polymorphisms and lumbar disc degeneration in a young Korean population."*Genetics and Molecular Research* 12.3 (2013): 2294-2305.

Hassett, G., et al. "Risk factors for progression of lumbar spine disc degeneration: the Chingford Study." *Arthritis & Rheumatism* 48.11 (2003): 3112-3117.

Heliövaara, Markku. "Risk factors for low back pain and sciatica." *Annals of medicine* 21.4 (1989): 257-264.

Hirose, Yuichiro, et al. "A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation." *The American Journal of Human Genetics* 82.5 (2008): 1122-1129.

Holm, S., et al. "Nutrition of the intervertebral disc: solute transport and metabolism." *Connective tissue research* 8.2 (1981): 101-119.

Horner, Heather A., and Jill PG Urban. "2001 Volvo Award Winner in Basic Science Studies: effect of nutrient supply on the viability of cells from the nucleus pulposus of the intervertebral disc." *Spine* 26.23 (2001): 2543-2549.

Horner, Heather A., et al. "Cells from different regions of the intervertebral disc: effect of culture system on matrix expression and cell phenotype."*Spine* 27.10 (2002): 1018-1028

Hunter, C. J., J. R. Matyas, and N. A. Duncan. "The notochordal cell in the nucleus pulposus: a review in the context of tissue engineering." *Tissue engineering* 9.4 (2003): 667-677.

Ishihara, Hirokazu, and Jill PG Urban. "Effects of low oxygen concentrations and metabolic inhibitors on proteoglycan and protein synthesis rates in the intervertebral disc." *Journal of Orthopaedic Research* 17.6 (1999): 829-835.

Jensen, Maureen C., et al. "Magnetic resonance imaging of the lumbar spine in people without back pain." *New England Journal of Medicine* 331.2 (1994): 69-73.

Johnson, William EB, et al. "Human intervertebral disc aggrecan inhibits nerve growth in vitro." *Arthritis & Rheumatism* 46.10 (2002): 2658-2664.

Jordan, Joanne L, Kika Konstantinou, and John O'Dowd. "Herniated Lumbar Disc." *BMJ Clinical Evidence* 2011 (2011): 1118. Print.

Kaigle, Allison M., Sten H. Holm, and Tommy H. Hansson. "1997 volvo award winner in biomechanical studies: Kinematic behavior of the porcine lumbar spine: A chronic lesion model." *Spine* 22.24 (1997): 2796-2806.

Kalichman, Leonid, and David J. Hunter. "The genetics of intervertebral disc degeneration. Associated genes." *Joint Bone Spine* 75.4 (2008): 388-396.

Karppinen, Jaro, et al. "Magnetic resonance imaging findings in relation to the COL9A2 tryptophan allele among patients with sciatica." *Spine* 27.1 (2002): 78-82.

Kashyap, Ajit Singh, et al. "Peripheral neuropathy." *The Lancet* 364.9446 (2004): 1664

Katz, Michael M., Alan R. Hargens, and Steven R. Garfin. "Intervertebral Disc Nutrition: Diffusion Versus Convection." *Clinical orthopaedics and related research* 210 (1986): 243-245.

Kawaguchi, Yoshiharu, et al. "Association between an aggrecan gene polymorphism and lumbar disc degeneration." *Spine* 24.23 (1999): 2456.

Kawaguchi, Yoshiharu, et al. "The association of lumbar disc disease with vitamin-D receptor gene polymorphism." *J Bone Joint Surg Am* 84.11 (2002): 2022-2028.

Kenny, David A., et al. "The statistical analysis of data from small groups."*Journal of personality and social psychology* 83.1 (2002): 126.

Kraemer, Juergen. "Natural Course and Prognosis of Intervertebral Disc Diseases." *Spine* 20.6 (1995): 635-639.

Lavy, Chris, et al. "Cauda equina syndrome." *Obstetric Anesthesia Digest*30.1 (2010): 16-17.

Le Maitre, C. L., et al. "Matrix synthesis and degradation in human intervertebral disc degeneration." *Biochemical Society Transactions* 35.4 (2007): 652-655.

Lee, Cynthia R., et al. "A phenotypic comparison of intervertebral disc and articular cartilage cells in the rat." *European Spine Journal* 16.12 (2007): 2174-2185.

Li, B., et al. "Reduced expression of insulin-like growth factor 1 receptor leads to accelerated intervertebral disc degeneration in mice." *International journal of immunopathology and pharmacology* 26.2 (2013): 337-347.

Liu, Kun, et al. "Foot Drop Caused by Lumbar Degenerative Disease: Clinical Features, Prognostic Factors of Surgical Outcome and Clinical Stage." *PloS one* 8.11 (2013): e80375.

Lotz, J. C., A. J. Fields, and E. C. Liebenberg. "The role of the vertebral end plate in low back pain." *Global Spine J* 3.3 (2013): 153-64.

Lotz, Jeffrey C., et al. "Compression-induced degeneration of the intervertebral disc: an in vivo mouse model and finite-element study." *Spine*23.23 (1998): 2493-2506.

Luo, Xuemei, et al. "Estimates and patterns of direct health care expenditures among individuals with back pain in the United States." *Spine*29.1 (2004): 79-86.

Maniadakis, Nikolaos, and Alastair Gray. "The economic burden of back pain in the UK." *Pain* 84.1 (2000): 95-103.

Martin, James A., et al. "Chondrocyte senescence, joint loading and osteoarthritis." *Clinical Orthopaedics and Related Research* 427 (2004): S96-S103.

Martin, Michael D., Christopher M. Boxell, and David G. Malone. "Pathophysiology of lumbar disc degeneration: a review of the literature."*Neurosurgical focus* 13.2 (2002): 1-6.

Matoulkova, Eva, et al. "The role of the 3'untranslated region in post-transcriptional regulation of protein expression in mammalian cells." *RNA biology* 9.5 (2012): 563-576.

Matsui, Hisao, et al. "Familial predisposition and clustering for juvenile lumbar disc herniation." *Spine* 17.11 (1992): 1323-1328.

Mayer, Jillian E., et al. "Genetic polymorphisms associated with intervertebral disc degeneration." *The Spine Journal* 13.3 (2013): 299-317.

McDevitt, Cahir A., et al. "Experimental osteoarthritic articular cartilage is enriched in guanidine-soluble type VI collagen." *Biochemical and biophysical research communications* 157.1 (1988): 250-255.

McNally, D. S., and M. A. Adams. "Internal intervertebral disc mechanics as revealed by stress profilometry." *Spine* 17.1 (1992): 66-73.

Mengoni, Marlène et al. "Derivation of Inter-Lamellar Behaviour of the Intervertebral Disc Annulus." *Journal of the Mechanical Behavior of Biomedical Materials* 48 (2015): 164–172. *PMC*. Web. 11 June 2016.

Mio, Futoshi, et al. "A functional polymorphism in COL11A1, which encodes the α1 chain of type XI collagen, is associated with susceptibility to lumbar disc herniation." *The American Journal of Human Genetics* 81.6 (2007): 1271-1277.

Modic, M. T., et al. "Imaging of degenerative disk disease." *Radiology* 168.1 (1988): 177-186.

Modic, Michael T., and Jeffrey S. Ross. "Lumbar Degenerative Disk Disease 1." *Radiology* 245.1 (2007): 43-61.

Moon, Hong Joo, et al. "Effects of secreted factors in culture medium of annulus fibrosus cells on microvascular endothelial cells: elucidating the possible pathomechanisms of matrix degradation and nerve in-growth in disc degeneration." *Osteoarthritis and Cartilage* 22.2 (2014): 344-354

Moon, Sung M., et al. "Evaluation of intervertebral disc cartilaginous endplate structure using magnetic resonance imaging." *European Spine Journal* 22.8 (2013): 1820-1828.

Moore, Robert J., et al. "The origin and fate of herniated lumbar intervertebral disc tissue." *Spine* 21.18 (1996): 2149-2155.

Nachemson, A., et al. "In vitro diffusion of dye through the end-plates and the annulus fibrosus of human lumbar inter-vertebral discs." *Acta Orthopaedica Scandinavica* 41.6 (1970): 589-607.

Neidlinger-Wilke, Cornelia, et al. "Regulation of gene expression in intervertebral disc cells by low and high hydrostatic pressure." *European Spine Journal* 15.3 (2006): 372-378.

Nerlich, Andreas G., Beatrice E. Bachmeier, and Norbert Boos. "Expression of fibronectin and TGF-ß1 mRNA and protein suggest altered regulation of extracellular matrix in degenerated disc tissue." *European Spine Journal* 14.1 (2005): 17-26.

Noponen-Hietala, Noora, et al. "Genetic variations in IL6 associate with intervertebral disc disease characterized by sciatica." *Pain* 114.1 (2005): 186-194.

Ohshima, H., and J. P. G. Urban. "The effect of lactate and pH on proteoglycan and protein synthesis rates in the intervertebral disc." *Spine*17.9 (1992): 1079-1082.

Paassilta, Petteri, et al. "Identification of a novel common genetic risk factor for lumbar disk disease." *Jama* 285.14 (2001): 1843-1849.

Pattappa, Girish et al. "Diversity of Intervertebral Disc Cells: Phenotype and Function." *Journal of Anatomy* 221.6 (2012): 480–496. *PMC*. Web. 11 June 2016.

Pearce, Richard H., et al. "Magnetic resonance imaging reflects the chemical changes of aging degeneration in the human intervertebral disk."*The Journal of rheumatology. Supplement* 27 (1991): 42-43.

Pfirrmann, Christian WA, et al. "Magnetic resonance classification of lumbar intervertebral disc degeneration." *Spine* 26.17 (2001): 1873-1878.

Pluijm, S. M. F., et al. "Collagen type I α1 Sp1 polymorphism, osteoporosis, and intervertebral disc degeneration in older men and women." *Annals of the rheumatic diseases* 63.1 (2004): 71-77.

Pope, Malcolm H., Kheng Lim Goh, and Marianne L. Magnusson. "Spine ergonomics." *Annual review of biomedical engineering* 4.1 (2002): 49-68.

Pope, Malcolm H., Marianne Magnusson, and David G. Wilder. "Low back pain and whole body vibration." *Clinical orthopaedics and related research*354 (1998): 241-248.

Pye, Stephen R., et al. "Lumbar disc degeneration: association between osteophytes, end-plate sclerosis and disc space narrowing." *Annals of the rheumatic diseases* 66.3 (2007): 330-333.

Raj, P. Prithvi. "Intervertebral Disc: Anatomy-Physiology-Pathophysiology-Treatment." *Pain Practice* 8.1 (2008): 18-44.

Resnick, Donald. "Degenerative diseases of the vertebral column."*Radiology* 156.1 (1985): 3-14.

Roberts, S. "Disc morphology in health and disease." *Biochemical Society Transactions* 30.6 (2002): 864-869.

Roberts, S., et al. "Type X collagen in the human invertebral disc: an indication of repair or remodelling?." *The Histochemical journal* 30.2 (1998): 89-95.

Roughley, P., et al. "The involvement of aggrecan polymorphism in degeneration of human intervertebral disc and articular cartilage." *Eur Cell Mater* 11.1 (2006): 1-7.

Saleem, Shafaq, et al. "Lumbar disc degenerative disease: disc degeneration symptoms and magnetic resonance image findings." *Asian spine journal* 7.4 (2013): 322-334.

Sambrook, P. N., A. J. MacGregor, and T. D. Spector. "Genetic influences on cervical and lumbar disc degeneration." *Arthritis Rheum* 42.2 (1999): 336.

Schroeder, Gregory D. et al "The epidemiology and pathophysiology of lumbar disc herniations." Semin Spine Surg 28 (2016): 2-7

Science Watch 2012 "Hottest research 2012" available via (http://archive.sciencewatch.com/newsletter/2012/201203/hottest_research_2012) [Accessed 3/14/16]

Shankar, Hariharan, Jeremy A. Scarlett, and Stephen E. Abram. "Anatomy and pathophysiology of intervertebral disc disease." *Techniques in Regional Anesthesia and Pain Management* 13.2 (2009): 67-75.

Solovieva, Svetlana, et al. "Interleukin 1 polymorphisms and intervertebral disc degeneration." *Epidemiology* 15.5 (2004): 626-633.

Spengler, D. M., et al. "Elective discectomy for herniation of a lumbar disc. Additional experience with an objective method." *The Journal of Bone & Joint Surgery* 72.2 (1990): 230-237.

Stockwell, R. A. "The interrelationship of cell density and cartilage thickness in mammalian articular cartilage." *Journal of anatomy* 109.Pt 3 (1971): 411.

Sun, Zheng-ming, et al. "Association between the-1562 C/T polymorphism of matrix metalloproteinase-9 gene and lumbar disc disease in the young adult population in North China." *Connective tissue research* 50.3 (2009): 181-185.

Szostak, Emilia, and Fátima Gebauer. "Translational control by 3′-UTR-binding proteins." *Briefings in functional genomics* (2012): els056.

Taher, Fadi, et al. "Lumbar degenerative disc disease: current and future concepts of diagnosis and management." *Advances in orthopedics* 2012 (2012).

Takahashi, M., et al. "The association of degeneration of the intervertebral disc with 5a/6a polymorphism in the promoter of the human matrix metalloproteinase-3 gene." *Bone & Joint Journal* 83.4 (2001): 491-495.

Tilkeridis, C., et al. "Association of a COL1A1 polymorphism with lumbar disc disease in young military recruits." *Journal of medical genetics* 42.7 (2005): e44-e44.

Uitterlinden, André G., et al. "Genetics and biology of vitamin D receptor polymorphisms." *Gene* 338.2 (2004): 143-156.

Urano, Tomohiko, et al. "Association of a single nucleotide polymorphism in the insulin-like growth factor-1 receptor gene with spinal disc degeneration in postmenopausal Japanese women." *Spine* 33.11 (2008): 1256-1261.

Urban, J. P. G., S. Roberts, and J. R. Ralphs. "The nucleus of the intervertebral disc from development to degeneration." *American Zoologist*40.1 (2000): 53-061.

Urban, J. P., and Sally Roberts. "Degeneration of the intervertebral disc."*Arthritis Research and Therapy* 5.3 (2003): 120-138.

Videman, Tapio and Eric Parent. "Lumbar disc degeneration: epidemiology and genetic influences." *Spine* 29.23 (2004): 2679-2690.

Videman, Tapio, et al. "Associations of 25 structural, degradative, and inflammatory candidate genes with lumbar disc desiccation, bulging, and height narrowing." *Arthritis & Rheumatism* 60.2 (2009): 470-481.

Walker, Matthew H., and D. Greg Anderson. "Molecular basis of intervertebral disc degeneration." *The Spine Journal* 4.6 (2004): S158-S166.

Wang, Xiu-Jie, et al. "Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets." *Genome biology* 5.9 (2004): R65.

Whatley, Benjamin R., and Xuejun Wen. "Intervertebral disc (IVD): structure, degeneration, repair and regeneration." *Materials Science and Engineering: C*32.2 (2012): 61-77.

Yu, Jing, et al. "Microfibrils, elastin fibres and collagen fibres in the human intervertebral disc and bovine tail disc." *Journal of anatomy* 210.4 (2007): 460-471.

Zhang, Yin'gang, et al. "Advances in susceptibility genetics of intervertebral degenerative disc disease." *Int J Biol Sci* 4.5 (2008): 283-290.

Zhao, Chang-Qing, et al. "The cell biology of intervertebral disc aging and degeneration." *Ageing research reviews* 6.3 (2007): 247-261.