Isolation and phenotypic characterisation of human notochordal cells

Implications for the development of cell-based therapies for intervertebral disc degeneration

Ricardo Rodrigues Pinto

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy (PhD) in the Faculty of Medical and Human Sciences

2015

Institute of Inflammation and Repair
School of Medicine
Chapter 1  Introduction

1.1 BACK PAIN

1.2 THE INTERVERTEBRAL DISC

1.2.1 Nucleus pulposus

1.2.2 Annulus fibrosus

1.2.3 Endplate

1.3 INTERVERTEBRAL DISC DEGENERATION

1.3.1 Pathogenesis of disc degeneration

1.3.2 Aetiology of intervertebral disc degeneration

1.3.2.1 Mechanical load

1.3.2.2 Glucose starvation

1.3.2.3 Low pH

1.3.2.4 Inherited predisposition
1.4 THERAPIES FOR INTERVERTEBRAL DISC DEGENERATION

1.4.1 Conservative treatments 50
1.4.2 Surgical treatments 50
1.4.3 Biological treatments 52
  1.4.3.1 The ideal cell for NP tissue engineering 53
    1.4.3.1.1 Autologous nucleus pulposus cells 54
    1.4.3.1.2 Allogeneic nucleus pulposus cells 54
    1.4.3.1.3 Stem cells 55
      1.4.3.1.3.1 Embryonic stem cells 55
      1.4.3.1.3.2 Induced pluripotent stem cells 56
      1.4.3.1.3.3 Adult stem cells 57

1.5 THE SEARCH FOR THE IDEAL CELL PHENOTYPE 58
  1.5.1 The nucleus pulposus versus the articular cartilage cell phenotype 58
  1.5.2 The nucleus pulposus cell phenotype: interspecies variations 61

1.6 THE ROLE OF NOTOCHORDAL CELLS IN THE INTERVERTEBRAL DISC 67

1.7 EMBRYOLOGIC AND FOETAL DEVELOPMENT OF THE IVD 69

1.8 ONTOGENY OF THE CELLS POPULATING THE ADULT NUCLEUS PULPOSUS 72

1.9 THE NOTOCHORDAL CELL PHENOTYPE 75

1.10 SUMMARY 83

1.11 HYPOTHESES AND OVERALL AIMS OF THE PROJECT 86
Chapter 2  Identification of a human notochord-specific marker

2.1  INTRODUCTION

2.2  HYPOTHESES AND AIMS

2.3  MATERIALS AND METHODS

2.3.1 Sample acquisition and staging

2.3.2 Human embryonic and foetal spine dissection

2.3.3 Sample preparation for immunohistology

2.3.4 Haematoxylin and eosin stain

2.3.5 Masson’s trichrome staining

2.3.6 Safranin-O staining

2.3.7 Immunohistochemistry

2.3.8 Optimisation of immunohistochemical staining

2.3.9 Image analysis

2.4  RESULTS

2.4.1 Morphology

2.4.2 Extracellular matrix

2.4.3 Identification of developmental markers

2.4.3.1 Notochord-specific markers

2.4.3.5 Proteins with variable expression with developmental stages

2.4.3.11 Proteins without expression in the developing spine

2.4.3.12 Background staining

2.4.5 Summary of results

2.5  DISCUSSION
Chapter 3 Isolation and characterisation of human notochordal cells isolated using the specific intracellular marker KRT18

3.1 INTRODUCTION
3.2 HYPOTHESES AND AIMS
3.3 SUMMARY OF THE EXPERIMENTAL DESIGN
3.4 MATERIALS AND METHODS
3.4.1 Acquisition of cells
3.4.2 Thawing and expansion of frozen cells
3.4.3 Expansion of cells
3.4.4 Cell counts
3.4.5 Freezing of cells
3.4.6 Development of a method to extract RNA from fixed, permeabilised, labelled and sorted cells
3.4.6.1 Labelling of cells
3.4.6.2 Sorting of KRT18\textsuperscript{+} and KRT18\textsuperscript{-} cells
3.4.6.3 RNA extraction, quantity and quality assessment and amplification to cDNA
3.4.6.4 Assessment of RNA quantity, quality and quality
3.4.7 RNA extraction from KRT18\textsuperscript{+} and KRT18\textsuperscript{-} foetal spine cells
3.4.7.1 Extraction, KRT18-labelling and sorting of foetal spine cells
3.4.7.2 RNA amplification and purification
| 3.4.7.3 | Real-time quantitative polymerase chain reaction analysis | 219 |
| 3.4.8  | Complimentary DNA (cDNA) microarrays                  | 222 |
| 3.4.9  | Microarray analysis                                    | 223 |
| 3.4.10 | Immunolocalisation of HGF and c-Met                   | 225 |

3.5 RESULTS

| 3.5.1  | Development of a method to extract RNA from fixed, permeabilised, labelled and sorted cells | 226 |
| 3.5.2  | Isolation and RNA extraction from notochordal and sclerotomal foetal spine cells          | 231 |
| 3.5.3  | Real time quantitative PCR validation of notochordal cell separation                     | 234 |
| 3.5.4  | Microarray analysis                                                                          | 235 |
| 3.5.5  | Expression of HGF and c-Met in the developing spine                                        | 243 |

3.6 DISCUSSION

3.7 IMPLICATIONS OF THIS INVESTIGATION

Chapter 4 Isolation and characterisation of human notochordal cells using the specific cell surface marker CD24

| 4.1    | INTRODUCTION                                      | 263 |
| 4.2    | HYPOTHESES AND AIMS                              | 265 |
| 4.3    | SUMMARY OF THE EXPERIMENTAL DESIGN               | 266 |
| 4.4    | MATERIALS AND METHODS                            | 268 |
| 4.4.1  | Cell labelling optimisation                      | 268 |
| 4.4.2  | Development of a methodology to isolate viable CD24<sup>+</sup> cells from collagenous tissues | 269 |
4.4.3 RNA extraction from viable CD24\textsuperscript{+ve} and CD24\textsuperscript{-ve} foetal spine cells

4.4.4 qRT-PCR

4.4.5 cDNA microarrays

4.4.6 Microarray analysis

4.5 RESULTS

4.5.1 Optimisation of Draq\textsuperscript{7®} and CD24 labelling

4.5.2 Methodology to isolate viable CD24\textsuperscript{+ve} cells from collagenous tissues

4.5.3 Separation of CD24\textsuperscript{+ve} and CD24\textsuperscript{-ve} foetal spine cells and qRT-PCR validation of cell separation

4.5.4 cDNA microarrays

4.5.4.1 Microarray quality control

4.5.4.2 Hierarchical clustering and principal clustering analysis

4.5.4.3 Identification of CD24\textsuperscript{+ve} and CD24\textsuperscript{-ve} markers

4.5.4.4 Gene pathway analysis

4.6 DISCUSSION

4.7 IMPLICATIONS OF THIS INVESTIGATION

Chapter 5 Validation of identified markers in a notochordal cell population and expression analysis in the adult NP

5.1 INTRODUCTION

5.2 HYPOTHESES AND AIMS

5.3 MATERIALS AND METHODS

5.3.1 Sample acquisition and processing

5.3.1.1 Embryonic and foetal samples
5.3.1.2 Adult NP tissue

5.3.2 RNA isolation from embryonic and foetal notochordal and sclerotomal cells

5.3.3 RNA isolation and cDNA synthesis from foetal notochordal and sclerotomal tissue
  5.3.3.1 Notochordal and sclerotomal tissue dissection
  5.3.3.2 RNA extraction, amplification, purification and cDNA synthesis

5.3.4 RNA isolation and cDNA synthesis from adult NP cells
  5.3.4.1 NP cell isolation
  5.3.4.2 RNA extraction
  5.3.4.3 Genomic DNA elimination
  5.3.4.4 cDNA synthesis

5.3.5 qRT-PCR

5.3.6 Immunohistochemistry

5.3.7 Statistical analysis

5.4 RESULTS

5.4.1 Validation of the markers identified in the KRT18 microarray analysis
  5.4.1.1 Gene expression of KRT18^+ve and KRT18^-ve markers
    5.4.1.1.1 Expression of KRT18^+ve and KRT18^-ve markers in CD24^+ve and CD24^-ve cells isolated from developing foetal spines
5.4.1.1.2 Expression of KRT18$^{+ve}$ and KRT18$^{-ve}$ markers in notochordal and sclerotomal tissues microdissected from an 18 WPC sample

5.4.1.2 Protein expression of the KRT18$^{+ve}$ marker ADORA3 in embryonic and foetal IVD tissues

5.4.2 Validation of the markers identified in the CD24 microarray analysis

5.4.2.1 Gene expression of CD24$^{+ve}$ and CD24$^{-ve}$ markers

5.4.2.1.1 Expression of CD24$^{+ve}$ and CD24$^{-ve}$ markers in sorted CD24$^{+ve}$ and CD24$^{-ve}$ cells isolated from the developing foetal spines used in the CD24 microarray analysis

5.4.2.1.2 Expression of CD24$^{+ve}$ and CD24$^{-ve}$ markers in sorted CD24$^{+ve}$ and CD24$^{-ve}$ cells isolated from developing foetal spines from two additional specimens

5.4.2.1.3 Expression of CD24$^{+ve}$ and CD24$^{-ve}$ markers in notochordal and sclerotomal tissues microdissected from an 18 WPC sample

5.4.2.2 Protein expression of CD24$^{+ve}$ (notochordal) markers in embryonic and foetal IVD tissues

5.4.7 Expression of notochordal KRT18$^{+ve}$ markers in adult NP samples

5.4.7.1 Gene expression of notochordal KRT18$^{+ve}$ markers in adult NP samples

5.4.7.2 Protein expression of notochordal KRT18 markers in adult NP samples
5.4.8 Expression of notochordal CD24 markers in adult NP samples

5.4.8.1 Gene expression of notochordal CD24 markers in adult NP samples

5.4.8.2 Protein expression of notochordal CD24 markers in adult NP samples

5.5 DISCUSSION

Chapter 6 Conclusions and future work

6.1 GENERAL DISCUSSION AND CONCLUSIONS

6.2 FUTURE WORK

Chapter 7 References

Word count: 82.324
**LIST OF TABLES**

**Table 1.1** Genes whose polymorphisms have been associated with disc degeneration. 49

**Table 1.2** NP markers identified in different animal species. 64

**Table 1.3** Mice notochordal/immature NP markers. 78

**Table 1.4** Rat notochordal/immature NP markers. 79

**Table 1.5** Pig notochordal/immature NP markers. 80

**Table 1.6** Bovine notochordal/immature NP markers. 81

**Table 1.7** Human notochordal/immature NP markers. 82

**Table 2.1** Human embryonic and foetal sample details. 97

**Table 2.2** Details of the antibodies and antigen retrieval methods used. 106

**Table 3.1** Fixation and permeabilisation agents tested. 209

**Table 3.2** MCF-7 and T/C28a2 cell mixtures used. 210

**Table 3.3** Thermal cycling programming for RNA amplification. 218

**Table 3.4** TaqMan qRT-PCR assay mastermix. 219

**Table 3.5** Thermal cycler programming for qRT-PCR. 220

**Table 3.6** Human oligonucleotide primers and probes used for qRT-PCR analysis. 221

**Table 3.7** Details of the antibodies and antigen retrieval methods used. 225

**Table 3.8** Top differentially expressed genes between KRT18<sup>+</sup> and KRT18<sup>-</sup> sorted foetal spine cells. 237

**Table 3.9** Top 10 cell surface KRT18<sup>+</sup> markers. 239
Table 3.10 Upstream regulators of differentially expressed genes between KRT18<sup>+</sup> and KRT18<sup>-</sup> cells.

Table 4.1 Identification and age of the samples used in the CD24 microarray analysis.

Table 4.2 Human oligonucleotide primers and probes used for qRT-PCR analysis.

Table 4.3 IPA biological function prediction.

Table 4.4 Proportion of viable cells (Draq7<sup>-</sup>) after incubation for 1, 2 or 16 hours in Accumax® or type II collagenase.

Table 4.5 Proportion of viable CD24<sup>+</sup> and CD24<sup>-</sup> events in each foetal sample.

Table 4.6 dchip analysis of samples used in the microarray analysis.

Table 4.7 Top (upregulated) and bottom (downregulated) differentially expressed genes between CD24<sup>+</sup> and CD24<sup>-</sup> cells.

Table 4.8 High-scoring networks (Score &gt;40) identified by IPA.

Table 4.9 Top 20 diseases and functions associated with the genes in the dataset.

Table 4.10 Upstream regulators of differentially expressed genes.

Table 4.11 Master regulators of differentially expressed genes in the CD24 microarray analysis.

Table 5.1 Samples used for microarray validation.

Table 5.2 Histological grading system for IVD degeneration.

Table 5.3 Details of adult NP samples used for gene (qRT-PCR) and protein (immunohistochemistry (IHC)) analysis.

Table 5.4 Reverse transcription mastermix.
Table 5.5 Thermal cycler programming for cDNA synthesis.

Table 5.6 Human oligonucleotide primers and probes used for qRT-PCR analysis.

Table 5.7 SYBR green qRT-PCR assay mastermix.

Table 5.8 Thermal cycler programming for qRT-PCR.

Table 5.9 Details of the antibodies and antigen retrieval methods used.

Table 5.10 Spearman r values for the correlation between the adult NP expression of GRB14, ADORA3, CDH6, ANPEP and CD69 and the degeneration grade.

Table 5.11 ADORA3 staining of adult NP samples. ADORA3 was expressed in all but two of the samples analysed.

Table 5.12 Spearman r values for the correlation between the adult NP expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 and the degeneration grade.

Table 5.13 MAP1B staining of adult NP samples. MAP1B was expressed in all but one of the samples analysed.

Table 5.14 Summary of the validation results for the markers identified in the KRT18 microarray analysis.

Table 5.15 Summary of the validation results for the markers identified in the CD24 microarray analysis.
LIST OF FIGURES

Figure 1.1 Disc degeneration cascade. 36

Figure 1.2 Illustration of a spinal segment and IVD depicting its major structures. 39

Figure 1.3 Overview of possible biological treatments to regenerate or repair the degenerated IVD. 53

Figure 1.4 Illustration depicting the development of the spine between the 3rd and the 10th WPC. 71

Figure 2.1 Photographs illustrating dissection of a human foetal spine - M662 (8 WPC). 99

Figure 2.2 H&E staining of a cohort of developing spines. 112

Figure 2.3 SAF-O staining of a cohort of developing spines. 118

Figure 2.4 Masson’s trichrome staining of a cohort of developing spines. 122

Figure 2.5 KRT8 immunostaining of a cohort of developing spines. 126

Figure 2.6 KRT18 immunostaining of a cohort of developing spines. 131

Figure 2.7 KRT19 immunostaining of a cohort of developing spines 136

Figure 2.8 CD24 immunostaining of a cohort of developing spines 140

Figure 2.9 GAL3 immunostaining of a cohort of developing spines. 144

Figure 2.10 CD55 immunostaining of a cohort of developing spines. 150

Figure 2.11 CTGF immunostaining of a cohort of developing spines. 155

Figure 2.12 BASP1 immunostaining of a cohort of developing spines. 160

Figure 2.13 T immunostaining of a cohort of developing spines. 165

Figure 2.14 Tie2 immunostaining of a cohort of developing spines. 171
Figure 2.15 CD90 immunostaining of a cohort of developing spines.

Figure 2.16 E-Cad immunostaining of a cohort of developing spines.

Figure 2.17 No primary antibody controls

Figure 2.18 Schematic representation of the protein staining in the developing human spine.

Figure 3.1 Experimental design to isolate notochordal from sclerotomal cells in the developing spine using the unique expression of KRT18 by notochordal cells.

Figure 3.2 Methodologies for RNA extraction from fixed-permeabilised-labelled-sorted cells.

Figure 3.3 Scatter plots depicting the gating utilised to identify and isolate KRT18^{+ve} and KRT18^{−ve} cells (example taken from sample M659 (15 WPC)).

Figure 3.4 Graphical explanation of IPA upstream regulator analysis.

Figure 3.5 Determination of ideal FITC-conjugated anti-KRT18 antibody concentration.

Figure 3.6 Analysis of FITC-conjugated anti-KRT18 antibody staining in different mixtures of MCF-7 and T/C28a2 cells.

Figure 3.7 Effect of different fixation and permeabilisation agents on the RNA quantity, quality and integrity.

Figure 3.8 Effect of antibody labelling on the RNA quantity, quality and integrity.

Figure 3.9 Effect of cell sorting on the RNA quantity, quality and integrity.

Figure 3.10 Microscopic view of foetal spine cells extracted from sample M644.
Figure 3.11 Sorting of sample M644.

Figure 3.12 Analysis of the RNA extracted from KRT18<sup>+</sup> and KRT18<sup>-</sup> sorted foetal spine cells.

Figure 3.13 Analysis of the cDNA amplified from RNA extracted from KRT18<sup>+</sup> and KRT18<sup>-</sup> sorted foetal spine cells.

Figure 3.14 qRT-PCR validation of KRT18 cell separation.

Figure 3.15 IPA of predicted interactions between the growth factors EGF, HGF and JAG1 and dataset genes.

Figure 3.16 HGF immunostaining of a cohort of developing spines.

Figure 3.17 c-Met immunostaining of a cohort of developing spines.

Figure 4.1 Schematic illustration of the methodology used to assess the ideal method to obtain viable CD24<sup>+</sup> cells.

Figure 4.2 Example of the gating to isolate viable CD24<sup>+</sup> and CD24<sup>-</sup> cells.

Figure 4.3 Flow cytometry overlay histograms depicting MCF-7 cells labelled with different concentrations of PE-conjugated anti-CD24 antibody and PE-conjugated anti-IgG1 isotype control at different concentrations.

Figure 4.4 Histograms and scatter plots used for Draq7® optimisation.

Figure 4.5 MCF-7 cells were encapsulated in collagen gels, incubated in Accumax® for different time periods and then labelled with Draq7® and PE-conjugated-anti-CD24 antibody or Draq7® and PE-IgG isotype control.

Figure 4.6 MCF-7 cells were encapsulated in collagen gels, incubated in type II collagenase for different time periods and then labelled with Draq7® and PE-conjugated-anti-CD24 antibody or Draq7® and PE-IgG isotype control.
**Figure 4.7** FACS plots of CD24 foetal samples used in the microarray analysis.

**Figure 4.8** Relative gene expression of CD24 and T in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells.

**Figure 4.9** Hierarchical clustering of the genes in the 5 CD24\(^{+}\)ve and 2 CD24\(^{-}\)ve samples analysed.

**Figure 4.10** PCA of the genes expressed by CD24\(^{+}\)ve and CD24\(^{-}\)ve samples.

**Figure 4.11** Volcano plot depicting the differential gene expression signals between CD24\(^{+}\)ve and CD24\(^{-}\)ve cells.

**Figure 4.12** Top growth factor signalling pathways that were significantly differentially expressed in CD24\(^{+}\)ve cells.

**Figure 5.1** Photomicrograph of a section through the IVD of sample M784 (18WPC).

**Figure 5.2** Relative gene expression of KRT18\(^{+}\)ve and KRT18\(^{-}\)ve markers in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from human embryonic and foetal developing spine cells.

**Figure 5.3** Relative gene expression of KRT18\(^{+}\)ve cell surface markers in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from human embryonic and foetal developing spine cells.

**Figure 5.4** H&E staining of an axial section through the IVD of specimen M784 (18 WPC).

**Figure 5.5** Relative gene expression of the KRT18\(^{+}\)ve and KRT18\(^{-}\)ve markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC).
**Figure 5.6** Relative gene expression of the KRT18^{+ve} cell surface markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC).

**Figure 5.7** ADORA3 immunostaining of a cohort of developing spines.

**Figure 5.8** Relative gene expression of CD24^{+ve} and CD24^{−ve} markers in CD24^{+ve} and CD24^{−ve} cells isolated from the same human embryonic and foetal specimens used in the CD24 microarray analysis.

**Figure 5.9** Relative gene expression of the CD24^{+ve} and CD24^{−ve} markers in CD24^{+ve} and CD24^{−ve} cells isolated from specimens M792 and M801.

**Figure 5.10** Relative gene expression of the CD24^{+ve} and CD24^{−ve} markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC).

**Figure 5.11** RTN1 immunostaining of a cohort of developing spines.

**Figure 5.12** PRPH immunostaining of a cohort of developing spines.

**Figure 5.13** MAP1B immunostaining of a cohort of developing spines.

**Figure 5.14** Box and whisker graph depicting the gene expression of GRB14, ADORA3, CDH6, ANPEP and CD69 in the adult NP.

**Figure 5.15** Scatter plots depicting the correlation between the gene expression of GRB14, ADORA3, CDH6, ANPEP and CD69 in the adult NP and the degeneration grade.

**Figure 5.16** ADORA3 immunostaining of adult NP tissues.

**Figure 5.17** Box and whisker graph depicting the relative gene expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 in the adult NP.
Figure 5.18 Scatter plots depicting the correlation between the expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 in the adult NP and the degeneration grade.

Figure 5.19 MAP1B immunostaining of adult NP tissues.
ABSTRACT

Back pain is a highly prevalent condition whose pathogenesis is associated with intervertebral disc (IVD) degeneration. Degeneration is driven by abnormal cell biology, particularly within the IVD’s inner core, the nucleus pulposus (NP). In recent years, there has been an ever-increasing search for cell-based therapies aimed at correcting the cell biology and thus repairing/regenerating the degenerate IVD. The success of these novel therapies, however, requires a thorough understanding of IVD development and of the phenotype of its cells. The embryonic, foetal and juvenile NP is populated by large vacuolated notochordal cells that with skeletal maturity are replaced by smaller NP cells. Since notochordal cells have been shown to display protective and anabolic roles in the IVD their loss in humans has often been suggested to initiate the degenerative process. As such, a detailed understanding of notochordal cells and their regulatory pathways may help identify factors involved in IVD homeostasis and aid the development of novel cell-based therapies targeting IVD degeneration. The study of human notochordal cells has, however, been hindered by ethical, logistical and technical difficulties in obtaining suitable samples and, as such, the human notochordal cell phenotype is, to date, unknown, constituting a major limitation in the field. The work presented here was conducted with the objective of developing a methodology to isolate human developing notochordal cells (NP progenitors) from adjacent sclerotomal cells (annulus fibrosus and vertebral body progenitors), to characterise the notochordal cell phenotype and identify potential factors involved in notochordal cell biology.
Initially, human embryonic and foetal spines were characterised to assess their suitability as a source of notochordal cells and to identify a notochord-specific marker that could be used to isolate notochordal cells for microarray studies. The human developing spine contained large vacuolated notochordal cells in all stages analysed (3.5-18 weeks post-conception (WPC)) that specifically expressed KRT8, KRT18 and KRT19 at all stages and CD24 between 5.5-18 WPC. KRT18 and CD24 were independently used to label notochordal cells (7.5-14 weeks post-conception) and separate them from sclerotomal cells. Methodologies were developed to allow extraction of RNA of sufficient quality for microarray analysis from fixed, permeabilised (in the case of KRT18) and/or, labelled and sorted cells (CD24). Microarray analysis identified and real-time qPCR and, for some markers, immunohistochemistry, validated GRB14, SLC19A1, FGF10, ADORA3, TBXA2R, CDH6, ANPEP, CD69, CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 as human notochordal cell markers. Ingenuity pathway analysis was performed to investigate the pathways/networks and upstream regulators and downstream effectors of notochordal cells. Inhibition of inflammation and angiogenesis were identified as relevant to notochordal cell biology, function and, possibly, to the known protective and anabolic role notochordal cells display in the IVD. Notochordal marker gene expression was identified in adult NP tissue, and negatively correlated with degeneration. Proteins encoded by ADORA3 and MAP1B were expressed by a proportion of adult NP cells, suggesting the presence of notochord-derived cells in the adult NP.

Importantly, this is the first study to detail a methodology and successfully isolate human notochordal cells. Such methodology has the potential to be used to culture and investigate the biology of viable human notochordal cells (CD24+ve). Future
studies aimed at developing cell-based therapies for IVD degeneration could also use these identified markers to assess appropriate stem cell differentiation to notochordal cells.
DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT

i) The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii) Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii) The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any
reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv) Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on presentation of Theses.
LIST OF ABBREVIATIONS

A2M – α-2-macroglobulin
ABC – Avidin-Biotin Complex
AC – Articular cartilage
ADAMTS – A Disintegrin And Metalloproteinase with Thrombospondin Motifs
ADORA3^ve – ADORA3 negative
ADORA3^{+ve} – ADORA3 positive
AF – Annulus fibrosus
ANXA3 – Annexin A3
AQP1 – Aquaporin 1
BASP1 – Brain abundant membrane attached signal protein
BMP – Bone morphogenetic protein
BSA – Bovine serum albumin
CA12 – Carbonic anhydrase XII
CD24^ve – CD24 negative
CD24^{+ve} – CD24 positive
cDNA – Complimentary DNA
COL2A1 – Collagen type II alpha 1
CS – Carnegie stage
CTGF – Connective tissue growth factor
CYTL1 – Cytokine-like 1
DAB – 3, 3’-diaminobenzidine
DMEM/F-12 – 1:1 Dulbecco’s modified Eagle medium/nutrient mixture F-12
Draq7<sup>−ve</sup> – Draq7 negative
Draq7<sup>+</sup>ve – Draq7 positive
DSC2 – Desmocolin-2
E-Cad – E-cadherin
ECM – Extracellular matrix
EDTA – Ethylenediaminetetraacetic acid
EGF – Epidermal growth factor
EMEM – Eagle’s minimal essential medium
ESC – Embryonic stem cell
FACS – Fluorescence activated cell sorting
FBLN1 – Fibulin 1
FCS – Foetal calf serum
FDR – False discovery rate
FGF – Fibroblast growth factor
FOXF1 – Forkhead box F1
FOXF2 – Forkhead box F2
GAL3 – Galectin 3
GDF – Growth differentiation factor
GLUT-1 – Glucose transporter 1
GPC3 – Glypican 3
H&E – Haematoxylin and eosin
HBB – Hemoglobin beta chain
HGF – Hepatocyte growth factor
HIF – Hypoxia inducible factor
IBSP – Integrin binding sialoprotein
IFN – Interferon
IGF – Insulin growth factor
IHC – Immunohistochemistry
IL – Interleukin
IL-1RA – Interleukin 1 receptor antagonist
IMS – Industrial methylated spirit
iPS – Induced pluripotent stem
IVD – Intervertebral disc
KRT8 – Keratin 8
KRT18 – Keratin 18
KRT18^ve – KRT18 negative
KRT18^+ve – KRT18 positive
KRT19 – Keratin 19
MAP1B^ve – MAP1B negative
MAP1B^+ve – MAP1B positive
MCF-7 – Michigan cancer foundation-7
MCT – Micro-centrifuge tube
MMP – Matrix metalloproteinases
MSC – Mesenchymal stem cell
N-Cad – N-cadherin
NCAM1 – Neural cell adhesion molecule
NP – Nucleus pulposus
NRP-1 – Neuropilin
OVOS2 – Ovostatin 2
PAX1 – Paired box 1
AUTHOR’S PREFACE AND ACKNOWLEDGMENTS

In 2007 I started a residency in Orthopaedic Surgery. I was eager to learn and to treat patients. I wanted to be a good doctor, I wanted to master surgery. I dedicated myself to learning anatomy, clinical diagnoses and surgical techniques and I attentively learnt from the best surgeons to make sure I didn’t miss anything, that I made no mistakes. Days were long and nights were short, but I was happy, I was learning my speciality. I gradually started developing a special interest for Spinal Surgery – in my opinion, the most challenging and demanding field in Orthopaedic Surgery.

Early in my surgical training I started being involved in research. I participated in a few clinical projects, started to learn statistics and published a few papers… but I always felt something was lacking, that I wasn’t doing enough, that I didn’t know how to do enough.

This eagerness to learn more and to pursue a higher goal found support in my Head of Department, Professor António Oliveira. He believed in me and encouraged to do what others thought would be a step back in my career and, over these years, his words of advice were always present. He is an example of leadership that I will always pursue.

In 2010 I embarked on a journey that is now about to end. I was awarded a PhD grant for medical doctors (Programme for the Advanced Medical Education) and spent 6 months learning the basics of research. Those were tough times, the demand was high and the expectations even higher. During this time I always had the support and well thought advice of Professor Leonor Parreira, who believed...
that an Orthopaedic Surgeon would succeed in the difficult world of research and would be able to make a difference. I was taught cellular and molecular biology, epidemiology, genetics, biostatistics, inflammation and repair. In all the modules I made myself think scientifically and tried to apply what I had learnt to practical scenarios within my speciality. That is when I realised that, in the very near future Biology would dictate much of what will be done in Orthopaedics. What surprised me the most was that, when treating disc degeneration related conditions, we often removed or fused the diseased intervertebral disc and that, whenever we attempted to replace it with a synthetic prosthesis the results were far from being optimal. There had to be a better, more physiological way!

That’s when I became familiar with the work of Professor Judith Hoyland and Dr Stephen Richardson on the phenotype of nucleus pulposus cells towards the development of cell-based strategies for intervertebral disc degeneration.

In April 2011 I emailed Professor Judith Hoyland and in the following month flew to Manchester for an interview; I started my PhD in October 2011.

These almost four years have been an amazing journey. The first months were extremely difficult. Everything took longer than I expected and there was so much repeating, going back to changing the smallest detail, working day after day to find out which small changes in temperature, incubation time, antibody concentration, centrifugation speed would allow me to give that next step into isolating and characterising the human notochordal cell. This, for a surgeon who’s used to treating a patient with a few hours of hard work, and to seeing results in a few days or weeks was very frustrating. But I never gave up. I always knew that I would achieve what I had set myself to and that this would help me become a
better doctor, understand better the diseases I treat and help researchers help us treating our patients.

This journey was made even more difficult because I was alone, in a foreign country, having left my wife, Frederica, back at home, in Portugal, and it would have never been possible without the help of some amazing people.

I would like to acknowledge my supervisor Professor Judith Hoyland for her dedication and continuous support, for guiding me through the maze that science often is and for always having a kind word of support. Her confidence, intelligence and sharpness will always be an example for me. I would also like to thank my supervisor Dr Stephen Richardson for having his office door always open for long scientific discussions, for always pushing me a little bit further and for never allowing me to give up – we phenotyped fixed-permeabilised-labelled-sorted cells!

I would also like to thank my advisor Professor Tony Freemont for grading the adult histological samples and for pushing me beyond my limits in my Literature Report assessment – it is an experience I will never forget and that I will fiercely avoid.

Learning all the techniques required for this PhD was a hard and long task that was only made possible due to the invaluable help from Mrs Pauline Baird, Mrs Sonal Patel and Mr Andy Fotheringham. Thank you Pauline for staying late and teaching almost everything I learnt, for the understanding and for the smile on your face; thank you Sonal for optimising the final antibodies and for the assistance with q-RT-PCR and thank you Andy for the assistance with sample processing and with the dissection. You were all very important and our tea breaks will never be forgotten.
I would also like to acknowledge the research fellow Dr Hamish Gilbert for the scientific support with qRT-PCR and for keeping me company in the lab during some of the weekends and the post-doc Dr Matthew Ablett for the advice on microarray analysis.

I would also like to thank Professor Neil Hanley, Dr Karen Piper-Hanley and Dr Andrew Berry for providing me with the foetal samples, Mr Michael Jackson for teaching me flow cytometry, Mr Michal Smiga for running the Agilent® bioanalyser analysis and Dr Leo Zeef for running the microarrays and giving me advice on their analysis.

I would also like to thank the PhD students Francesca, Shahnaz, Jude, Kim, Louise, Chris, Nicola with whom I shared an office for 2 and a half years. Some of you are Doctors already and others are on their way to becoming one and I am very proud to have had discussions with you and to have learnt from you.

Finally, I would like to thank my funding body, the Programme for the Advanced Education.

But because the times in Manchester were not just about work and were also about friendship I have to thank my dearest friend and housemate Matt Dale for making me feel at home from the first day, for always having time to listen to me, for the Friday night curries, for the beers in the Woodstock and in Folk – I am forever grateful for your friendship. I would also like to thank Mario for bringing me Portugal in Manchester.

But also at work I found great friends. Thank you Steve for the pints after work, for introducing me to your beautiful family and for all the parenthood advice. Thank you Hamish for the mornings at the gym and for the endless nights
clubbing, drinking and being happy. Thank you Matt for introducing me Mr Scruff and for being happy. Thank you Louise for being a bright light.

On a personal note, I have to thank my parents for giving me the tools that have made my career possible, for teaching me that hard work and dedication were the paths that would make me succeed in life and for never allowing me to give up. Their love and dedication are in every task I undertake. I would also like to thank my grandmother Avó Maria for being the strongest person in the world – you have succeeded in making your life bright despite the storm! To my little brother Eduardo for sometimes being bigger than me, for his cleverness and his integrity. The enormous personal accomplishment that this PhD is, however, would not have been possible without an enormous personal sacrifice that was being away from Frederica, my wife, for more than 2 years. It was her courage and never ending love and support that helped me through the years and the belief that our love would be stronger than the distance. You, Frederica, have made all this possible and your love for me has made distance bearable. Thank you for always being there in the most important occasions. I love you for your kindness and your self-respect, for being true to what you are and believe and for always wanting more.

Finally, it was during the time that I was conducting the research that is being brought to life in these pages, that Frederica brought to life Francisca, our beautiful daughter. Francisca was born out of the love, sacrifice and bravery that her mom and I had to overcome during the first 2 years of this PhD – and she is the greatest achievement of all!
CHAPTER 1

Introduction
1.1 BACK PAIN

Back pain is a major health problem with a point, one-month and lifetime prevalences respectively being 80, 11.9 and 23.2 percent (Hoy et al., 2012). Despite the enormous investment made in its diagnosis and treatment, the continuous population ageing is likely to increase its global prevalence even further (Deyo and Weinstein, 2001, Hoy et al., 2012). Costs associated with the treatment and those associated with work loss, insurance and disability costs mount up to £12 billion yearly in the UK (Maniadakis and Gray, 2000) and $85.9 billion in the USA (Martin et al., 2008). Using complex survey regression methods and analysing the nationally representative Medical Expenditure Panel Survey in the USA between the years of 1997 and 2005, Martin and colleagues estimated that the number of people suffering from back pain had increased by 20.7% during that period and that the overall health expenditure (adjusted for inflation) to treat back and neck pain had increased by 65%. However, and during the same period of time, the authors found that self-reported measurements of mental health, physical functioning, work, school and social limitations among adult patients with back pain had worsened (Martin et al., 2008), indicating that, despite the investment made in this field, clinical progress has been limited.

To date, the pathogenesis of back pain is not completely understood. It is believed to be multifactorial in origin, with factors such as heredity (Balague et al., 2012, Livshits et al., 2011, Patel et al., 2011), but also obesity, occupation, level of physical activity, smoking, alcohol consumption and diabetes (Balague et al., 2012, Macfarlane et al., 2009, Shiri et al., 2010, Tegeder and Lotsch, 2009) all
being implicated to various degrees. Studies analysing magnetic resonance imaging, have found that degeneration of the intervertebral disc (IVD) is associated with approximately 40% cases of back pain (Cheung et al., 2009, Livshits et al., 2011) and a relation of causality between IVD degeneration and back pain has also been shown (Bogduk, 1991, Freemont, 2009, Freemont et al., 1997).

As described by Kirkaldy-Willis, IVD degeneration is thought to be the trigger for a cascade of events that will be in the origin of most of the spinal conditions treated in medical practice – disc bulging and herniation, spinal stenosis with or without spondylolisthesis and degenerative scoliosis (Kirkaldy-Willis et al., 1978, Williams et al., 2011b) (Figure 1.1).

Figure 1.1 Disc degeneration cascade. (adapted from the concept of Kirkaldy-Willis (Kirkaldy-Willis et al., 1978)).
Pain arising from disc degeneration is manifested in several ways:

1. Inflammatory responses triggered by the herniated IVD material into the spinal canal (Ahn et al., 2002).

2. Nerve ingrowth into the IVD following annular and endplate ruptures (Freemont et al., 1997, Friberg and Hirsch, 1949).

3. Effects of altered biomechanics (Battie et al., 2007).
1.2 THE INTERVERTEBRAL DISC

The human spine is composed of five major regions: cervical, formed by 7 vertebrae (C1-C7); thoracic, formed by 12 vertebrae (T1-T12); lumbar, formed by 5 vertebrae (L1-L5); sacral, containing the sacrum, a single bone formed by the fusion of 5 sacral vertebrae, during adolescence; coccygeal, formed by the fusion of 4-6 small coccygeal vertebrae during adolescence. The cervical, thoracic, lumbar vertebrae and the 5th lumbar vertebrae and the sacrum are separated from each other by an IVD; each IVD and its upper and lower vertebrae constitute one spinal unit.

The IVD is responsible for shock absorption and mobility of the spinal unit (Hunter et al., 2003, Urban and McMullin, 1988). It is the largest avascular (Freemont, 2009, Moore, 2006, Raj, 2008) and aneural structure (Freemont et al., 1997) in the human body and is composed of three major structures (Figure 1.2).
1.2.1 Nucleus pulposus

The nucleus pulposus (NP) forms the gelatinous inner core of the IVD. It is composed of large quantities of the proteoglycan aggrecan within an irregular mesh of type II collagen fibres. The glycosaminoglycan side chains of aggrecan carry negative charges and generate swelling hydrostatic pressure that are responsible for maintaining disc hydration and function (Urban, 1996).

The cellular density of the NP is approximately $4 \times 10^6$ cells cm$^{-3}$ (Maroudas et al., 1975, Roughley, 2004). The morphology of NP cells varies between different animal species and, within the same species, with age and maturation state. In humans, the embryonic, foetal and immature NP is populated by large vacuolated notochordal cells that, soon after birth, become replaced by smaller, rounded cells that are frequently described as chondrocyte-like cells (Hunter et al., 2003) and...
will be here referred to as small NP cells. These morphological differences seen in humans are not common to all animal species, with some species retaining notochordal cells throughout life or at least during a proportionally longer period of their lifespan (Hunter et al., 2003).

1.2.2 Annulus fibrosus

The annulus fibrosus (AF) is a heterogeneous structure surrounding the NP. It is subdivided into outer AF, which is formed by distinct lamellae, composed of type I collagen fibres oriented obliquely between each lamellae (Cassidy et al., 1989, Marchand and Ahmed, 1990) and a less fibrous and less organized inner AF, characterized by a transition to type II collagen and increased proteoglycan content (Humzah and Soames, 1988). This architecture enables the AF to constrain the hydrostatic pressures generated within the NP upon compression, facilitating mobility between the spinal segments (Guerin and Elliott, 2007, Heuer et al., 2008, Johannessen et al., 2006, O'Connell et al., 2007).

The cellular density of the AF is approximately $9 \times 10^6$ cells cm$^{-3}$, with a higher density in the outer than in the inner AF (Maroudas et al., 1975, Roughley, 2004). Outer AF cells are fusiform fibroblast-like while inner AF cells are more rounded (Hunter et al., 2003).
1.2.3 Endplate

Endplates form the interface between the IVD and the upper and lower platforms of the vertebral bodies (VB) and are constituted by an osseous and a cartilaginous part. While the former is in contact with the VB, the latter is in contact with the NP and the AF and is usually termed cartilaginous endplate. Cartilaginous endplates are hyaline cartilage structures that protect inferiorly and superiorly the NP and AF and are responsible for diffusing oxygen and nutrients to and waste products from the NP. Their structure is formed by a network of type II collagen fibrils and proteoglycans. Adjacent to the vertebral body (VB) the endplate is richer in collagen, whereas adjacent to the NP it is richer in proteoglycans (Hamilton et al., 2006).

The cellular density of the endplate is approximately $15 \times 10^6$ cells cm$^{-3}$ (Maroudas et al., 1975, Roughley, 2004), a density that is similar to that found in articular cartilage (Pattappa et al., 2012). The cells populating the endplate are small and round, resembling chondrocytes found in articular cartilage and while subtle local variations in cell distribution have been described, no distinct layers exist (Moore, 2000).
1.3 INTERVERTEBRAL DISC DEGENERATION

It is the NP that is thought to be required for generation and maintenance of the disc’s structural integrity (Setton and Chen, 2006). Occupying only 1% of the IVD volume (Anderson et al., 2005), but with a vital role in maintaining IVD integrity are its cells, which synthesize their surrounding extracellular matrix (ECM) and also release catabolic and anabolic factors. An intricate balance between these factors is required to maintain the ECM homeostasis seen in healthy discs.

1.3.1 Pathogenesis of disc degeneration

IVD degeneration is characterised by a cell-driven imbalance between ECM synthesis and breakdown. In the NP, type II collagen is gradually replaced by type I collagen, which confers a more fibrous nature to this tissue. Proteoglycan synthesis, particularly aggrecan, also decreases (Buckwalter, 1995). Concurrently, there is an up-regulation of ECM degrading enzymes, namely MMP (matrix metalloproteinases) -1, -3, -7, -9, -10 and -13 and ADAMTS (A Disintegrin And Metalloproteinase with Thrombospondin Motifs) -1, -4, -5, -9 and -15 (Le Maitre et al., 2004, Pockert et al., 2009, Richardson et al., 2009).

The pro-inflammatory cytokine interleukin 1 (IL-1) is important for the normal function of the IVD, with its effects being controlled by its natural inhibitor, IL-1
receptor antagonist (IL-1RA) (Le Maitre et al., 2005a). During degeneration, however, there is up-regulation of IL-1, which is not counteracted by IL-1RA. This imbalance between the cytokine and its antagonist leads to many features common to those occurring in disc degeneration: up-regulation of ECM degrading enzymes (MMP and ADAMTS (Le Maitre et al., 2007d)), decreased aggrecan synthesis and replacement of type II by type I collagen (Le Maitre et al., 2007d), angiogenesis and neurogenesis into the normally avascular and aneural IVD (Voronov et al., 2007), and apoptosis (Zhao et al., 2007)). The importance of IL-1 regulation of the IVD is further confirmed by the fact that exogenous applications of IL-1RA to human IVD cells or tissue, reverses the molecular findings of degeneration (Hoyland et al., 2008, Le Maitre et al., 2007c).

Tumour necrosis factor-α (TNF-α), another pro-inflammatory cytokine, is also expressed by degenerate IVD cells, especially cells from herniated or prolapsed discs (Olmarker and Larsson, 1998), and has been implicated in the pathogenesis of nerve root damage and radicular pain (Igarashi et al., 2000). Its role in disc’s ECM degenerative changes is, however, not as clear. Although TNF-α expression increases with the degree of degeneration (Weiler et al., 2005), unlike with IL-1, studies have shown that IVD cells do not express its receptor (Le Maitre et al., 2007b), and its inhibitor fails to reverse ECM degradation molecular changes (Hoyland et al., 2008).

Together with the catabolic changes, the degenerating IVD also undergoes cellular changes, such as cell senescence and apoptosis. With IVD degeneration, there is an decrease in telomere length and an upregulation of the senescence markers P16INK4A and β-galactosidase (Le Maitre et al., 2007a); studies have shown that the senescence observed with IVD degeneration is more likely to be
stress-induced premature senescence (as seen by the expression of caveolin-1) rather than replicative senescence (Heathfield et al., 2008, Kepler et al., 2013).

Concurrently, it has also been shown that degenerated discs have increased cell death by apoptosis (Johnson and Roberts, 2007). There is consequently a cell number decrease; this cell number decrease has been proposed to be transient (Urban et al., 2001) and to be followed by increased cellular proliferation and cell cluster formation (Boos et al., 2002).

These ECM and cellular changes lead to reduced osmotic pressure, reduced hydration and ultimately impaired mechanical function (Boxberger et al., 2006, Costi et al., 2008). The NP eventually loses its ability to distribute the compressive forces between the VB, which are non-uniformly transferred to the AF, generating areas of increased pressure and subject to micro-trauma (O'Connell et al., 2009). This altered force distribution and micro-trauma results in tears and fissures along the AF and later in loss of disc height (Videman and Nurminen, 2004); in advanced stages of degeneration the NP herniates outwards through the AF fissures and neo-vessels and nerves grow inwards. Neo innervation and neo vascularisation into the normally aneural and avascular IVD are, therefore, two simultaneous features of IVD degeneration. Analysis of degenerate discs has identified nerves extending as far as the NP concurrently with the expression of substance P (a nociceptive neurotransmitter) by those isolated nerve fibres, therefore implicating this neo-innervation in the pathogenesis of back pain (Freemont et al., 1997, Richardson et al., 2009). Those nerves were furthermore accompanied by blood vessels expressing neural growth factor, another nociceptive neurotransmitter (Freemont et al., 2002). This neovascularisation has been shown in degenerate discs (Freemont et al., 1997,
Freemont et al., 2002) and in herniated tissues retrieved from surgeries (Yasuma et al., 1993) and has been shown to be driven by VEGF (Kokubo et al., 2008), IL-1β (Lee et al., 2011) and connective tissue growth factor (CTGF) (Ali et al., 2008). The inability to sustain and distribute load also induces micro-trauma to the cartilaginous endplate, which progressively undergoes calcification that will impair diffusion of nutrients to the IVD and lead to accumulation of waste products (Benneker et al., 2005).

Degeneration of the IVD impacts not only on this tissue, but also on the surrounding anatomical structures, such as spinal facets, spinal cord, nerve roots, spinal musculature and ligaments, which further impair IVD function leading to back pain (Daffner and Wang, 2009).

1.3.2 Aetiology of intervertebral disc degeneration

Although several factors have been implicated in the initiation and progression of disc degeneration, its exact aetiology has not been clearly defined yet. Mechanical load, glucose starvation, low pH and heritability are among the most studied events.
1.3.2.1 Mechanical load

Disc degeneration has long been assumed to be a product of “wear and tear” and, therefore, mechanical insults have always been linked to its initiation and progression (Luoma et al., 1998, Viikari-Juntura, 1997). This has been demonstrated at the molecular level, as cell survival and ECM synthesis are both sensitive to compressive stress (Maclean et al., 2004, Stokes and Iatridis, 2004, Walter et al., 2011). Mechanical loading, however, is not always detrimental as, although excessive loading results in localized tissue injury that alters strain distribution, physiological loads are required to induce diffusion and promote ECM synthesis (Stokes and Iatridis, 2004). In fact, twin studies have failed to establish causality between routine exposure to loading and disc degenerative changes (Battie et al., 2009). In vitro studies suggest that disc ECM anabolism is associated with increased magnitude of loading while its catabolism is more related to sustained periods of loading (Sowa et al., 2011).

1.3.2.2 Glucose starvation

The adult IVD, being practically devoid of blood vessels, is the largest avascular structure in the human body and, therefore, relies on diffusion through the endplate for nutrient and oxygen supply. There is consequently a gradient of nutrient and oxygen availability from the peripheral AF to the central NP, where nutrient and oxygen concentrations are the lowest (Bibby and Urban, 2004). As cells in the NP have to survive in oxygen-deprived conditions, they rely mostly on
glycolysis for cell metabolism and to produce energy. With degeneration, and as endplates calcify, glucose diffusion progressively decreases leading to up-regulation of catabolic and down-regulation of anabolic genes (Bibby and Urban, 2004, Junger et al., 2009, Neidlinger-Wilke et al., 2012, Urban et al., 2004)

1.3.2.3 Low pH

During glycolysis, glucose is metabolized into lactic acid, which is removed from the NP by diffusion through the endplate. With endplate calcification seen in IVD degeneration, however, the removal of lactic acid to the adjacent VB progressively decreases, resulting in its accumulation and, therefore, an acidic pH (Benneker et al., 2005). Kitano and colleagues have found that the pH in the NP of adult patients with diseased discs was 6.65±0.07 and that of asymptomatic patients with non-diseased discs was 7.14±0.05(Kitano et al., 1993). Studies in vitro have shown that NP cells cultured in an acidic medium display down-regulation of ECM anabolic factors such as aggrecan and type I and type II collagens (Neidlinger-Wilke et al., 2012).
1.3.2.4 Inherited predisposition

Inherited predisposition is also a contributing factor to the development of disc degeneration (Ikegawa, 2013, Videman et al., 1998, Videman et al., 2009). There is a significant correlation between the prevalence of back pain and the existence of reports of back pain in relatives (Bijkerk et al., 1999, Postacchini et al., 1988). In a population study including a database of over 2.4 million people, Patel and colleagues have identified a strong contribution of heredity to the development of symptomatic lumbar disc disease (Patel et al., 2011). Twin studies suggest that disc degeneration and its progression are highly genetically influenced with a heritability of 75% being found in Australian and UK women (Sambrook et al., 1999, Williams et al., 2011b) and approximately 30-55% in Finnish men (Battie et al., 2008). Several research groups have focused on identifying susceptibility genes that may be associated with disc degeneration. Results from those studies are summarised in Table 1.1.
Table 1.1 Genes whose polymorphisms have been associated with disc degeneration (adapted from (Mayer et al., 2013)).

<table>
<thead>
<tr>
<th>Gene category</th>
<th>Gene name</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECM genes</td>
<td>Aggrecan</td>
<td>(Roughley et al., 2006, Kawaguchi et al., 1999, Kim et al., 2011, Solovieva et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Asporin</td>
<td>(Song et al., 2008a)</td>
</tr>
<tr>
<td></td>
<td>Cartilage intermediate layer protein</td>
<td>(Seki et al., 2005, Virtanen et al., 2007b, Min et al., 2010, Min et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>Type I collagen</td>
<td>(Pluijm et al., 2004, Tilkeridis et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Type IX collagen</td>
<td>(Videman et al., 2009, Annunen et al., 1999, Jim et al., 2005, Higashino et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Type XI collagen</td>
<td>(Solovieva et al., 2006, Videman et al., 2009, Noponen-Hietala et al., 2003)</td>
</tr>
<tr>
<td></td>
<td>Fibronectin</td>
<td>(Anderson et al., 2010, Oegema et al., 2000)</td>
</tr>
<tr>
<td></td>
<td>Hyaluronan and proteoglycan link protein 1</td>
<td>(Urano et al., 2011)</td>
</tr>
<tr>
<td>ECM catabolic enzymes</td>
<td>MMP 1</td>
<td>(Valdes et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>MMP 2</td>
<td>(Song et al., 2008b)</td>
</tr>
<tr>
<td></td>
<td>MMP 3</td>
<td>(Dong et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Parkinson protein 2, E3 ubiquitin protein ligase</td>
<td>(Takahashi et al., 2001, Yuan et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>Proteosome subunit b type 9</td>
<td>(Williams et al., 2013)</td>
</tr>
<tr>
<td>ECM anabolic enzymes</td>
<td>Tissue inhibitors of MMP</td>
<td>(Valdes et al., 2005)</td>
</tr>
<tr>
<td>Pro-inflammatory</td>
<td>IL-1</td>
<td>(Eskola et al., 2010, Virtanen et al., 2007a, Solovieva et al., 2004)</td>
</tr>
<tr>
<td>cytokines</td>
<td>IL-6</td>
<td>(Eskola et al., 2010, Noponen-Hietala et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Cyclooxygenase-2</td>
<td>(Valdes et al., 2005)</td>
</tr>
<tr>
<td>Other</td>
<td>Vitamin D receptor</td>
<td>(Videman et al., 1998, Jones et al., 1998, Cheung et al., 2006, Valdes et al., 2005, Videman et al., 2001)</td>
</tr>
<tr>
<td></td>
<td>Growth differentiation factor 5</td>
<td>(Williams et al., 2011a)</td>
</tr>
</tbody>
</table>

While all the factors described above have been implicated, to various degrees, in the pathogenesis of IVD degeneration, their exact contribution to it is not clear. Complex arrays of genetic and environmental factors display a role and set the degenerative cascade in the IVD. To what extent they influence disc degeneration, their exact effect and how they interact with each other has not yet been clearly demonstrated.
There is still no definitive cure for disc degeneration-related clinical conditions. Although most of the established treatments are effective in relieving patient’s complaints and yield good clinical short- to mid-term results, they are not devoid of complications, especially in the long term. Most importantly, they aim at relieving symptoms and fail to address the genesis of the disease itself.

1.4.1 Conservative treatments

The first approach to the disease is generally conservative and consists of non-steroidal anti-inflammatory drugs, opioids, physical therapy and spinal manipulation. These, however, have little effect on the disease course and outcome (Balague et al., 2012) and, therefore, their efficacy remains a subject of debate (van der Roer et al., 2005).

1.4.2 Surgical treatments

When conservative treatments fail and a structural problem such as nerve root entrapment, spinal canal stenosis or deformity is identified, surgical treatment is indicated (discectomy, spinal fusion or disc arthroplasty).
Discectomy consists of the removal of the herniated or migrated disc material and aims at decompressing the spinal cord or nerve roots.

Fusion and arthroplasty (prosthetic disc replacement) are based on the concept that the pain arises from a diseased spinal segment, unable to perform its functions: load bearing and mobility. Fusion aims at immobilizing the vertebrae adjacent to the diseased disc. Published results demonstrate superiority of this treatment to non-operative treatments (Brox et al., 2003, Fairbank et al., 2005). It is, however, not devoid of complications: the immobilization of a spinal segment increases the stress placed on adjacent levels, potentially leading to their degeneration (Lund and Oxland, 2011) and the immobilization of long spinal segments reduces spinal mobility and impacts on the patient quality of life.

To obviate this problem and preserve spinal biomechanics, prosthetic disc replacements, such as those successfully used in bigger joints (hip and knee) have been developed. Although initially disappointing, the results of the newer implants are currently being published (Hellum et al., 2011, Zigler et al., 2007) and seem to be superior to fusion, triggering a new interest in this treatment (Fairbank, 2011). They are, however, not devoid of complications, with prosthetic disc’s migration, extrusion or failure being among the most frequent (Errico, 2005).
1.4.3 Biological treatments

Due to the fact that the aforementioned treatments address the disease symptoms and not disease itself, recent research has focused on possible biologic approaches that can tackle the disease process and, therefore, either prevent or reverse degeneration of the IVD.

Biological treatments can either aim to regenerate or repair the IVD. Regeneration is the process of reviving existing disc tissue. It consists on the introduction of growth factors (Masuda, 2008), gene transfection (Sobajima et al., 2004) or cell transplantation (Sakai, 2008) into the degenerate disc and relies on the principle that the cellular population within it can be rescued to a non-degenerate state. Regeneration is ideally aimed at early stages of degeneration, mostly when it is still confined to the NP.

When the disease has progressed to more advanced stages, biological regeneration may not be feasible, and repair may be required. The primary method for this repair is through tissue engineering, in which cells are implanted within a biodegradable biomaterial support. These cells will subsequently generate new, normal disc tissue and maintain its integrity and function over extended periods. This treatment offers the potential of a long-term solution to disc degeneration and low back pain (Figure 1.3).
1.4.3.1 The ideal cell for NP tissue engineering

The ideal cell for NP implantation would have to be able to survive in its niche and in the biomaterial in which it was implanted and to lay down a functional ECM that would mimic or even improve the characteristics of the native tissue. Several researchers have focused on identifying the most suitable cell source for NP tissue engineering. While some authors advocate implanting fully mature NP cells (autologous or allogeneic), others favour the use of cells in a more immature state (stem cells) differentiated towards a NP-like phenotype.
1.4.3.1.1 Autologous nucleus pulposus cells

Although autologous NP cell implantation would appear a good solution to repopulate the disc and while studies have shown this can retard disc degeneration (Hohaus et al., 2008, Meisel et al., 2007), concerns have been raised by this approach. As the IVD is relatively hypocellular (Liebscher et al., 2011, Maroudas et al., 1975) a single biopsy would not be sufficient to obtain enough cells for expansion. This would be further complicated by the fact that the method used to harvest disc cells (needle puncture) has been shown to induce degeneration in healthy discs and accelerate degeneration in degenerate ones (Carragee et al., 2009, Michalek et al., 2010, Nassr et al., 2009). Additionally, degenerate disc cells display increased senescence (Le Maitre et al., 2007a), increased expression of ECM catabolic and degrading enzymes (Le Maitre et al., 2004, Le Maitre et al., 2005a) and decreased expression of ECM components (Pearce et al., 1987, Sive et al., 2002), factors that would make them functionally inadequate for transplantation.

1.4.3.1.2 Allogeneic nucleus pulposus cells

An alternative to autologous cells would be the use of allogeneic cells. Nomura and colleagues have transplanted either allogeneic NP tissue or allogeneic NP cells into a rabbit model of disc degeneration with both treatments demonstrating slowing of the degenerative process (as depicted by the intensity of staining for type II collagen) without inducing graft-versus-host response. They hypothesized
that the relative avascularity of this tissue would confer an immunoprivileged state to the IVD (Nomura et al., 2001). The use of such cells or tissue, however, would require a donor bank of healthy human samples, which would be difficult to obtain.

1.4.3.1.3 Stem cells

Stem cells have been proposed as the ideal cell source for regenerating diseased and aged tissues due to their ability to self-renew and to produce specialised progeny (Lutolf et al., 2009).

1.4.3.1.3.1 Embryonic stem cells

Embryonic stem cells (ESCs) have the potential to become any cell type in the human body and, as many of the diseases that most afflict society are diseases of cellular deficiency, their use raises high expectations. Their pluripotency, however, is also a reason of concern, as the plasticity that allows them to become any cell is also difficult to control. Undifferentiated ESCs have the potential to form teratomas (tumours derived from the three germ layers). Additionally, the fact that these cells are obtained from human embryos raises ethical concerns that have been difficult to overcome (Murry and Keller, 2008). To date, no studies have shown differentiation of ESCs to NP cells. Using a combination of growth factors, an in vitro study has shown differentiation of mouse ESCs to cells expressing node/ notochordal markers (Winzi et al., 2011).
1.4.3.1.3.2 Induced pluripotent stem cells

In a groundbreaking study by Takahashi and Yamanaka, mature somatic cells (fibroblasts) were reprogrammed to generate pluripotent cells (Takahashi and Yamanaka, 2006). These induced pluripotent stem (iPS) cells have all the properties of ESCs in adopting a specific phenotype but are not embryo-derived, thus not raising the same ethical concerns. As they offer the potential of an autologous treatment, for example dermal fibroblasts being induced to pluripotency and then differentiated to NP cells, they do not raise immunological concerns. Vierbuchen and colleagues maintained the enthusiastic pace of the research in this field by direct reprogramming between cells of different progeny. Importantly, as these cells do not go through an undifferentiated state, they are thought to have no teratogenicity (Vierbuchen et al., 2010). Using the knowledge from these studies, Chen and colleagues found that a CD24 positive (CD24\textsuperscript{+ve}) enriched population of mouse iPS cells cultured in a laminin-rich culture medium and in the presence of low oxygen tension and notochordal cell conditioned medium differentiated towards cells with a NP-like phenotype (Chen et al., 2013). In another study, Liu and colleagues found that human iPS cells cultured with porcine NP tissue could differentiate towards a healthy NP-like phenotype (Liu et al., 2014). These promising studies open an avenue for the research in this field, as the generation of healthy NP-like cells may be used to further investigate the behaviour of these cells, the pathogenesis of degeneration and to develop new therapies. Both studies have limitations, however, particularly the markers chosen to assess NP cell differentiation – a matter of current debate that will be reviewed in section 1.5.
1.4.3.1.3.3 Adult stem cells

Adult stem cells are present in niches located next to specific adult tissues and maintain their ability to differentiate. They are already primed to generate specialised cell types, therefore lacking the plasticity of ESC or iPS cells, but they are not tumorigenic and there are no concerns related to their clinical use.

Mesenchymal stem cells (MSCs) are mesoderm-derived adult stem cells, which are present in many tissues (bone marrow (Pittenger et al., 1999), adipose tissue, umbilical cord, muscle, dermis, periosteum, synovial membrane, synovial fluid and cartilage (Barry and Murphy, 2004, Sonoyama et al., 2006)). In 2006, the International Society for Cellular Therapy published a position statement in which they defined the minimal criteria for defining MSCs: MSCs should be adherent to plastic in standard culture conditions, express CD105, CD73 and CD90, and lack the expression of CD45, CD34, CD14 or CD11b, CD79α or CD19 and HLA-DR and differentiate in vitro to osteoblasts, chondroblasts and adipocytes (Dominici et al., 2006). MSCs are capable of differentiation to several connective tissue lineages such as cartilage (Barry and Murphy, 2004, Caplan, 2007, Noel et al., 2002), bone (Arinzeh, 2005, Hong et al., 2006, Noel et al., 2002), fat (Barry and Murphy, 2004, Helder et al., 2007), ligament (Sonoyama et al., 2006, Trubiani et al., 2005) and muscle (Barry and Murphy, 2004). Furthermore, there is a growing body of evidence confirming that they can be differentiated to adult NP cells (Helder et al., 2007, Henriksson et al., 2009b, Minogue et al., 2010a, Richardson et al., 2008, Richardson et al., 2006b, Risbud et al., 2004, Sakai et al., 2005, Stoyanov et al., 2011, Strassburg et al., 2012, Strassburg et al., 2010), suggesting they may be useful for novel cell-based tissue engineering therapies.
1.5 THE SEARCH FOR THE IDEAL CELL PHENOTYPE

While cell-based therapies offer a promising and exciting future for IVD repair and regeneration, for them to be successful, it is paramount that implanted cells have the correct phenotype to replace native disc cells and to produce an appropriate functioning ECM \textit{in vivo}. While culture conditions and growth factors can be used to direct stem cell fate (Lutolf et al., 2009, Murry and Keller, 2008), the conditions or factors necessary for differentiation towards a NP cell have not yet been clearly identified. Furthermore, to assess stem cell differentiation to a cell capable of repairing or regenerating the NP it is essential to know the desired end-stage (target cell) phenotype.

1.5.1 The nucleus pulposus versus the articular cartilage cell phenotype

Many studies have attempted to differentiate stem cells, particularly MSC, to NP cells. Those studies, however, have assessed differentiation by analysing the expression of chondrogenic genes or markers, such as collagen type II α1 (COL2A1), aggrecan (ACAN) and sex determining region Y (SRY)-box 9 (SOX9) (Henriksson et al., 2009b, Le Maitre et al., 2009, Richardson et al., 2006a, Richardson et al., 2008, Tao et al., 2008), that are also known to be expressed by healthy adult human NP cells (Sive et al., 2002). However, despite the similarities between the NP and articular cartilage (AC), there are important differences
between these two tissues and the cells that reside in them. Such differences have been shown \textit{in vivo} when the transplantation of rabbit’s auricular cartilage into its IVD led to the formation of a solid tissue, which resembled AC and lacked the osmotic and gelatinous nature of the NP (Gorensek et al., 2004). Indeed, while the AC and NP are both load-bearing connective tissues, whose cells reside in similar hypoxic, acidic niches, there are also important differences between both tissues:

- **Morphology**: While the chondrocytes that populate AC are small and round, NP cells can be (depending on the animal species and on the maturation stage) large and vacuolated (notochordal) or small and round, resembling AC chondrocytes (small NP cells).

- **Ontogeny**: AC cells derive from the lateral plate mesoderm; while the ontogeny of NP cells is not yet clearly understood, it has been suggested that, at least a subpopulation of the NP cells populating the adult IVD are notochord (axial mesoderm) derived (Choi et al., 2008, McCann et al., 2012).

- **Aggrecan/ type II collagen ratio**: While aggrecan and type II collagen are the main constituents of the AC and NP, its ratio is 27/1 in the NP and 2/1 in the AC (Mwale et al., 2004).

- **Collagen network**: In the AC, type II collagen is organised to form a rigid network, while in the NP is organised to form a loose network.

- **ECM proteoglycans**: The proteoglycans in the AC ECM can be organised to form large aggregates (hyaluronic acid and central filaments), multiple monomers and large non-aggregated monomers. The proteoglycans in the NP are organised in short non-aggregated monomers and in clusters of monomers without apparent central filaments (Buckwalter et al., 1989).
• **Mechanical loads**: While the AC is subjected mostly to compressive loading, the NP is subjected to compressive and shear loading (Iatridis et al., 1996).

• **Biomechanical behaviour**: AC behaves like a viscoelastic solid in response to shear transient and dynamic deformation, while the NP behaves like a fluid under transient deformation and like a viscoelastic solid under dynamic deformation (Iatridis et al., 1996).

The understanding that the NP cell differs from the AC and that a thorough understanding of the NP cell and its phenotype is paramount for the development of cell-based therapies for IVD degeneration, has driven researchers to investigate the ideal phenotype to aim for in IVD regeneration strategies. Two important questions arise from these investigations:

1. What is the NP cell phenotype and how does that phenotype differ between animal species?
2. Is the ideal target cell the one that populates the adult NP, or should research focus on the cells populating the NP at earlier developmental stages, where animal and *in vitro* studies have shown that they are relatively resistant to degeneration?

Studies addressing these questions will be reviewed in sections 1.5.2 and 1.6.
1.5.2 The nucleus pulposus cell phenotype: interspecies variations

The search for the NP phenotype has been a topic of intensive research for the last decade with numerous laboratories focusing a substantial amount of work on this topic. Some authors have investigated markers that are representative of the NP microenvironment and that may be a reflection of the NP physiology and metabolism. Within these, hypoxia inducible factor 1 (HIF-1), a transcription factor involved in glycolysis, the main metabolic pathway utilised by NP cells, and glucose transporter 1 (GLUT-1), a transporter that is responsible for glucose uptake by cells, were found to, together with MMP-2 and vascular endothelial growth factor (VEGF), have higher expression in rat NP in comparison with its adjacent AF and cartilaginous endplate (Fujita et al., 2008, Rajpurohit et al., 2002). The importance of HIF-1 for NP physiology has been further highlighted by a recent study in which the conditional deletion of this gene in the mouse notochord led to NP cell death and its replacement by a fibrocartilaginous tissue lacking the biomechanical properties of the NP (Merceron et al., 2014). These findings have led to these markers being recently included in a consensus statement from the Orthopaedic Research Society Spine Research Interest Group as NP markers relevant to the normal regulation of these cells (Risbud et al., 2014). However, one can also argue that these markers reflect physiological adaptation of these cells to the microenvironment in which they reside rather than marking a distinct cellular phenotype.

Some of the more important advances in this field have been achieved by unbiased molecular phenotyping studies, in which large-scale gene expression profiling approaches were used to distinguish the NP from similar or
neighbouring tissues, such as AC, AF and endplate. Those studies have allowed the identification of markers that could otherwise have been missed and also to highlight important differences and similarities between different species. While such markers may not have a previously described role in the IVD, they likely reflect some of the functions of NP cells. Importantly, such studies have shed light over a long lasting debate regarding the ontogeny of the cells populating the adult NP and this will be discussed later in this introduction.

The first study investigating the molecular phenotype of NP cells used microarrays to compare the genes expressed by 8 week-old rat NP with those expressed by the AF, tendon, skeletal muscle, skin, blood, bone marrow, spinal cord and brain. This study identified and proposed CD24 as an NP marker due to its higher differential expression in the NP than in the other tissues analysed (Fujita et al., 2005). Further confirming the specificity of this marker to the rat NP, a different study identified it, together with, n-cadherin (N-Cad), neuropilin 1 (NRP-1), brain abundant membrane attached signal protein (BASP1), CD221 and CD155, as having a higher differential expression in the rat NP in comparison with AF cells (Tang et al., 2012). Finally, another study investigated the rat NP phenotype by using microarrays to compare its gene expression with that of the AF and AC, having identified and validated at the protein level keratin 19 (KRT19), pleiotrophin (PTN), annexin A3 (ANXA3) and glypican 3 (GPC3) as rat NP markers; CD24 was also highly differentially expressed in the rat NP but did not meet the cut-off for protein validation (Lee et al., 2007).

Using a larger animal model, Sakai and colleagues compared the phenotype of the dog’s NP, AF and AC cells and identified keratin18 (KRT18), neural cell adhesion molecule (NCAM1) and α-2-macroglobulin (A2M) as NP markers.
Another microarray study compared the phenotype of bovine NP, AF and AC cells. When comparing the NP with the AF and AC, sclerostin domain containing 1 (SOSTDC1), keratin 8 (KRT8), KRT18 and keratin 19 (KRT19), synaptosomal-associated protein 25 (SNAP25), and N-Cad were found to have higher differential expression in the NP. When comparing the NP and AF with the AC, KRT8, KRT18, BASP1, N-Cad, forkhead box F1 (FOXF1), TNFα induced protein 6 (TNFAIP6), forkhead box F2 (FOXF2), tenomodulin (TNMD), aquaporin 1 (AQP1), SNAP25, and SOSTDC1 were found to have higher differential expression in the IVD tissues. Fibulin 1 (FBLN1) and integrin binding sialoprotein (IBSP) were found to have lower differential expression in the cow’s NP, being classified as negative NP markers (Minogue et al., 2010b).

A study from the same group compared the gene expression profile of healthy adult NP tissue with that of its AC. The authors identified and validated (using qRT-PCR) the genes hemoglobin beta chain (HBB), carbonic anhydrase XII (CA12), FOXF1, paired box 1 (PAX1) and ovostatin 2 (OVOS2) as having higher differential expression in the NP and cytokine-like 1 (CYTL1), IBSP and growth differentiation factor 10 (GDF10) as having higher differential expression in the AC (Minogue et al., 2010a). Using a similar strategy, Power and colleagues compared the gene expression of the human NP, AF and AC tissue, and confirmed the differential expression of the aforementioned NP marker CA12 as a cell surface marker with a higher differential expression in the NP than in the other tissues analysed (Power et al., 2011).

Table 1.2 summarises the markers identified in these studies.
Table 1.2 NP markers identified in different animal species.

<table>
<thead>
<tr>
<th>Species</th>
<th>Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>KRT19, GPC3, ANXA3, PTN, CD24, KRT8, N-Cad</td>
</tr>
<tr>
<td>Dog</td>
<td>A2M, KRT18, NCAM1, DSC2</td>
</tr>
<tr>
<td>Bovine</td>
<td>KRT8, KRT18, KRT19, N-Cad, SNAP25, SOSTDC1, FOXF1, FOXF2</td>
</tr>
<tr>
<td>Human</td>
<td>FOXF1, OVOS2, HBB, CA12, PAX1, KRT18, KRT19</td>
</tr>
</tbody>
</table>

It is clear from the analysis of Table 1.2 that there are marked phenotypic differences between species. There are several potential reasons for these differences:

1. The height, thickness and shape of the IVD vary greatly between rats, dogs, cows and humans. It is therefore expected that waste product removal and nutrient intake from and to the NP (performed through diffusion though the endplate (Bibby and Urban, 2004)) to be much harder in the larger and thicker human IVD, possibly leading to it being subjected to much more stressful conditions (acidic pH due to lactic acid accumulation, hypoxia and glucose starvation) than, for example, the rat IVD. This is supported by the fact that different studies using the same animal models often identified the same markers (CD24 for the rat and CA12 for the human).

2. The mechanical forces applied through the IVD of the quadruped and small rat, the larger quadrupeds dog and cow and the bipedal human are very different in terms of magnitude and direction (Alini et al., 2008) and such differences will certainly influence the microenvironment in which NP cells reside.
3. In humans, cows and chondrodystrophic dogs (such as those used in the study by Sakai and colleagues (Sakai et al., 2009)) notochordal cells are replaced by small NP cells soon after birth, while rats retain notochordal cells through most of their lifespan (Miyazaki et al., 2009). This difference in cell morphology has important implications for tissue engineering strategies to repair the degenerate IVD and will be the main focus of the following sections of this Introduction.

All these differences will likely impact on the phenotype of the cells populating the NP in the different animal species. Therefore, and although the understanding of each animal’s NP phenotype is important to understand disease processes in that specific animal, to formulate hypotheses for studies in humans and even design and interpret experimental treatments in a given animal species, translation of such findings to human research should be done with care.

Despite the interspecies differences, some genes were also found to be common to more than one species. KRT8 was identified as a rat and bovine NP marker, KRT18 as a dog, bovine and human NP marker and KRT19 as a rat, bovine and human NP marker. NCAD was identified as a rat and dog NP marker and FOXF1 as a bovine and human NP marker. It is interesting that KRTs, which together with cadherins (e-cadherin (E-Cad) and n-cadherin (N-Cad)) have previously been reported to be expressed by the human embryonic notochord (Gotz et al., 1995) and chordoma (notochord-derived tumours) (Gottschalk et al., 2001), were identified as NP-specific markers in all species, independently of whether they retain notochordal cells or not. It is possible that, while some markers are more indicative of the niche in which NP cells reside and, therefore, differ between
species, others, such as KRTs, may correspond to a signature profile common to several animal species and that is maintained independently of the microenvironment, anatomical and morphology differences between the different animals.
1.6 THE ROLE OF NOTOCHORDAL CELLS IN THE INTERVERTEBRAL DISC

As aforementioned, the morphology of the cells populating the NP differs between different animal species and, within the same species, between different developmental stages. While some animals retain notochordal cells throughout most of their lifespan, in others notochordal cells are replaced by small NP cells soon after birth. While notochordal cells are large (25-85µm), occur in clusters, have well-demarcated Golgi and “immature” mitochondria associated with large endoplasmatic reticulum and are filled with vacuoles, small NP cells are much smaller (10µm) and contain no vacuoles (Hunter et al., 2004).

The morphology of the cells populating the NP has important implications for the understanding of IVD physiology and degeneration due to:

1. **Observations in humans:** In humans, the disappearance of notochordal cells to be replaced by small NP cells at the end of the first decade of life correlates with the appearance of the first histological signs of IVD degeneration (Walmsley, 1953).

2. **Observations in animals:** Like humans, certain breeds of dogs (named chondrodystrophic dogs) have notochordal cells replaced by small NP cells and are reported to develop spinal disorders, IVD herniation and degeneration, whereas non-chondrodystrophic dogs, who retain notochordal cells until much later in life, are relatively resistant to degeneration (Rufai et al., 1995, Stevens et al., 2000). Studies investigating the content of the NP of these animal species found that dogs
which retain notochordal cells have a higher NP proteoglycan and water content whereas those of chondrodystrophic dogs are much more fibrous (Bray and Burbidge, 1998a, Bray and Burbidge, 1998b) – this loss of NP hydration is one of the major features of IVD degeneration.

3. *In vitro* experiments: Notochordal cells *in vitro* produce more proteoglycans than small NP cells (Cappello et al., 2006, Miyazaki et al., 2009); when cultured together, notochordal cells stimulate NP cells to produce more proteoglycans themselves (Aguiar et al., 1999, Erwin and Inman, 2006); this stimulatory effect of notochordal cells on small NP cells has been attributed to soluble factors secreted by notochordal cells (Oegema, 2002) and to be potentiated in hypoxic cultures (Erwin et al., 2009). Finally, while IL-1 induced small NP cell death *in vitro*, addition of notochordal cells prevented this effect (Erwin et al., 2011).

4. Experiments in animals: Puncturing of the mouse IVD induces degeneration, which is accompanied by a morphological change from large vacuolated to small NP cells (Yang et al., 2009).

For these reasons, a notochordal NP has frequently been associated with a healthier and more hydrated IVD, whereas small NP cell-rich NP has been associated with a more fibrous IVD, which is more prone to degeneration. This has led to a growing interest in understating notochordal cell function, morphology and phenotype. Since in humans notochordal cells are present during the embryologic and foetal stages section 1.7 will describe the development of the IVD.
1.7 EMBRYOLOGIC AND FOETAL DEVELOPMENT OF THE IVD

The structures forming the IVD are first seen in the human embryo after the 3rd week post-conception (WPC), during gastrulation. In gastrulation, the inner cell mass rearranges itself and originates the ectoderm, mesoderm and endoderm. These three germ layers will give rise to all the tissues and organs (Nakaya and Sheng, 2008). Towards the end of the 3rd WPC, mesoderm, which will give rise to most connective tissues in the body, becomes sub-divided according to its relation to the centre of the embryo into axial mesoderm (or notochord), paraxial mesoderm (or somites), intermediate mesoderm and lateral plate mesoderm (Solnica-Krezel and Sepich, 2012) with the two most central being fundamental for IVD and VB formation. The central notochord extends along a midline cranio-caudal localisation, ventral to the neural tube, and with a row of somites on each side (Figure 1.4, week 3) and provides longitudinal support to the embryo, directs neural tube formation (Greene and Copp, 2009) and induces paraxial mesoderm segmentation into pairs of cranio-caudally distributed somites (Mead and Yutzey, 2012). While being mesoderm-derived, at this stage notochordal and somite cells display epithelioid characteristics such as cell polarity and cell-cell contact (Hay, 2005).

The 4th WPC marks the beginning of organogenesis, in which the cells acquire the characteristics of the tissues and systems they will originate (Muller and O'Rahilly, 2004). Somite cells will be induced by notochordal cells to lose their epithelioid characteristic and to migrate: somite cells closer to the embryo axis will migrate
centrally and become sclerotomal cells, which will originate the AF and the VB; somite cells localised further away from the embryo axis will migrate laterally and become dermomyotomal cells, which will originate the dermis and muscle (Hay, 2005). The central notochord and sclerotome will be in the origin of the spinal skeleton (Figure 1.4, week 4).

The 5th and 6th WPC are characterised by the condensation of sclerotomal cells around the notochord and neural tube and by the migration of dermomyotomal cells further away from the midline (Peacock, 1951). At the end of the 6th WPC, somite segmentation will be very evident in the sclerotomal cells and highly condensed sclerotomal cell regions (AF anlagens) intercalate with looser sclerotomal regions (VB anlagens) (Figure 1.4, weeks 5-6).

Between the 7-9th WPC the cells occupying the less condensed regions will exert osmotic pressure on the centrally located notochord, pushing its cells to the adjacent segments, where sclerotomal cells will condense even further to accommodate all migrating notochordal cells (Figure 1.4, weeks 7-9). This process finalises around the 10th WPC, in which the loosely organised sclerotomal cells have occupied all their segments and the notochordal cells become restricted to the adjacent segments, encircled by densely organised sclerotomal cells. While the former sclerotomal cells will become VB cells, the later will become AF cells (Aszodi et al., 1998, Peacock, 1951) (Figure 1.4, week 10). At this stage remnants of the notochordal sheath may still be present in the VB anlagen segments.

From the 11th WPC onwards, all the main spinal structures have a morphology and shape that resembles the tissue they will originate in the fully formed skeleton. VB anlagen cells are rounded and larger and will later become hypertrophic (to ossify after birth); AF anlagens are localised to the periphery of the adjacent
segments and have a characteristic lamellar shape; in the IVD centre are the vacuolated notochordal NP cells.

Figure 1.4 Illustration depicting the development of the spine between the 3rd and the 10th WPC. The notochord and neural tube are two rod like structures that occupy the centre of the embryo at the 3rd WPC. In the 4th WPC the notochord induces somite cell migration towards the midline (sclerotomal cells) and to the periphery (dermatomyotomal cells). Sclerotomal cells condense around the notochord and neural tube and adopt a segmented pattern (5-6th WPC). Between the 7-9th WPC sclerotomal cells in the VB regions expand towards the midline, and push notochordal cells away from its centre; in the adjacent segments, sclerotomal cells condense at their periphery to accommodate the notochordal cells that are being pushed away from the VB anlagens. From the 10th WPC onwards, notochordal cells are restricted to the central NP region of the IVD, being encircled by AF cells in their periphery. Superiorly and inferiorly, sclerotomal cells will occupy the entire segment which will later become the VB (Rodrigues-Pinto et al., 2014).
1.8 ONTOGENY OF THE CELLS POPULATING THE ADULT NUCLEUS PULPOSUS

As the morphology of the cells populating the human embryonic, foetal and immature NP differs from that of the fully mature NP, authors have for long debated the origin of the cells populating the adult NP. Some argue that notochordal cells die and are replaced by cells migrating from adjacent tissues (Butler, 1989), such as the endplate (Kim et al., 2003) or the AF (Butler, 1989).

Other authors, however, state that adult NP cells are notochordal in origin and have changed their morphology as an adaptation to the IVD niche. This is supported by the aforementioned study by Yang and colleagues where needle-puncture induced notochordal cell morphology change to small NP cells (Yang et al., 2009) and also because chordomas can differentiate into chondrocyte-like cells, expressing aggrecan and type II collagen (Gottschalk et al., 2001). Such differences in morphology and size in cells with the same ancestry could be due to differences in metabolic rates, cell cycle status or signal variations in RNA/DNA ratios (Risbud et al., 2010).

This controversy has been elucidated by two studies in which the recombinase gene Cre was knocked into the mice notochordal markers sonic hedgehog (Shh) (Choi et al., 2008) or Noto (McCann et al., 2012) in mice. By mating them with reporter mice the authors in both studies were able to track notochordal cells during the embryonic and foetal development and until maturity and found that all NP cells in all stages analysed were notochord-derived. Further supporting that the adult NP contains notochord-derived cells is the aforementioned fact that adult
bovine NP cells (Minogue et al., 2010b), chondrodystrophic dog NP cells (Sakai et al., 2009) and both the non-degenerate and degenerate human adult NP (Minogue et al., 2010a, Rutges et al., 2010a, Stosiek et al., 1988) – all known to be predominantly or even exclusively populated by small NP cells – express KRT8, KRT18 and KRT19 or brachyury (T), which have been reported to be expressed by the embryonic notochord (Gotz et al., 1995), and to be highly expressed in the notochordal NP of the young rat (Lee et al., 2007). This was further confirmed when notochordal and small NP cells were separated by size from the bovine NP and were found to have an overlapping expression of several genes, including that of the transcription factor T, often also associated with the notochordal phenotype (Minogue et al., 2010b, Rodrigues-Pinto et al., 2013).

While these studies confirmed the presence of notochord-derived cells in the adult NP that, in some species, such as the human, have undergone differentiation to small NP cells, others have suggested that this population of notochord-derived adult NP cells may co-exist with a sub-population of morphologically identical cells with a different ontogeny. Gilson and colleagues found that the expression of KRT8 in the adult bovine IVD was restricted to a subset of cells, found within a morphologically similar population of KRT8 negative (KRT8^{ve} cells) (Gilson et al., 2010). Using human autopsy samples, Weiler and colleagues identified KRT8, KRT18 and KRT19 and also Galectin 3 (GAL3) (also expressed by the embryonic notochord (Gotz et al., 1997)) in a large proportion of NP cells in patients under the age of 30, but restricted to subsets of clustered cells in older patients (Weiler et al., 2010). Further supporting the presence of a sub-population of cells in the adult NP, Tanaka and colleagues induced IVD degeneration in transgenic mice
with GFP-tagged notochordal cells, and identified a second population of small round GFP-negative cells in the NP (Tanaka et al., 2012).

Parallel to the controversy of the ontogeny of adult NP cells and of the number of cells populating the NP, several groups have also focused on identifying, within the adult IVD, tissue specific stem cells that could be recruited to repair the degenerate IVD. Using the panel of MSC markers validated by the International Society for Cell Therapy, Blanco and colleagues identified and isolated within human degenerate discs a population of cells capable of differentiating along the chondrogenic and osteogenic lineages (Blanco et al., 2010) and Risbud and colleagues identified a similar population of cells that were capable of differentiating along the chondrogenic, adipogenic and osteogenic lineages (Risbud et al., 2007). Henriksson and colleagues identified, near the AF borders, a population of cells that were capable of undergoing cell proliferation, hypothesising that they could play a role in IVD function (Henriksson et al., 2009a). Sakai and colleagues identified a population of multipotent cells (Tie2 and GD2 positive) in the rat and human NP that were capable of differentiating along mesenchymal lineages and of expressing a NP-like phenotype (Sakai et al., 2012). While these studies identifying cells with stem cell properties in the adult IVD are interesting and may indicate that a resident population of cells may be recruitable to regenerate the IVD, none of the studies described the ontogeny of such cells and, importantly, if they were notochord- or sclerotome-derived.
1.9 THE NOTOCHORDAL CELL PHENOTYPE

Due to the complexity of the NP cell, to develop cell-based therapies to repair or regenerate the degenerate IVD it is important to understand the population of the human adult NP, and particularly, if different sub-populations of cells co-exist in the same tissue, which is their ontogeny and whether they play different roles in the NP. Such study could in theory be possible by replicating the aforementioned mice fate mapping studies in which embryonic notochordal cells were labelled and tracked along their development until maturity. This would allow an understanding of the fate of notochordal cells in the human NP, the existence of sub-populations of NP cells and, if so, the role of each cell subtype in IVD physiology and disease. Due to obvious ethical reasons, however, such this study will never be possible in humans.

An alternative would be to isolate and characterise the phenotype of notochordal cells and possibly identify notochord-specific markers that could be useful to assess the presence and even isolate notochord-derived cells in the adult NP. If, as suggested by animal and in vitro studies, notochordal cells or factors they produce are the ideal cell type to aid in regenerating the degenerate IVD, the identification of the notochordal cell phenotype and of specific notochordal markers would constitute a large step towards understanding and treating IVD degeneration.

To date, several groups have attempted to identify notochordal cell markers. The strategies used have been by comparing the phenotype of notochordal rat NP cells with that of its costal cartilage (Leung et al., 2010), the phenotype of the NP tissue from rats of different ages (Chen et al., 2009, Chen, 2007, Tang et al., 2012), the
phenotype of immature notochordal pig NP and AF cells (Chen et al., 2009, Gilchrist et al., 2011), the phenotype of sub-populations of pig NP cells separated by size and granularity (Chen et al., 2006b), immature tissue and isolated cells of rat, porcine and human IVD (Chen et al., 2009) and the phenotype of human juvenile (scoliotic) and adult NP tissue (Tang et al., 2013). While the markers identified in those studies can be specific to the animal model in which they were performed, they are also helpful in understanding the notochordal cell phenotype. These studies identified and proposed T, BASP-1, CD221, and NRP-1, (Tang et al., 2012), CD55 (Leung et al., 2010), N-Cad (Leung et al., 2010, Tang et al., 2012), transforming growth factor β (TGF-β), bone morphogenetic protein 6 (BMP-6) and connective tissue growth factor (CTGF) (Chen, 2007) as rat notochordal markers and T, CD24 and CD54 (Tang et al., 2013) as human notochordal/ immature markers. CD90 was not expressed either in immature rat (Tang et al., 2012) or in scoliotic juvenile NP (Tang et al., 2013) and has been proposed as a negative notochordal cell marker.

There are, however, important caveats to these studies and to their conclusions. Firstly, and as discussed above, there are important anatomical differences between the rat and the human spine and IVD which impact on their cell’s microenvironment and, possibly, phenotype. Secondly, Tang and colleagues failed to show the presence of notochordal cells in the juvenile discs used in the study and therefore it is not clear if the juvenile phenotype described is that of the human notochordal cell. Thirdly, the use of discs from children undergoing surgery for scoliosis may also be criticised as IVDs from scoliotic patients are subjected to eccentric loads and are prematurely degenerated which may impact
on the gene expression of their cells (Rajasekaran et al., 2010, Hristova et al., 2011).

Interestingly, Minogue and colleagues identified the rat notochordal marker BASP1 in the non-notochordal NP of adult humans, with its expression decreasing with degeneration (Minogue et al., 2010a). Unpublished results from the host laboratory also suggest that the rat notochordal marker CD55 may also be an adult human NP marker (50 fold higher differential expression in the NP than in the AF or AC). Finally, and as aforementioned, KRT8, KRT18, KRT19, GAL3 and T, which have all been identified in the NP of animals that retain notochordal cells, have also been identified in the human adult NP.

All these data suggest that, while notochordal cell gene expression may differ between species, it is also possible that some notochordal markers identified in animal models may also be specific to human notochordal or notochord-derived cells, independently of their morphology, maturation and disease state. The knowledge of such markers would allow to “track” and possible identify notochord-derived cells in human IVDs from different maturation and disease states and to understand the fate and function of notochordal and notochord-derived cells in the human IVD.

Tables 1.3, 1.4, 1.5, 1.6 and 1.7 depict the genes identified in these studies and discuss their relevance to the IVD field.
<table>
<thead>
<tr>
<th>GENE</th>
<th>DESCRIPTION OF THE PROTEIN ENCODED AND RELEVANCE TO THE IVD BIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Noto</strong> (notochord homeobox) (McCann et al., 2011)</td>
<td>Transcription factor whose expression is restricted to the organiser node and the nascent notochord during gastrulation and axis elongation. It regulates node morphogenesis, notochord ciliogenesis and left-right patterning.</td>
</tr>
<tr>
<td><strong>Shh</strong> (sonic hedgehog) (Choi et al., 2008)</td>
<td>Signalling molecule required for patterning in the early embryo and that is highly expressed in the embryonic notochord. It is hypothesised to act in the induction of the floor plate and in neural tube patterning.</td>
</tr>
</tbody>
</table>
Table 1.4 Rat notochordal/ immature NP markers.

<table>
<thead>
<tr>
<th>GENE</th>
<th>DESCRIPTION OF THE PROTEIN ENCODED AND RELEVANCE TO THE IVD BIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>T (Tang et al., 2012)</td>
<td>Embryonic transcription factor that regulates the transcription of genes required for mesoderm formation and differentiation. It is localised to notochord-derived cells where it is thought to mediate cartilage development in the developing embryo, and also in chordomas.</td>
</tr>
<tr>
<td>BASP1 (Tang et al., 2012)</td>
<td>Membrane-bound protein initially described as being specific to neuronal and spinal cord cells, where it participates in neurite outgrowth and synaptic plasticity, but later found to be also expressed by human endothelium, developing mammary gland, kidney, testis, and lymphoid tissues.</td>
</tr>
<tr>
<td>BMP-6 (Chen, 2007)</td>
<td>Belongs to the BMP family of signalling molecules that are implicated in inducing bone growth and dorsal axis patterning.</td>
</tr>
<tr>
<td>CD24 (Tang et al., 2013)</td>
<td>Cell surface protein expressed by mature granulocytes and many B cells. It is highly expressed by chordomas but not by chondrosarcomas (Fujita et al., 2005).</td>
</tr>
<tr>
<td>CD55 (Leung et al., 2010)</td>
<td>Membrane protein involved in the regulation of the complement cascade, which is broadly distributed among haematopoietic and non-haematopoietic cells.</td>
</tr>
<tr>
<td>NRP-1 (Tang et al., 2012)</td>
<td>Transmembrane glycoprotein involved in neuronal guidance and angiogenesis. This ligand, together with its receptor (semaphorin) are highly expressed in the healthy human NP, but their expression decreases with degeneration, having therefore been proposed to act as barriers to neuronal ingrowth into the healthy IVD (Tolofari et al., 2010).</td>
</tr>
<tr>
<td>CTGF (Chen, 2007)</td>
<td>Cysteine-rich secretory protein that mediates ECM interactions. It is secreted by non-chondrodystrophic dog’s notochordal cells and has been shown to up-regulate proteoglycan synthesis by small bovine NP cells (Erwin, 2008).</td>
</tr>
<tr>
<td>TGF-ß2 (Chen, 2007)</td>
<td>Extracellular glycosylated protein. It has a vital role during embryonic development and suppresses the effects of interleukin dependent T-cell growth. Its expression in IVD cells (together with the expression of TGF-ß1, TGF-ß2 and their receptors) have been shown to decrease with mice ageing (Matsumaga et al., 2003).</td>
</tr>
<tr>
<td>Integrin sub-units α3, α6 and β4 (Chen et al., 2009, Gilchrist et al., 2011)</td>
<td>Integrins are expressed by NP cells and are responsible for their adhesion to laminins.</td>
</tr>
<tr>
<td>CD239 (Chen et al., 2009, Gilchrist et al., 2011)</td>
<td>CD239 is expressed by NP cells and is responsible for their adhesion to laminins.</td>
</tr>
</tbody>
</table>
Table 1.5 Pig notochordal/ immature NP markers.

<table>
<thead>
<tr>
<th>GENE</th>
<th>DESCRIPTION OF THE PROTEIN ENCODED AND RELEVANCE TO THE IVD BIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laminin (LM511, LM322)</td>
<td>Laminins are heterodimeric proteins involved in cell adhesion, migration, proliferation, differentiation and survival.</td>
</tr>
<tr>
<td>Integrin sub-units α1 and α6</td>
<td>See table 1.4.</td>
</tr>
<tr>
<td>Integrin sub-units α3, α6 and β4</td>
<td>See table 1.4.</td>
</tr>
<tr>
<td>CD239</td>
<td>See table 1.4.</td>
</tr>
</tbody>
</table>
Table 1.6 Bovine notochordal/immature NP markers.

<table>
<thead>
<tr>
<th>GENE</th>
<th>DESCRIPTION OF THE PROTEIN ENCODED AND RELEVANCE TO THE IVD BIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td>T (Leung et al., 2010)</td>
<td>See Table 1.4.</td>
</tr>
<tr>
<td>N-Cad (Leung et al., 2010)</td>
<td>Calcium dependent cell–cell adhesion glycoprotein expressed during gastrulation and that is required for the establishment of neuronal polarity.</td>
</tr>
<tr>
<td>BASP1 (Leung et al., 2010)</td>
<td>See Table 1.4.</td>
</tr>
<tr>
<td>KRT8 (Leung et al., 2010)</td>
<td>Keratins are typically expressed by epithelial cells but are found in a wide range of tissues including the developing notochord.</td>
</tr>
<tr>
<td>KRT18 (Leung et al., 2010)</td>
<td>KRT8 is involved in maintaining cellular structural integrity and in signal transduction and cellular differentiation.</td>
</tr>
<tr>
<td>KRT19 (Leung et al., 2010)</td>
<td>KRT18 typically dimerises with KRT8 to form an intermediate filament in simple single-layered epithelial cells. KRT19 is the smallest known acidic cytokeratin. It is found in the periderm, a transient structure that functions as a protective layer for the embryonic skin. Keratins 8 and 18 are involved in resistance against TNF-α-induced apoptosis in hepatocytes (Caulin et al., 2000). Since TNF-α is implicated in nerve ingrowth into the IVD (Hayashi et al., 2008), and hence in back pain pathogenesis (Freemont et al., 1997), they may act as regulators of this ingrowth. Since intermediate filaments are usually present in cells subjected to mechanical stress it is possible that, in the IVD, these proteins act to withstand the hydrostatic pressures in the NP (Hunter et al., 2003).</td>
</tr>
</tbody>
</table>
Table 1.7 Human notochordal/immature NP markers.

<table>
<thead>
<tr>
<th>GENE</th>
<th>DESCRIPTION OF THE PROTEIN ENCODED AND RELEVANCE TO THE IVD BIOLOGY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CD24</strong> (Tang et al., 2013)</td>
<td>See Table 1.4.</td>
</tr>
<tr>
<td><strong>CD54</strong> (Tang et al., 2013)</td>
<td>Cell surface glycoprotein expressed by endothelial cells and cells of the immune system. In cultured IVD cells, its expression is upregulated by the exposure to IL-17, IFNγ, and TNFα (Gabr et al., 2011).</td>
</tr>
<tr>
<td><strong>CD55</strong> (Leung et al., 2010)</td>
<td>See Table 1.4.</td>
</tr>
<tr>
<td><strong>IGF1R/CD221</strong> (insulin-like growth factor 1 receptor)</td>
<td>Receptor for insulin-like growth factor, which has been found overexpressed in malignant tissues where it functions as an anti-apoptotic agent by enhancing cell survival. It is expressed by human adult NP and inner AF cells in both non-degenerate and degenerate discs (Le Maitre et al., 2005b). A single nucleotide polymorphism in this gene has been associated with disc degeneration (Urano et al., 2008).</td>
</tr>
<tr>
<td><strong>KRT18</strong> (Minogue et al., 2010a, Weiler et al., 2010)</td>
<td>See Table 1.6.</td>
</tr>
<tr>
<td><strong>KRT19</strong> (Minogue et al., 2010a, Weiler et al., 2010)</td>
<td>See Table 1.6.</td>
</tr>
<tr>
<td><strong>Integrin sub-units α3, α6 and β4</strong> (Chen et al., 2009, Gilchrist et al., 2011)</td>
<td>See Table 1.4.</td>
</tr>
<tr>
<td><strong>CD239</strong> (Chen et al., 2009, Gilchrist et al., 2011)</td>
<td>See Table 1.4.</td>
</tr>
</tbody>
</table>
1.10 SUMMARY

The high prevalence of back pain and its association with IVD degeneration has led to a growing interest in the development of cell-based therapies that could be used to repair or regenerate the degenerate IVD. Since IVD degeneration is initiated by cell-driven ECM changes in the NP, a significant amount of research has focused on its cells.

Although for a long time NP cells were perceived as being similar to AC cells, there are significant differences between both tissues that impact on the cell’s behaviour, metabolism and phenotype. The molecular differences between those cells have been addressed by phenotyping studies that have elucidated the NP phenotype. Such studies have been performed using several animal models and human tissue. Importantly, they highlighted that, although some genes are common to two or more species, there are also important interspecies differences that should be understood when translating findings from animal studies to human research. It is possible that such differences are a reflection of the anatomical, postural and developmental singularities of each animal species.

NP phenotyping studies also highlighted that genes that had previously been associated with a notochordal cell phenotype were present in the NP of animals devoid of cells with the characteristic large vacuolated notochordal cell morphology. That, together with fate mapping studies in mice elucidated a long lasting controversy regarding the ontogeny of NP cells, suggesting that NP cells are derived from the embryonic and foetal notochord.
As such all these findings suggest that the adult NP contains notochord-derived cells, which in some animal species have undergone a morphological change from large vacuolated to small NP cells. While the reason for this morphology change is not clear, since is it species- and age-specific, it is possible that it occurs as an adaptation of notochordal cells to the microenvironment in the IVD of each species or at different developmental stages (such as progressive endplate calcification leading to reduced nutrient intake and waste product removal from the NP or failure to sustain mechanical load with advancing age). It is not known how the morphology change impacts on the cell phenotype and function.

Furthermore, studies suggest that notochord-derived cells may co-exist in the NP with a second population of morphologically identical cells that have migrated from adjacent tissues, such as the VB, endplate or AF.

It is possible that the small round morphology adopted by the cells populating the adult human NP (whether it is a single population of notochord-derived cells or two populations of notochord and sclerotomal derived cells) may itself reflect the unique physicochemical microenvironment in which the cells reside. If, as suggested in the study by Tanaka and colleagues, the cell migration to the NP occurs as a response to an insult inducing degeneration (Tanaka et al., 2012), it remains to be elucidated whether those migrating cells are involved in the degeneration process or in an attempt to repair.

Clarification on this subject is important as histological studies in humans, observations in animals and in vitro and animal experiments have suggested that notochordal cells confer resistance to IVD degeneration and could therefore be the ideal cells to repair or regenerate the degenerate IVD.
Thus, a more thorough understanding of notochordal cells would allow a better understanding of basic biological processes regulating IVD function and would also allow an improved knowledge of its degeneration. This, however, will only be possible once the human notochordal cell phenotype is thoroughly characterised. A few groups have attempted to characterise the notochordal cell phenotype in animal species that retain such cells. However, to date, human notochordal cells have not been isolated and their gene expression has never been characterised.
1.11 HYPOTHESES AND OVERALL AIMS OF THE PROJECT

The hypotheses for this project were that:

1. The human developing spine is composed of notochordal and sclerotomal cells.
2. Human notochordal cells can be identified by the expression of unique markers and separated from sclerotomal cells using fluorescence activated cell sorting (FACS).
3. Human notochordal cells have a unique molecular phenotype that is distinct from that of sclerotomal cells. The identification/characterisation of their transcriptome and proteinome will allow:
   a. An understanding of the phenotype of human notochordal cells
   b. Investigation of the pathways, regulators and mechanisms leading to notochordal cell differentiation.
   c. Assessment of the expression of notochordal cell markers in the adult NP and how their expression changes with degeneration.

To address these hypotheses the aims of the study were to:

1) Characterise the tissue morphology of the human embryonic and foetal spine and determine its suitability as a source of human notochordal cells.
2) Identify, using immunohistochemistry, a notochord-specific marker that could be used to label and isolate (using FACS) notochordal cells from their adjacent sclerotomal cells.
3) Characterise the human notochordal cell phenotype using microarrays to identify:
   a. Notochord-specific markers.
   b. Cell surface markers that can be used to isolate viable notochordal cells.
   c. The pathways, networks, upstream regulators and downstream functions of notochordal cells.

4) Validate the identified markers at the gene and protein level.

5) Analyse the presence of notochord-specific markers in IVD tissue harvested from patients undergoing surgery for IVD degeneration-related conditions and correlate notochordal gene expression with degeneration grade.

In 2010 Professor Judith Hoyland and Dr Stephen Richardson initiated a collaboration with Professor Neil Hanley and his research group to obtain human spinal tissue from embryos and foetuses donated to science by women undergoing pregnancy terminations. Such specimens were obtained and collected within 2-4 hours after pregnancy termination. These samples provided a unique opportunity to obtain and study human notochordal cells, which are reported to be present at these developmental stages.

The work undertaken analysing such samples and the results presented in this PhD thesis represent, to the author’s knowledge, the first study to ever isolate human notochordal cells, to characterise their phenotype and to investigate the pathways, networks, upstream regulators and downstream functions of such cells. Due to the small size of the samples and the difficulties in separating notochordal
cells from their adjacent sclerotomal tissues, even using microsurgical techniques, a significant proportion of the work undertaken was to develop methodologies that would allow the isolation of a pure population of notochordal cells.
1.12 THESIS STRUCTURE

The results obtained from the investigation are detailed in chapters 2, 3, 4 and 5. The work in each chapter however was often conducted simultaneously, rather than in a sequential manner.

In chapter 2, immunohistochemistry was used in a cohort of developing spines to identify a notochord-specific marker that could be used to label and sort notochordal cells using FACS. While the identification of 4 notochord-specific markers will be described in this chapter – KRT8, KRT18 and KRT19 having an intracellular localisation and CD24 being localised to the outer cell membrane – the identification of CD24 occurred at a later stage of investigation, subsequent to the development of an isolation procedure utilising KRT18.

For this reason, and while the immunohistochemical studies that led to the identification of CD24 as a notochord-specific markers continued, the first marker to be used to isolate notochordal cells was KRT18 – this is described in chapter 3. The continuous pursuit to identify a cell surface marker was justified by the complex and time consuming methodology needed to extract high quality RNA from cells that had previously been fixed, permeabilised, labelled with the intracellular marker KRT18 and sorted. This methodology, however, was developed, optimised and used to isolate notochordal cells in an attempt to identify potential novel cell surface markers that would obviate many of the technical difficulties posed by KRT18 labelling and also to isolate viable notochordal cells. However, due to complex technical limitations and the time gap in sample acquisition, only one foetal specimen was used in chapter 3 and all
the cells and their RNA were depleted in the microarrays and before any validation was possible.

The immunohistochemical identification of CD24 as a notochord-specific cell surface marker occurred while the microarray analysis of KRT18 sorted cells was still being undertaken. The isolation, phenotypic characterisation and microarray analysis of notochordal cells using CD24 are described in chapter 4.

Chapter 5 describes the validation of the markers identified in chapters 3 and 4. Since all samples sorted using KRT18 were depleted at this stage, notochordal cells isolated using CD24 were used for gene expression validation of the markers identified using both methodologies. Protein validation of the same markers was performed in a cohort of developing spines. Finally, a panel of markers was selected and their expression in adult NP tissues was investigated and correlated with degeneration.
CHAPTER 2

Identification of a human notochord-specific marker

Ricardo Rodrigues Pinto
2.1 INTRODUCTION

Knowledge of the development of the human NP and its phenotype underpins the understanding of the role of this important structure in maintaining the IVD homeostasis and the development of stem cell based therapies targeting its degeneration. This, however, has been hindered by limitations in obtaining human samples, particularly during the first trimester of gestation, when the IVD forms. Thus, current knowledge is, to date, mostly restricted to studies in animals, particularly mice and rat (Bedore et al., 2013, Choi et al., 2008, Dahia et al., 2012, Dahia et al., 2009, DiPaola et al., 2005, Hayes et al., 2011, Maier et al., 2013, Tang et al., 2012).

During embryonic and foetal development, the human NP is composed of large vacuolated notochordal cells encircled by smaller non-vacuolated sclerotomal cells, which will later give rise to the AF and to the VB with its cartilaginous endplate [reviewed in (Rodrigues-Pinto et al., 2014)]. Soon after birth, however, the notochordal NP cell population gradually becomes replaced by a population of small non-vacuolated chondrocyte-like NP cells (Hunter et al., 2003). While recent research suggests that notochord-derived cells may persist in the adult NP, despite having acquired a different morphology, it is unclear whether they coexist with a second population of morphologically identical cells – possibly migrated from the neighbouring sclerotome-derived tissues (Gilson et al., 2010, Minogue et al., 2010b, Minogue et al., 2010a, Weiler et al., 2010). Furthermore, it is not known if this morphological change is accompanied by either a phenotypic or functional change of these cells.
Importantly, *in vitro* and animal studies have shown that notochordal cells and small NP cells are functionally distinct, with notochordal cells having a fundamental role in IVD homeostasis, as they produce more proteoglycans than small NP cells (Cappello et al., 2006), stimulate small NP cells during co-culture to produce a healthier ECM (Aguiar et al., 1999, Erwin et al., 2006) and prevent IL-1-induced cell death in small NP cells (Erwin et al., 2011). Furthermore, cadaveric studies have shown that notochordal cell loss in humans correlates with the appearance of the first histological signs of degeneration (Walmsley, 1953). As such, their loss with maturity in humans has been suggested to initiate the degenerative process.

To clarify this, identification of unique developmental notochordal markers is needed. In order to identify such markers several studies have investigated the notochordal / immature NP cell phenotype (reviewed in chapter 1 and summarised in Tables 1.3, 1.4, 1.5, 1.6 and 1.7) with KRT8, KRT18, KRT19, T, GAL3, CD24, CD55, BASP1, CTGF and E-Cad being proposed as notochordal/ immature NP markers, Tie2 as a NP progenitor cell marker and CD90 as negative NP marker.

Most of this data, however, derives from studies in animal models (mouse, rat, pig and cow) and, therefore, translation to human research should be done with caution, as vertebrate embryogenesis is not fully conserved between species, and differences in phenotype, function and developmental pathways have been shown to exist in other systems (Irie and Kuratani, 2011, Richardson et al., 1997) and also in the fully mature IVD (Minogue et al., 2010b, Minogue et al., 2010a, Rodrigues-Pinto et al., 2013).

The identification of unique human notochordal cell markers would allow researchers to trace the fate of notochordal cells during human IVD development,
maturation and degeneration and to understand if, despite having acquired a
different morphology, notochord-derived cells still persist in the adult human NP.
To date, however, such studies have not been conducted and this is a major
limitation in the field.
2.2 HYPOTHESES AND AIMS

The hypotheses for this study were that:

1. The human embryonic and foetal spine contains morphologically identifiable notochordal cells.
2. Notochordal cells have a phenotype that is distinct from the adjacent sclerotomal cells.

To address the hypotheses the aims for this study were to:

1. Characterise the morphology of the human spine/IVD between the developmental stages of 3.5 to 18 WPC.
2. Analyse the protein expression of KRT8, KRT18, KRT19, GAL3, CD24, CD55, BASP1, T, CTGF, CD90, Tie2 and E-Cadherin in the developing spine.
2.3 MATERIALS AND METHODS

2.3.1 Sample acquisition and staging

Human embryonic and foetal samples were obtained with approval from the local ethical committee (Manchester Royal Infirmary, Ref. No: 08/H1010/28 Early Pregnancy Tissue Collection) and with full informed consent following medical or surgical pregnancy termination (Table 2.1 details the samples used). Embryonic staging was performed according to the Carnegie stage (CS) classification and converted to weeks post-conception (WPC); foetal staging was estimated by hand and foot length measurements and converted to WPC (Bullen P, 1997, O'Rahilly and Muller, 2010).
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Stage (WPC)</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>M234</td>
<td>3.5 (CS10)</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M320</td>
<td>5.5 (CS16)</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M396</td>
<td>6</td>
<td>Histology</td>
</tr>
<tr>
<td>M741</td>
<td>7</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M607</td>
<td>7</td>
<td>Histology</td>
</tr>
<tr>
<td>M779</td>
<td>7.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M680</td>
<td>7.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M669</td>
<td>7.5</td>
<td>Histology</td>
</tr>
<tr>
<td>M490</td>
<td>7.5</td>
<td>Histology</td>
</tr>
<tr>
<td>M494</td>
<td>8</td>
<td>Histology</td>
</tr>
<tr>
<td>M736</td>
<td>8</td>
<td>Histology</td>
</tr>
<tr>
<td>M744</td>
<td>8</td>
<td>Histology</td>
</tr>
<tr>
<td>M682</td>
<td>8.5</td>
<td>Histology</td>
</tr>
<tr>
<td>M414</td>
<td>8.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M397</td>
<td>9</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M782</td>
<td>9.5</td>
<td>Histology</td>
</tr>
<tr>
<td>M636</td>
<td>9.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M483</td>
<td>10</td>
<td>Histology</td>
</tr>
<tr>
<td>M685</td>
<td>10</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M565</td>
<td>10</td>
<td>Histology</td>
</tr>
<tr>
<td>M415</td>
<td>10.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M425</td>
<td>11</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M404</td>
<td>11.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M742</td>
<td>12</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M439</td>
<td>12</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M394</td>
<td>12.5</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M776</td>
<td>13</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M739</td>
<td>14</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M409</td>
<td>14</td>
<td>Histology</td>
</tr>
<tr>
<td>M777</td>
<td>17</td>
<td>Histology and immunohistochemistry</td>
</tr>
<tr>
<td>M784</td>
<td>18</td>
<td>Histology and immunohistochemistry</td>
</tr>
</tbody>
</table>
2.3.2 Human embryonic and foetal spine dissection

Samples were processed within 2-4 hours of acquisition. Due to the small size of the younger embryos (3.5- 5.5 WPC) no attempt was made to dissect the developing spine anlagens from the remaining tissues and they were mounted as a whole. In embryos with 6 WPC and older and in foetuses (8-18 WPC), the whole spine (vertebrae and IVD) was dissected from the adjacent tissues. Dissection was performed under sterile conditions, using microsurgical instruments and a stereomicroscope (Stem 2000, Carl Zeiss®), as follows. First, the spines together with their adjacent tissues (ligaments, ribs and spinal cord) were carefully dissected from the embryo/foetus and transferred to a Petri dish containing PBS (Figure 2.1 A-C); second, using microsurgical forceps and scissors, the ribs (at their costovertebral joints) and the spinal cord, were gently separated from the spine (Figure 2.1 D and E); third, the anterior and posterior longitudinal ligaments were gently separated from the spine; finally, the resulting whole foetal spine containing the vertebrae and intervertebral discs (Figure 2.1 F) was washed in phosphate buffered saline (PBS) (Dulbecco’s PBS without Ca&Mg, PAA Laboratories®, H15-002).
Figure 2.1 Photographs illustrating dissection of a human foetal spine - M662 (8 WPC). A: photograph of the undissected spine immediately after harvesting from the foetus. B: Posterior and C anterior view of the undissected spine depicting the ribs (+), the spinal cord (-) and the spine (=). D: dissection of the ribs from the spine, E: detachment of the spine from the spinal cord and ribs and F: view of a dissected spine depicting the vertebral bodies (VB) and intervertebral discs (IVD); note the more condensed sclerotomal cell region in the centre of the developing VB (arrow).

2.3.3 Sample preparation for immunohistology

Embryonic and foetal spines were fixed immediately after harvest in 4% (w/v) paraformaldehyde (PFA, Sigma®, 36148) in PBS at 4°C for 24 hours, decalcified in 20% ethylenediaminetetraacetic acid (EDTA), pH 7.4 (Tennaquest®) at room temperature for 3 days and washed in running tap water for 1 day. Decalcified foetal spines were processed overnight inside a histology cassette on a tissue processor (Thermo Electron Corporation®). For processing, the cassette was sequentially immersed in 50% industrial methylated spirit (IMS, Fisher Scientific®, M/4450/17) (90 minutes), 70% IMS (60 minutes), 99% IMS (60
minutes), 99% IMS (60 minutes), 99% IMS (60 minutes), 99% IMS (60 minutes),
99% IMS (60 minutes), xylene (Fisher Scientific, X/0200/17) (90 minutes),
xylene (90 minutes), xylene (90 minutes), molten wax (90 minutes) and molten
wax (120 minutes). After processing, the cassettes were placed in molten wax in a
vacuum oven (Binder®) for 2x 30 minutes for adequate wax penetration into the
tissue. Finally, samples were embedded in molten wax and placed on ice to
solidify. A microtome (Shandon Finesse 325, Thermo Scientific®) was used to
cut 5µm sections.

The 3.5 WPC whole-mounted specimen was sectioned transversally while all
other samples were sectioned longitudinally along the whole spinal length.
Sections were briefly placed in a water bath at 42°C to allow for sections to flatten
and then mounted on positively charged slides (Thermo Scientific®,
J1800AMNZ). Slides were placed in a hot oven at 37°C for 2 days to allow for
evaporation of excess water. Before staining, slides were placed on a hot plate at
60°C for 30 minutes to remove air bubbles trapped under the wax.

2.3.4 Haematoxylin and eosin stain

For morphological characterisation of the human spine and to assess whether the
developing human NP contained notochordal cells, foetal spine sections were
stained with haematoxylin and eosin (H&E).

Five micrometre tissue sections were prepared as described in 2.3.3. First,
sections were deparaffinised in xylene for 3x 5 minutes, re-hydrated in IMS for 4x
2 minutes and washed in running tap water for 5 minutes. Re-hydrated sections were stained in freshly filtered Mayer’s haematoxylin (Solmedia Laboratory Supplies®, HST011) for 1 minute, rinsed in running tap water for 5 minutes and counterstained in freshly filtered eosin (stock solution: 100mL Eosin A (Merck®, 341972Q), 10mL Eosin Phloxin B (Merck® 341972Q), 4mL glacial acetic acid (VWR International®, 10001CU), 780mL IMS) for 10 seconds. Stained sections were dehydrated in IMS for 4x 2 minutes, cleared in xylene for 3x 5 minutes and mounted with a coverslip in mounting medium (Shand Consul-Mount, Thermo Scientific®, 9990440).

2.3.5 Masson’s trichrome staining

Masson’s trichrome stain was used to analyse collagen content in the developing spine. Due to the limited size and number of the younger specimens (3.5 and 5.5 WPC), this stain was only performed in specimens between 6 - 18 WPC.

Five micrometre tissue sections were prepared as described in 2.3.3. First, slides were deparaffinised in xylene for 3x 5 minutes, re-hydrated in IMS for 4x 1 minute and washed in running tap water for 5 minutes. Re-hydrated slides were stained in Weigert's haematoxylin (equal volumes of Weigert's iron haematoxylin A (TCS Biosciences®, HS375) and B (TCS Biosciences®, HS380)) for 10 minutes and washed in running tap water for 5 minutes to remove excess stain. Then, slides were counter-stained with Ponceau acid fuchsin solution (Biostain®, RRSP131-c) for 10 minutes, after which excess stain was removed by washing in
running water for 5 minutes, followed by a 10 minutes incubation in phosphomolybdic acid solution (Biostain®, RRSP172-c) and a 5 minute wash in distilled water. Next, slides were stained with Masson aniline blue (GCC diagnostics®), after which acetic acid (GCC Diagnostics®, UN: 3265) was added for 5 minutes. Finally, slides were washed in deionised water to remove excess stain, dehydrated in IMS for 4x 2 minutes, cleared in xylene for 3x 5 minutes and mounted with a coverslip in mounting medium.

### 2.3.6 Safranin-O staining

Safranin-O/ fast green (SAF-O) stain was used to analyse the proteoglycan content of the developing spine. Due to the limited size and number of the younger specimens (3.5 and 5.5 WPC), this stain was only performed in specimens between 6 - 18 WPC.

Five micrometre tissue sections were prepared as described in 2.3.3. First, slides were deparaffinised in xylene for 3x 5 minutes, re-hydrated in IMS for 4x 2 minutes and washed for 10 seconds in tap water. Re-hydrated tissue sections were then stained in Weigerts haematoxylin (prepared as described in 2.3.5) for 3 minutes and then washed for 10 minutes in running tap water. Sections were then stained with 0.1% (w/v) aqueous fast green (Difco®, 42053) for 5 minutes, followed by a 10 seconds wash with 1% (v/v) aqueous acetic acid. Next, slides were rinsed in tap water for 10 seconds and then stained in filtered 0.1% (v/v) aqueous safranin-O (Sigma®, S2255) for 4 minutes. Finally, sections were
dehydrated in IMS for 5x 2 minutes, cleared in xylene for 3x 5 minutes and mounted with coverslips in mounting medium.

2.3.7 Immunochemistry

Protein expression of KRT8, KRT18, KRT19, GAL3, CD24, CD55, BASP1, T, CTGF, CD90, Tie2 and E-Cad was assessed using immunohistochemistry and disclosed with the Avidin-Biotin Complex (ABC) method.

Five micrometre tissue sections were prepared as described in 2.3.3. First, slides were deparaffinised in xylene for 3x 5 minutes, re-hydrated in IMS for 4x 2 minutes and washed in tap water for 10 seconds to remove excess IMS. For the antibodies where enzyme-only antigen retrieval methods were used, endogenous peroxidase blockade was performed prior to the antigen retrieval and with a 2x 5 minute Tris Buffered Saline (TBS) (0.5M Tris Base (Fisher Bioreagent®, BP152-1), 9% w/v NaCl (Fisher Chemical®, S/3160/65), pH 7.6) wash between both steps.

Where heat antigen-retrieval methods were used, antigen retrieval was performed prior to blocking endogenous peroxidases (2x 5 minute TBS washes between the two steps). Endogenous peroxidase blockade was performed by immersing slides in 100% IMS containing 0.3% (v/v) hydrogen peroxide and 25mM HCl for 30 minutes at room temperature. Antigen retrieval methods were determined empirically as described in 2.3.8.
After these steps, non-specific binding sites were blocked by incubating slides in 25% (v/v) goat serum (Sigma-Aldrich®, G9023) in 1% w/v bovine serum albumin (BSA, Sigma-Aldrich®, A9647) in TBS for 30 minutes at room temperature (2% and 5% BSA were respectively used for GAL3 and KRT18); BSA concentrations were determined empirically with higher concentrations being used to eliminate background staining.

Following blocking, slides were stained overnight at 4ºC (for KRT8, KRT18 and KRT19, GAL3, CD55 and BASP1) or for 2 hours at room temperature (for CD24, CTGF, Tie2, CD90, E-Cad and T) with primary antibodies diluted in 1% w/v BSA in TBS (2% and 5% BSA for GAL3 and KRT8, respectively).

Following primary antibody incubation, slides were washed in TBS for 3x 5 minutes. Primary antibodies were localised by incubating with biotinylated secondary antibodies diluted in 1% BSA in TBS (2% and 5% BSA for GAL3 and KRT8, respectively) for 30 minutes at room temperature. A goat anti-mouse secondary antibody (Santa Cruz Biotechnology®, sc-3795) was used to localise primary antibodies raised in mice and a goat anti-rabbit secondary antibody (Santa Cruz Biotechnology®, sc-3840) was used for primary antibodies raised in rabbit.

Slides were then washed in TBS for 3x 5 minutes. Amplification was performed by incubating with 4 drops of ABC-Amplification reagent (Vectastain®) for 30 minutes at room temperature. Following amplification, slides were washed in TBS for 3x 5 minutes and detection of the Avidin-Biotin complex was performed by incubating for 18 minutes with 3, 3’-diaminobenzidine (DAB) (Sigma-Aldrich®, D5905-50TAB). Sections were then washed in deionised water for 4 minutes and nuclei were counterstained by incubating for 90 seconds in freshly filtered Mayer’s Haematoxylin. Excess staining was removed by rinsing slides in...
running tap water for 5 minutes. Stained sections were finally dehydrated in IMS for 4x 5 minutes, cleared in xylene for 3x 5 minutes and mounted with a coverslip in mounting medium.

Table 2.2 details primary antibodies and antigen retrieval methods used. Unless where specified, all procedures were performed at room temperature and incubations were performed in a wet box to prevent the slides from drying out.
Table 2.2 Details of the antibodies and antigen retrieval methods used.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Antibodies</th>
<th>Antigen retrieval method</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRT8</td>
<td>0.5µg/mL mouse monoclonal anti-KRT8 IgG1 (Zytromed Systems®, 603-2156)</td>
<td>Pepsin + Pronase</td>
</tr>
<tr>
<td>KRT18</td>
<td>0.338µg/mL mouse monoclonal anti-KRT18 IgG1 (DakoCytomation®, M7010)</td>
<td>Heat TrisEDTA</td>
</tr>
<tr>
<td>KRT19</td>
<td>0.016µg/mL mouse monoclonal anti-KRT19 IgG1 (DakoCytomation®, M0888)</td>
<td>Heat TrisEDTA</td>
</tr>
<tr>
<td>CD24</td>
<td>0.20mg/mL mouse monoclonal anti-CD24 IgG1 (Abcam®, ab31622)</td>
<td>Heat Citrate</td>
</tr>
<tr>
<td>GAL3</td>
<td>4µg/mL rabbit polyclonal anti-Galectin-3 IgG (Santa Cruz Biotechnology®, sc-20157)</td>
<td>Pepsin + Pronase</td>
</tr>
<tr>
<td>CD55</td>
<td>25µg/mL mouse monoclonal anti-CD55 IgM (Sigma-Aldrich®, SAB4700249)</td>
<td>Heat TrisEDTA</td>
</tr>
<tr>
<td>CTGF</td>
<td>1µg/mL mouse monoclonal anti-CTGF IgG1 (R&amp;D Systems®, MAB660)</td>
<td>Heat Citrate</td>
</tr>
<tr>
<td>BASP1</td>
<td>0.67µg/mL rabbit polyclonal anti-BASP1 IgG (Santa Cruz Biotechnology®, sc-66994)</td>
<td>Heat Citrate</td>
</tr>
<tr>
<td>Tie2</td>
<td>4µg/mL mouse monoclonal anti-Tie2 IgG1 (Novus Biologicals®, NB110-60986)</td>
<td>Proteinase K</td>
</tr>
<tr>
<td>CD90</td>
<td>0.184µg/mL rabbit monoclonal anti-CD90 IgG (Abcam®, ab133350)</td>
<td>Heat Citrate</td>
</tr>
<tr>
<td>E-Cad</td>
<td>1.3µg/mL mouse monoclonal anti-ECad IgG1 (Abcam®, ab1416)</td>
<td>Heat Citrate</td>
</tr>
<tr>
<td>T</td>
<td>2µg/mL rabbit polyclonal anti-Brachyury IgG (Abcam®, ab20680)</td>
<td>Heat Citrate</td>
</tr>
</tbody>
</table>
2.3.8 Optimisation of immunohistochemical staining

Prior to use in this study, immunohistochemistry staining methods (including antigen retrieval and antibody concentrations) were optimised using foetal spine sections. To exclude non-specific binding additional foetal samples were also stained with isotype immunoglobulin controls (instead of primary antibodies) at the same concentration as the primary antibodies and using the same antigen retrieval methods. In cases where background staining was detected slides were stained without primary antibody (using BSA in TBS). For antibodies where no staining was detected, antigen retrieval method and antibody concentration were optimised on positive control tissue sections. To identify ideal positive control tissues the Human protein atlas project (www.proteinatlas.org) was consulted.

To identify the optimal antigen retrieval method for each antibody, enzymatic (pepsin, pronase, pepsin + pronase and proteinase K) and heat (citrate and Tris-EDTA) antigen retrieval methods were empirically tested. The antigen retrieval protocols were:

a. Pepsin antigen retrieval:
   Slides were incubated in 0.25% pepsin (Sigma Aldrich®, P7012) (97066U/g) (w/v) in 10mM ice-cold HCl for 10 minutes at room temperature.

b. Pronase antigen retrieval:
   Slides were incubated in 0.1% pronase (Calbiochem®, 53702) (3200U/mg) (w/v) in TBS, pH 7.4 (50mM Tris-EDTA, 150mM NaCl) for 10 minutes at 37°C.
c. Pepsin + Pronase antigen retrieval:

Slides were incubated in 0.25% pepsin (97066U/g) (w/v) in 10mM ice-cold HCl for 10 minutes at room temperature. After pepsin incubation, slides were washed for 2x 5min in TBS and then incubated in 0.1% pronase (3200U/mg) (w/v) in TBS, pH 7.4 for 10 minutes at 37°C.

d. Proteinase K antigen retrieval:

Slides were incubated in 20µg/mL proteinase K (Promega®, V3021) in TBS for 15 minutes at room temperature.

e. Heat citrate antigen retrieval:

Heat induced antigen retrieval was performed using a steamer (filled with 1L of deionised water and pre-heated for 10 minutes). A box containing slides immersed in 800mL 10mM citrate buffer (pH 6.0 (10mM citric acid)) previously heated to 95°C was incubated in the steamer for 10 minutes and then transferred to the bench (at room temperature) for 10 minutes.

f. Heat TrisEDTA antigen retrieval:

Heat TrisEDTA antigen retrieval was performed as described for heat citrate but using TrisEDTA (10mM Tris, 1mM EDTA, pH 8.0) instead of citrate as buffer.

To determine the optimal antibody concentration each antibody was tested at a range of dilutions, starting at 1/10, 1/25, 1/50 and 1/100. If intense staining was still seen with 1/100, further dilutions were performed (1/250; 1/500; 1/750; 1/1000; 1/2000, etc.). Equivalent concentrations of the respective isotype control IgG were used as negative controls.
Secondary antibodies (goat anti-mouse and goat anti-rabbit) were used at 1.33 µg/mL, as this concentration has been thoroughly tested and used in the host laboratory.

### 2.3.9 Image analysis

All preliminary histological analysis was undertaken using a Leica® Leitz DMRB microscope. The final cohort of foetal spine sections was scanned using the Pannoramic 250 Flash II digital slide scanner (3DHistech®) and visualised using the Pannoramic Viewer software (3DHistech®) (Krenacs et al., 2010). For each stain/antibody sufficient images were chosen to depict the labelling along the developmental stages analysed.
2.4 RESULTS

2.4.1 Morphology

Notochord cells were present in all embryonic and foetal spine sections analysed and were initially along the embryo axis to later become restricted to the developing NP; notochordal cells were typically round and vacuolated (Figure 2.2 A and B-M panels 2).

In the 3.5WPC specimen the notochord was localised anteriorly to the neural tube and with a row of somites on each side; morphologically, notochordal cells were vacuolated and organized side by side forming a cylindrical midline epithelioid-like structure (Figure 2.2 A). With embryonic development, the notochord and somites extended along the embryo axis (Figure 2.2 B-D, panels 1); while notochordal cells maintained their central localisation, lateral somite cells (dermatomyotomal cells, dermis and muscle precursors) migrated laterally whereas those located more centrally (sclerotomal cells) migrated towards the midline (Figure 2.2 B2). Furthermore, sclerotomal cells adopted a segmented morphology pattern: segments with higher cell density (precursors of the AF in the IVD region) intercalated with less densely organised segments (precursors of the VB) (Figure 2.2 C-D panels 1).

Between the 8th and 10th WPC, notochordal cells progressively disappeared from the less densely organised sclerotomal segments (VB anlagens) to occupy a wider midline region within the adjacent denser segments (IVD anlagens) (Figure 2.2 E,
F and G); this was accompanied by a progressive enlargement of the vacuole size (Figure 2.2 G2). After the 10th WPC the VB segments were practically devoid of notochordal cells, with only remnants of the notochordal sheath being present there; at that and in all stages thereafter notochordal cells were restricted to the IVD segments (Figure 2.2 G-M panels 2). Sclerotomal cells within the IVD anlagen segments were organised in lamellae around the central notochordal region; the lamellae were tighter in the more peripheral IVD regions (outer AF anlagen) and looser in the more central regions (inner AF anlagen) (Figure 2.2. G-M panels 3). Sclerotomal cells in the adjacent VB anlagen segments were round and more loosely organised. With advancing stages, the VB anlagen cells became hypertrophic, in a centripetal manner (Figure 2.2 G1 and H-M panels 4).

A more pronounced demarcation between the foetal structures was noted in the older specimens analysed, with a clear demarcation between the NP, AF and VB precursor cells (Figure 2.2 L-M panels 1).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.2 H&E staining of a cohort of developing spines. The notochord, develops as a rod-like centrally located structure formed by large and vacuolated notochordal cells (arrows) and surrounded by somites (S) which will later become sclerotomal cells. During the analysed stages, the notochord involutes to become localised to the central IVD region (NP anlagen) and sclerotomal cells adopt a segmented pattern; densely organised sclerotomal regions will form the developing AF (formed by the outer AF (oAF) and the inner AF (iAF)). Sclerotomal cells in the adjacent regions have a round morphology and will later form the VB. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M396 (6 WPC); D: M741 (7 WPC); E: M736 (8 WPC); F: M397 (9 WPC); G: M685 (10 WPC); H: M425 (11 WPC); I: M742 (12 WPC); J: M776 (13 WPC); K: M739 (14 WPC); L: M777 (17 WPC); M: M784 (18 WPC). For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one, two or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP, panels 3 are centred to the developing sclerotomal AF and panels 4 highlight the sclerotomal VB.
2.4.2 Extracellular matrix

SAF-O staining pattern was consistent through all the development stages analysed. Sclerotomal cells in the developing VB stained strongly positive, demonstrating the high proteoglycan content of these cells – this staining became more intense with advancing stages (Figure 2.3 A-I panels 1). Notochordal cells were weakly positive for this stain in all stages analysed; the notochordal sheath, lining the notochordal cells and separating them from the adjacent sclerotomal cells, however, was strongly positive (Figure 2.3 A-I panels 2). SAF-O did not stain outer AF cells in all stages analysed and inner AF cells were only weakly positive (Figure 2.3 A-C panels 2 and D-I panels 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.3 SAF-O staining of a cohort of developing spines. Sclerotomal cells of the developing VB showed high proteoglycan content; notochordal cells (arrows) and inner AF cells were only weakly positive and outer AF cells were strongly positive. A: M741 (7 WPC); B: M779 (7.5 WPC); C: M636 (9.5 WPC); D: M685 (10 WPC); E: M404 (11.5 WPC); F: M776 (13 WPC); G: M739 (14 WPC); H: M777 (17 WPC); I: M784 (18 WPC).

For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF.
Similar to SAF-O, Mason’s trichrome staining was also relatively consistent throughout the developmental stages analysed, with only slight variations in staining intensity with different stages. Intense blue staining was seen in the AF in all stages analysed; that staining was more prominent in the outer AF than in the inner AF and became accentuated with advancing stage (Figure 2.4 A-C panels 2 and D-I panels 3). Reflecting the less collagenous nature of notochordal cells, these cells stained in light red (Figure 2.4 A-I panels 2). VB chondrocytes stained blue in the earlier developmental stages, but became red in the later stages (Figure 2.4 A-I panels 1).
Masson’s trichrome

6-7 WPC

7-8 WPC

9-10 WPC

10-11 WPC
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Figure 2.4. Masson’s trichrome staining of a cohort of developing spines. Intense collagenous (blue) staining is seen in the developing AF, particularly in the outer AF; notochordal cells (arrows) stain red showing the less fibrous nature of this tissue A: M741 (7 WPC); B: M779 (7.5 WPC); C: M636 (9.5 WPC); D: M685 (10 WPC); E: M404 (11.5 WPC); F: M766 (13 WPC); G: M739 (14 WPC); H: M777 (17 WPC); I: M784 (18 WPC). For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF.
2.4.3 Identification of developmental markers

No differences in staining intensity were noted for any of the antibodies analysed in the different spinal levels (cervical, thoracic and lumbar) at each specific developmental stage.

2.4.3.1 Notochord-specific markers

KRT8 was specifically expressed by all notochordal cell in all developmental stages analysed (3.5-18 WPC); staining was localised to the cytoplasm and around, but not inside, the vacuoles (Figure 2.5 A and B-K panels 2). Somite (3.5 WPC) and sclerotomal (5.5-18 WPC) cells in the developing AF and VB did not express this protein in any of the analysed stages (Figure 2.5 A, B- E panels 2, and F-K panels 2 and 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.5 KRT8 immunostaining of a cohort of developing spines showing notochord-specific expression (arrows) of this marker. No KRT8 expression was seen in the surrounding sclerotomal AF and VB cells. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M414 (8.5 WPC); E: M397 (9 WPC); F: M425 (11 WPC); G: M439 (12 WPC); H: M776 (13 WPC); I: M739 (14 WPC); J: M777 (17 WPC); K: M784 (18 WPC); L: M415 (10.5 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. No staining was seen when the antibody was substituted for an equal concentration of its isotype control (-ve control).
KRT18 staining pattern was identical to that of KRT8, with cytoplasmic notochordal cell expression in all stages analysed (Figure 2.6 A and B-K panels 2) and no expression in any somite/sclerotomal cells (Figure 2.6 A, B-E panels 2, and F-K panels 2 and 3).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto

KRT18

3.5 WPC

5.5 WPC

6-7 WPC

7-8 WPC
Isolation and phenotypic characterisation of human notochordal cells
Ricardo Rodrigues Pinto
Figure 2.6 KRT18 immunostaining of a cohort of developing spines showing notochord-specific expression (arrows) of this marker. No KRT-18 expression was seen in the surrounding sclerotomal AF and VB cells. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M414 (8.5 WPC); E: M397 (9 WPC); F: M483 (10 WPC); G: M425 (11 WPC); H: M776 (13 WPC); I: M739 (14 WPC); J: M777 (17 WPC); K: M784 (18 WPC); L: M483 (10 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
Similar to KRT8 and KRT18, KRT19 was also notochord-specific in all developmental stages analysed (Figure 2.7 A and B-J panels 2). The staining was localised to the cytoplasm of notochordal cells (Figure 2.7 A, B-E panels 2, and F-J panels 2 and 3).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.7 KRT19 immunostaining of a cohort of developing spines showing notochord-specific expression (arrows) of this marker. No KRT-19 expression was seen in the surrounding sclerotomal cells in the developing AF and VB. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M414 (8.5 WPC); E: M397 (9 WPC); F: M715 (10.5); G: M439 (12 WPC); H: M739 (14 WPC); I: M777 (17 WPC); J: M784 (18 WPC); K: M483 (10 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
CD24 was not expressed by notochordal or somite cells in the 3.5 WPC sample (Figure 2.8 A). From 5.5 WPC to 18 WPC, however, CD24 was specifically expressed in the extracellular membrane of all notochordal cells (Figure 2.8 B-H panels 2). No sclerotomal staining was seen in any of the developmental stages analysed (Figure 2.8 A, B-D panels 2 and E-H panels 2 and 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells
Ricardo Rodrigues Pinto
Figure 2.8 CD24 immunostaining of a cohort of foetal spines showing notochord-specific expression (arrows) of this marker between 5.5-18 WPC. No expression is seen in somite or notochordal cells at 3.5 WPC. No CD24 expression was seen in the surrounding sclerotomal cells in the developing AF and VB in any analysed stage. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M744 (8 WPC); D: M636 (9.5 WPC); E: M685 (10 WPC); F: M404 (11.5 WPC); G: M739 (14 WPC); H: M784 (18 WPC); I: M636 (9.5 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
2.4.3.5 Proteins with variable expression with developmental stages

GAL3 was not expressed by notochordal or somite cells in the 3.5 and 5.5 WPC samples (Figure 2.9 A and B2). Between 7-18 WPC, however, GAL3 was expressed in the cytoplasm of all notochordal cells (Figure 2.9 C-J panels 2). The expression was notochord-specific between 7-9 WPC but was co-expressed in the cytoplasm of sclerotomal VB cells between 10-18 WPC (Figure 2.9 E-J panels 4); no expression was noted in the sclerotomal AF cells at any stages analysed (Figure 2.9 C-D panels 2 and A-J panels 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.9: GAL3 immunostaining of a cohort of foetal spines. GAL-3 is notochord-specific (arrows) between 7-9 WCP after which is becomes co-expressed by sclerotomal VB cells. 

A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M680 (7.5 WPC); D: M414 (11 WPC); E: M415 (10.5 WPC); F: M425 (11 WPC); G: M776 (13 WPC); H: M739 (14 WPC); I: M777 (17 WPC); J: M784 (18 WPC); K: M404 (11 WPC) rabbit isotype IgG control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one, two or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP, panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
CD55 was not expressed by notochordal or somite cells in the 3.5 and 5.5 specimens (Figure 2.10 A and B). Between 7–9 WPC CD55 was specifically expressed on the extracellular membrane of all notochordal cells (Figure 2.10 B-J panels 2); however, in these stages sclerotomal cells did not express CD55 (Figure 2.10 C2 and D3). After 10 WPC, CD55 became co-expressed by the sclerotomal cells in the developing AF (Figure 2.10 E-J panels 2); sclerotomal cells in the VB anlagen never expressed CD55 (Figure 2.10 E-J panels 3).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells
Ricardo Rodrigues Pinto
Figure 2.10 CD55 immunostaining of a cohort of developing spines. CD55 was notochord (arrows) specific between 5.5-10 WPC, after which it became co-expressed by sclerotomal cells in the developing AF. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M779 (7.5 WPC); D: M636 (9.5 WPC); E: M685 (10 WPC); F: M425 (11 WPC); G: M776 (13 WPC); H: M739 (14 WPC); I: M777 (17 WPC); J: M784 (18 WPC); K: M425 (11 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
In the 3.5 and 5.5 WPC specimens, CTGF was not expressed by notochord or somite/sclerotomal cells (Figure 2.11 A and B2). Between 6-18 WPC CTGF expression was found in all notochordal (Figure 2.11 C-I panels 2) and sclerotomal cells in the developing VB (Figure 2.11 C-D panels 2 and E-I panels 4). The intensity of VB staining was weak in all stages (Figure 2.11 C-D panels 2 and E-I panels 4) and that of notochordal cells was weak between 11-17 WPC (Figure 2.11 F-H panels 2). Sclerotomal AF cells did not express CTGF at any stage analysed (Figure 2.11 C-D panels 2 and E-I panels 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.11 CTGF immunostaining of a cohort of foetal spines showing notochord (arrows) and VB co-expression between 6-18 WPC. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M779 (7.5 WPC); E: M685 (10 WPC); F: M404 (11 WPC); G: M776 (13 WPC); H: M777 (17 WPC); I: M784 (18 WPC); J: M779 (7.5 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one, two or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP, panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
BASP1 was expressed by all notochordal and somite cells in the 3.5 WPC specimen (Figure 2.12 A). In the 5.5 WPC specimen no BASP1 staining was seen in notochordal or sclerotomal cells (Figure 2.12 B2). Between 6-18 WPC BASP1 was co-expressed by notochordal and sclerotomal cells (AF and VB anlagen); notochordal and sclerotomal staining, however, was weak between 6-8 WPC (Figure 2.12 C-D panels 2) and became more intense between 9-18 WPC (Figure 2.12 E-I panels 2 for notochordal cells, E-I panels 3 for AF anlagen cells and E-I panels 4 for VB anlagen cells).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.12 BASP1 immunostaining of a cohort of developing spines. BASP1 was localised to all notochordal (arrows) and somite/sclerotomal cells at all stages analysed, except in the 5.5 WPC specimen, where no developing spine anlagen staining was found. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M680 (7.5 WPC); E: M685 (10 WPC); F: M404 (11 WPC); G: M739 (14 WPC); H: M777 (17 WPC); I: M784 (18 WPC); J: M404 (11 WPC) rabbit isotype IgG control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one, two or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP, panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
T was notochord-specific in the 3.5 WPC specimen (Figure 2.13 A). In the 5.5 WPC specimen T was expressed by notochordal and sclerotomal cells (Figure 2.13 B2). Between 6-18 WPC T was expressed by all notochordal (Figure 2.13 C-K panels 2) and sclerotomal cells in the VB anlagens (Figure 2.13 C-E panels 2 and F-K panels 4); no expression was seen in the AF anlagens at any stage analysed (Figure 2.13 C-E panels 2 and F-K panels 3).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells
Ricardo Rodrigues Pinto
Figure 2.13 T immunostaining of a cohort of developing spines showing notochord (arrows) specificity in the 3.5 WPC sample and notochordal and sclerotomal (VB anlagen) co-expression between 5.5-18 WPC. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M779 (7.5 WPC); E: M636 (9.5 WPC); F: M685 (10 WPC); G: M742 (12 WPC); H: M776 (13 WPC); I: M739 (14 WPC); J: M777 (17 WPC); K: M784 (18 WPC); L: M742 (12 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one, two or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP; panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
2.4.3.11 Proteins without expression in the developing spine

Tie2 was not expressed by any developing spine cell anlagen in any of the stages analysed: notochordal cells (Figure 2.14 A and B-J panels 2), sclerotomal AF anlagen cells (Figure 2.14 A, B-D panels 2 and E-J panels 3) and sclerotomal VB anlagen cells (Figure 2.14 A, B-D panels 2 and E-J panels 3) were always negative for this protein.
RicaRdo RodRigues Pinto
Isolation and phenotypic characterisation of human notochordal cells
10-11 WPC

11-12 WPC

13 WPC

Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 2.14 Tie2 immunostaining in a cohort of developing spines showing absence of notochord or somite/sclerotomal expression in all stages analysed. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M680 (7.5 WPC); D: M636 (9.5 WPC); E: M404 (11 WPC); F: M742 (12 WPC); G: M776 (13 WPC); H: M739 (14 WPC); I: M777 (17 WPC); J: M784 (18 WPC); K: placenta positive control, L: placenta negative control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. Staining was optimised on placental tissue: positive control (+ve control) was stained with Tie2 antibody and negative control (-ve control) was stained with an equivalent concentration of isotype IgG.
CD90 was not expressed by any developing spine cell anlagen in any of the stages analysed: notochordal cells (Figure 2.15 A and B-H panels 2), sclerotomal AF anlagen cells (Figure 2.15 A, B-C panels 2 and C-H panels 3) and sclerotomal VB anlagen cells (Figure 2.15 A, B-C panels 2 and C-H panels 3) were always negative for this protein.
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto

176
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
**Figure 2.15** CD90 immunostaining in a cohort of developing spines showing lack of notochord or sclerotomal expression in all stages analysed (in figure E2 antibody trapping, rather than true staining, was found around notochordal cells). A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M680 (7.5 WPC); D: M685 (10 WPC); E: M404 (11 WPC); F: M739 (14 WPC); G: M777 (17 WPC); H: M784 (18 WPC); I: kidney positive control; J: kidney negative control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. CD90 staining was optimised in kidney sections stained with CD90 antibody (**+ve control**) and in kidney sections stained in an equal concentration of its isotype control (**–ve control**).
E-Cad was not expressed by any developing spine cell anlagen in any of the stages analysed: notochordal cells (Figure 2.16 A and B-K panels 2), sclerotomal AF anlagen cells (Figure 2.16 A, B-E panels 2 and F-K panels 3) and sclerotomal VB anlagen cells (Figure 2.16 A, B-E panels 2 and F-K panels 3) were always negative for this protein.
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO

14 WPC

17 WPC

18 WPC
Figure 2.16  E-Cad immunostaining of a cohort of developing spines showing absence of notochordal or somite/sclerotomal expression in all stages analysed. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M680 (7.5 WPC); E: M636 (9.5 WPC); F: M685 (10 WPC); G: M439 (12 WPC); H: M776 (13 WPC); I: M739 (14 WPC); J: M777 (17 WPC); K: M784 (18 WPC); L: M439 (12 WPC) isotype control; M: kidney positive control; N: kidney negative control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. E-Cad staining was optimised in kidney sections stained with E-Cad antibody (+ve control) and in foetal and kidney sections stained in an equal concentration of its isotype control (–ve control).
2.4.3.12 Background staining

Whenever heat citrate or pepsin + pronase antigen retrievals were used as antigen retrieval methods, specimens M776 (13 WPC), M739 (14 WPC), M734 (14 WPC) M777 (17 WPC) and M784 (18 WPC) showed a patchy, non-cellular staining. To investigate the meaning of this staining, the same samples were incubated using the same experimental conditions (heat citrate or pepsin + pronase antigen retrievals) but with PBS instead of primary antibody.

The same patchy staining was found to be present in the same regions when using this negative control, suggesting that this was background staining rather than antibody-antigen binding (Figure 2.17).
No primary antibody controls

13 WPC

13 WPC

14 WPC

14 WPC
Figure 2.17 No primary antibody controls. Background staining in specimens 13-18 WPC after heat citrate or pepsin + pronase antigen retrieval and incubation without primary antibody (PBS). A: M776 (13 WPC) heat citrate antigen retrieval; B: M776 (13 WPC) pepsin + pronase antigen retrieval; C: M739 (14 WPC) heat citrate antigen retrieval; D: M739 (14 WPC) pepsin + pronase antigen retrieval; E: M777 (17 WPC) heat citrate antigen retrieval; F: M777 (17 WPC) pepsin + pronase antigen retrieval; G: M784 (18 WPC) heat citrate antigen retrieval; H: M784 (18 WPC) pepsin + pronase antigen retrieval. For each panel 1, one higher magnification (panels 2) highlights an area of background staining.
2.4.5 Summary of results

The immunohistochemical staining allowed an understanding of the different pattern of staining along the different developmental stages analysed. While some markers were notochord-specific during all stages, others had expression that varied with the stage analysed and others showed no spinal anlagen expression. Figure 2.18 is a schematic illustration of the pattern of intensity for each of the antibodies analysed.
Figure 2.18 Schematic representation of the protein staining in the developing human spine. KRT8, KRT18 and KRT19 were notochord-specific markers at all stages analysed and CD24 was notochord-specific between 5.5-18 WPC. The expression of GAL3, CD55, CTGF, BASP1 and T varied with development stage. CD90, Tie2 and E-Cad were not expressed by any developing spine cell anlagen.
2.5 DISCUSSION

The search for novel therapies for IVD degeneration has motivated an increased interest in the understanding of the native nucleus pulposus phenotype (and hence ontogeny its cells) to guarantee that implanted cells have the correct phenotype to ensure adequate function. While the human developing NP is composed of large vacuolated notochordal cells, in the adult the NP is composed of small non-vacuolated cells whose ontogeny, despite lineage tracing studies in mice, is still a matter of debate. It is unclear whether the original population of notochordal cells dies to be replaced by cells migrating from adjacent tissues, differentiates into cells with a different morphology or both. To clarify this controversy and, since cell size and morphology differences are not uncommon in cells with common ancestry (Risbud et al., 2010), specific molecular markers for notochordal cells are needed.

Several studies have investigated the NP cell phenotype in rats (Fujita et al., 2005, Lee et al., 2007, Tang et al., 2012), dogs (Sakai et al., 2009, Smolders et al., 2013), cows (Gilson et al., 2010, Minogue et al., 2010b) and – since the NP phenotype differs between species (Rodrigues-Pinto et al., 2013) – in humans (Minogue et al., 2010a, Power et al., 2011, Weiler et al., 2010). Interestingly, some of the genes identified in the human adult NP had previously been identified within larger, notochordal cells of bovine IVD (Minogue et al., 2010b). These studies, however, could not clarify how specific to notochordal cells those genes were and, therefore, how indicative of notochordal ontogeny they could be.
To adequately clarify the ontogeny of the cells populating the adult NP it is fundamental to understand IVD development and to identify unique notochordal cell markers that may allow the identification of notochord-derived cells in humans, even after a morphological change. Studies have investigated the role of notochordal cells in IVD development and degeneration in rats (Bedore et al., 2013, Choi et al., 2008, Dahia et al., 2012, Dahia et al., 2009, DiPaola et al., 2005, Hayes et al., 2011, Maier et al., 2013, Tang et al., 2012). Only a few studies, however, have investigated the notochordal cell phenotype in humans (Rutges et al., 2010b, Shu et al., 2013, Smith et al., 2009b, Hayes et al., 2013). Furthermore, these studies have either had access to very limited number of samples and/or have focused on the investigation of the expression of ECM proteins. This, although informative regarding the microenvironment and the physicochemical characteristics of the cells populating the IVD at these developmental stages does not elucidate the phenotype of the developing notochordal cells, does not provide unique notochordal markers and, hence, does not clarify the ontogeny of the adult NP cells.

In this study, 31 human samples between the embryonic stages of 3.5-8 WPC (CS 10 to 23) and foetal stages of 8-18 WPC were used. Equating these to mouse \textit{in utero} development, which lasts for approximately 18-20 days, the equivalent to mouse embryonic days 9.5 (human CS 10, 3.5 WPC) to 18-20 (human CS 23, 8 WPC) was analysed, added by the human foetal stages of 8-18 WPC which find no parallel in mouse development (Downs and Davies, 1993).

This is, to the author’s knowledge, the first description of the morphology and phenotypic markers of the developing human spine. The large number of samples provided comprehensive staining across the different developmental stages...
allowing for a detailed description of the developing spine marker staining during the first and part of the second trimester of gestation.

There were no identifiable differences in marker staining between the IVD anlagens located in the cervical, thoracic and lumbar regions possibly reflecting a common microenvironment to the different spinal regions at each developmental stage; it is unclear if in the mature adult spine, in which, at the same time, different spinal segments are subjected to different mobility and load, differences in phenotype occur. In contrast, for most markers, the staining differed with developmental stage. The staining of some of the markers, however, remained constant throughout the developmental stages analysed, with some showing notochord-specificity.

KRT8, KRT18 and KRT19 were found to be human notochord-specific markers. Keratins are intermediate filaments mostly found in epithelial cells (Moll et al., 1982). Due to their unique cytoskeletal role, keratins act to provide structural integrity; they have also been found to regulate Fas-mediated apoptosis and modulate cell size and protein synthesis (Sun et al., 2013). During the developmental stages analysed, the notochord is the main axial element of the embryo and foetus, allowing the embryo to elongate (Stemple, 2005). Later, its cells are subjected to continuous hydrostatic pressure (Adams et al., 1990) exerted by the adjacent sclerotomal cells, leading to notochordal cell disappearance from the centre of the developing VB to become localised in the central IVD. It is possible that, given the cytoskeletal properties of keratins, these proteins act to provide the notochord and its cells the structural role they exert during these developmental stages. Furthermore, their expression by notochordal cells confirms the developing notochord as an epithelial tissue. Relevant to the IVD
field, they have also been found in the NP of immature rats (Lee et al., 2007), chondrodystrophic dogs (Sakai et al., 2009), adult cows (Minogue et al., 2010b) and adult humans, with and without histological signs of degeneration (Minogue et al., 2010a, Rutges et al., 2010a, Stosiek et al., 1988). Importantly, the expression of these notochordal markers by cells within the adult human NP indicates that, at least a subpopulation of adult NP cells is of notochordal ontogeny. The expression of these intermediate filaments in the NP of various animal species, at different stages of development, maturation and disease and in cells with or without vacuoles and also, as shown here, in the human developing notochord, indicates that these proteins may be involved in a role that is intrinsic to all notochord-derived NP cells. It is possible that keratins act to provide structural integrity and resistance to pressure to notochordal and NP cells.

This study also identified CD24 as a notochord-specific marker between 5.5-18 WPC (only notochordal cells of the 3.5 WPC specimen did not express this cell surface marker). CD24 is a cell surface protein that is expressed by immature cells and is usually absent after they reach maturity. It is expressed during B and T-cell maturation and in differentiating neuroblasts (Bruce et al., 1981). While some studies have linked it to cancer cell growth, proliferation and survival (Smith et al., 2006), others have found it to be a marker of regenerative muscle cells (Figarella-Branger et al., 1993) and of hepatocyte (Qiu et al., 2011) and renal (Ivanova et al., 2010) progenitor cells. During human development, CD24 is expressed by the intestinal mucosal, nasal, salivary gland, bronchial and renal tubular epithelia and by hair follicles (Rougon et al., 1991). Relevant to the IVD field, CD24 has previously been identified in rat notochordal NP cells (Fujita et al., 2005, Tang et al., 2012) and in human chordomas (Fujita et al., 2005), which supports the
findings in this study confirming this protein as a notochordal marker. Recently, it has also been identified in NP cells from children and adolescents undergoing surgery for scoliosis (Tang et al., 2013), again supporting a notochordal ontogeny for the NP cells in these patients; it is not known, however, if NP cells retain a notochordal phenotype in scoliotic children.

GAL3 and CD55 displayed an identical expression pattern: they were not expressed by any developing spine cell in the earlier stages, were notochord-specific between 7-9 WPC and became co-expressed by sclerotomal cells after 10 WPC. The co-expression of GAL3 by NP and VB anlagen is in disagreement with the findings by Oguz and colleagues which have found this marker to be expressed by NP and AF cells in the new-born rat (Oguz et al., 2007). The co-expression of CD55 by notochordal and sclerotomal AF anlagen cells is also in disagreement with the findings by Leung and colleagues, which had proposed it as a rat notochordal marker (Leung et al., 2010). These findings further stress the important differences in marker expression between species.

BASP1 and CTGF were not notochord-specific at any stage analysed. BASP1, a membrane-bound protein involved in nerve growth and synaptic plasticity and that has been identified in the NP but not in the AF of rat (Tang et al., 2012) and bovine (Minogue et al., 2010b), was ubiquitously expressed in the developing spine in all stages analysed except in the 5.5 WPC specimen. CTGF is a growth factor involved in ECM interactions that is secreted by dog notochordal cells and stimulates NP cells to upregulate proteoglycan production (Erwin, 2008); it is present in the conditioned medium of pig notochordal cells (Gantenbein et al., 2014). In this study, CTGF was co-expressed by notochordal and VB anlagen cells after 6 WPC and until the later stage analysed (18 WPC); this is in
agreement in recent findings by Bedore and colleagues which have identified the expression of this protein in the developing NP and VB (but also the AF) of embryonic mice; the authors of this study further confirm an important role of this protein in regulating IVD development (Bedore et al., 2013).

T is an embryonic transcription factor required for mesoderm formation and differentiation (Vujovic et al., 2006) and notochord-development (Fujita et al., 2015). This protein has often been used as a marker of a notochordal phenotype (Saggese et al., 2014, Liu et al., 2015, Spillekom et al., 2014). In this study T was found to be expressed by all notochordal cells in all stages analysed, which suggests that, also in humans, this transcription factor has a role in notochordal cell development; however, its expression was not specific to notochordal cells, with the developing sclerotomal AF anlagen cells between 5.5-18 WPC also expressing this marker.

CD90, Tie2 and E-Cad expression was not found in any developing spine anlagen cell at all developmental stages analysed. CD90 has previously been proposed as a rat (Tang et al., 2012) and immature human (Tang et al., 2013) AF marker. The findings in this study do not support those observations, as this marker was not expressed by sclerotomal AF anlagen cells at any developmental stage. Sakai and colleagues have identified the presence of Tie2 positive (Tie2⁺ve) cells within NP cells from 8-10 week-old mice and 18-69 year-old humans and have found that those cells, in culture, exhibited multipotency and self-renewal capacity (Sakai et al., 2012); the authors, however, failed to demonstrate the ontogeny of such Tie2⁺ve cells. In this study, notochordal cells did not express Tie2, which suggests that the cells isolated by Sakai and colleagues may be of mesenchymal and not of notochordal origin. Finally, E-Cad, which has previously been identified in
chordomas (Mori et al., 2002, Horiguchi et al., 2004), was not expressed by developing human notochordal cells.
2.6 IMPLICATIONS OF THIS INVESTIGATION

Results presented in this chapter, identify, for the first time, human notochordal cell markers. A developmental pattern of marker protein expression was seen in the different specimens analysed, with marker expression changing with developmental stages. This is possibly due to differences in physicochemical environment acting in the embryo and foetus at each specific developmental stage. While some of these markers were, at least at some stages, notochord-specific, their notochord-specificity was not constant, making them unsuitable as notochord-specific markers.

The expression of some of the markers, however, remained constant through the developmental stages analysed. This suggests that those markers may be more intrinsic to the notochordal cells and their function, rather than the forces and stimuli acting on them. Four human notochord-specific markers (KRT8, KRT18, KRT19 and CD24) were identified. The fact that these proteins, which were previously shown to be present in the human adult NP, were specific to human developing notochordal NP cells suggests that the human adult NP contains notochord-derived cells that have acquired a smaller non-vacuolated morphology. Those markers can potentially be used to identify and isolate notochordal cells in the foetal spine and notochord-derived cells in adult discs.
CHAPTER 3

Isolation and characterisation of human notochordal cells isolated using the specific intracellular marker KRT18
3.1 INTRODUCTION

To develop cell-based therapies for IVD degeneration implanted cells must be differentiated into the correct phenotype and adequately replace or stimulate the native aberrant NP cell population to produce an adequate hydrated ECM. For such, a complete understanding of the target cell phenotype is fundamental.

Microarrays are a powerful technique used to measure the expression of all genes in a given cell population at a specific time point; information from such studies has been used to characterise disease gene profiles, predict disease progression and develop new therapies (Karlsson et al., 2010, Saito et al., 2010, Yang et al., 2010). In the IVD field they have been used to characterise the NP phenotype and to identify novel markers and these have successfully been used to assess stem cell differentiation towards adult NP cells (Fujita et al., 2005, Lee et al., 2007, Minogue et al., 2010b, Minogue et al., 2010a, Power et al., 2011, Sakai et al., 2009, Smolders et al., 2013, Tang et al., 2012, Tang et al., 2013).

Adult NP cells, however, may not have the ideal phenotype to regenerate the degenerate disc as they have been shown to display increased senescence (Le Maitre et al., 2007a), autophagy-mediated cell death (Ye et al., 2011), as well as increased expression of ECM catabolic and degrading enzymes (Le Maitre et al., 2004, Le Maitre et al., 2005a) and decreased expression of ECM components (Pearce et al., 1987, Sive et al., 2002). Conversely and as notochordal cells have anabolic and anti-catabolic properties which suggests that they, or factors they produce, are capable of synthesising a more hydrated ECM that can better perform the disc’s functions, notochordal cells may be a better choice. Thus,
understanding the phenotype and biology of human notochordal cells is fundamental to developing novel biological or cell-based strategies for IVD regeneration.

However, as morphologically distinct vacuolated notochordal cells are only present in the embryonic, foetal and juvenile spine, research on these cells has been impaired by ethical, logistical and technical difficulties and no studies have, to date, characterised the phenotype of human notochordal cells.

In chapter 2 of this thesis, the human developing spine was characterised in terms of morphology and protein marker expression. It was found that the developing NP is formed by large vacuolated notochordal cells, therefore being a suitable source of human notochordal cells. However, due to the small size of the human embryonic and foetal developing spine, and to the proportionally smaller size of the notochordal region, isolation of a pure population of notochordal cells without contaminating them with adjacent sclerotomal cells would not be possible even using microdissection techniques.

FACS uses the principles of light scattering, and excitation and fluorescence emission of fluorochromes attached to specific molecules or expressed by cells, to identify and sort different populations of cells. This technique has previously been used to identify and separate cells based on their specific expression of given proteins and such cells have been used for microarray phenotyping studies (Gallardo and Behra, 2013, Smith et al., 2009a).

During the embryonic and foetal stages analysed in chapter 2 (3.5-18WPC), notochordal cells have a unique expression of KRT8, KRT18 and KRT19. The unique expression of such markers by notochordal cells could potentially be used to separate them from neighbouring sclerotomal cells using FACS for phenotypic
characterisation. KRT18 has previously been identified in the adult human NP (Minogue et al., 2010a), being restricted to sub-populations of NP cells (Weiler et al., 2010), indicating that this embryonic and foetal notochord-specific marker may maintain its notochord-specificity with NP maturation and after notochordal cell morphology change to small NP cells. For this reason, KRT18 was selected to identify and separate foetal notochordal from sclerotomal cells.

Studies utilising FACS to separate cells for microarray studies have most commonly labelled those cells with antibodies directed to cell surface markers. KRT18, however, is an intracellular protein and, therefore, requires the cells to be fixed and permeabilised prior to antibody labelling to allow for antibody penetration into the cell. Extracting RNA from fixed-permeabilised-labelled-sorted cells is challenging due to RNA degradation and the only literature available reporting attempts to do it dates to the 1990’s (Diez et al., 1999, Esser et al., 1995) where using alcohol-based fixation and permeabilisation agents was proposed. In neither of these studies, however, did the obtained RNA have sufficient quality to use for microarrays.

As such, to identify, separate using FACS and characterise with microarrays the gene expression of notochordal cells using their unique expression of KRT18, novel methodologies that would allow fixing, permeabilising, labelling and sorting cells with minimal detrimental impact on its RNA had to be developed.
3.2 HYPOTHESES AND AIMS

The hypotheses for this study were that:

1. KRT18 is a unique notochordal cell marker that can be used to identify and separate notochordal from sclerotomal cells in the developing human spine.
2. RNA with sufficient quality for microarrays can be extracted from notochordal and sclerotomal cells that have previously been fixed, permeabilised, labelled with KRT18 and separated using FACS.
3. Notochordal and sclerotomal differential gene expression will allow an identification of the notochordal cell phenotype and of notochord-specific cell surface markers that can potentially be used to isolate viable notochordal cells.
4. Microarray analysis will allow the identification of the factors regulating notochordal cell development, which may explain the role of these cells in the IVD.

The aims for this study were to:

1. Develop methodologies to allow the isolation of RNA with sufficient quality for microarray studies from cells that have previously been fixed, permeabilised, labelled with an intracellular marker and sorted.
2. Utilise the developed methodology to fix, permeabilise, label with anti-KRT18 antibody, sort and extract good quality RNA from human developing notochordal and sclerotomal cells.
3. Characterise the molecular phenotype of human notochordal cells using microarrays and identify notochord-specific cell surface markers.

4. Analyse microarray data to identify notochordal cell regulators.
3.3 SUMMARY OF THE EXPERIMENTAL DESIGN

To isolate notochordal cells from their adjacent sclerotomal cells, foetal spines were obtained and dissected from their adjacent tissues (ribs, ligaments and spinal cord). Dissected spinal tissues were then digested and dissociated to obtain single notochordal and sclerotomal cells, which were then labelled with anti-KRT18 antibody. FACS was used to identify and separate KRT18 positive (KRT18\textsuperscript{+ve}, notochordal) from KRT18 negative (KRT18\textsuperscript{-ve}, sclerotomal) cells. KRT18\textsuperscript{+ve} and KRT18\textsuperscript{-ve} cells were collected into separate micro-centrifuge tubes (MCTs) and used for RNA extraction. Gene expression profiles of KRT18\textsuperscript{+ve} and KRT18\textsuperscript{-ve} cells were analysed and compared using microarrays (Figure 3.1).

As KRT18 is an intracellular protein, labelling with this marker required cells to be fixed and permeabilised to allow for antibody penetration into the cell.
cytoplasm. As cell fixation and permeabilisation affects RNA quality, different fixation and permeabilisation methods were tested to identify the methodology that yielded the RNA with the highest quality from fixed-permeabilised-labelled-sorted cells.

Due to the difficulties in obtaining sufficient numbers of foetal samples for optimisation procedures, these were conducted using MCF-7 cells (a commercially available cell line known to express KRT18).

Traditionally a protocol to extract RNA from fixed-permeabilised-labelled-sorted cells would comprise the following steps: 1) fixation; 2) permeabilisation; 3) labelling with primary antibody; 4) labelling with secondary antibody; 5) FACS; 6) RNA extraction. Between each step, cells need to be centrifuged to allow for pellet formation, washed and centrifuged again before re-suspending in the subsequent solution (Figure 3.2. A). In preliminary experiments, however, it was found that cell numbers reduced with each step. To minimise cell loss, the following adaptations to the protocol were performed: 1) an agent capable of fixing and permeabilising cells in a single step was used; 2) a directly conjugated antibody was used. It was also found that washing cells after antibody incubation did not affect staining intensity or background staining and, for that reason, this wash was eliminated. Thus, the following protocol was used: 1) one-step fixation and permeabilisation; 2) labelling with directly conjugated antibody; 3) FACS; 4) RNA extraction (Figure 3.2. B). To determine the ideal methodology to be used, RNA was extracted and analysed at the end of step 1 to identify the appropriate fixation/ permeabilisation agent; at the end of step 2 to assess whether KRT18 labelling would affect the RNA; and at the end of step 3 to assess the effect of FACS on RNA quantity and quality.
Figure 3.2 Methodologies for RNA extraction from fixed-permeabilised-labelled-sorted cells. (A) Traditional methodology used to extract RNA from cells labelled with an intracellular marker. This methodology was shortened where possible to reduce the number of steps and washes, as detailed in (B). 1\textsuperscript{st} ab (primary antibody), 2\textsuperscript{nd} ab (secondary antibody).
3.4 MATERIALS AND METHODS

3.4.1 Acquisition of cells

MCF-7 is a human Caucasian breast carcinoma cell line and is the acronym for Michigan Cancer Foundation-7. A frozen vial containing 4x10^6 MCF-7 cells was purchased from the European Collection of Cell Cultures.

T/C28a2 is a juvenile human chondrocyte cell line and was a kind gift from Dr Mary Goldring (Hospital for Special Surgery, Weill Cornell Medical College, New York, USA).

3.4.2 Thawing and expansion of frozen cells

Frozen cells were rapidly thawed in a water bath (37°C) and added to a culture flask (BD Biosciences®, 355001) containing an adequate volume of cell culture medium.

MCF-7 cell culture medium was comprised of Eagle’s minimal essential medium (EMEM; Sigma-Aldrich®, M4655) containing 10% foetal calf serum (FCS, Gibco®, 10270), 1% GlutaMAX (containing glucose, pyruvate and 2mM glutamine) (Gibco®, 35050-038), 1% non-essential amino acids (Sigma-Aldrich®, M7145) and 1% (v/v) antibiotic/antimycotic solution (containing 100U/ml penicillin, 100µg/ml streptomycin, and 0.5ng/ml amphotericin) (Sigma-Aldrich®,
A5955. T/C28a2 cells were grown in 1:1 Dulbecco’s Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12, Gibco®, 11320-033) supplemented with 10% FCS and 1% (v/v) antibiotic/antimycotic solution.

Cells were incubated overnight (37°C, 5% CO2) (Galaxy S®, Wolf Laboratories) for adherence and medium was changed after 24 hours in culture. Cells were then cultured in appropriate medium in an incubator at 37°C and 5% CO2. Cells were passaged when 70-80% confluence was reached.

3.4.3 Expansion of cells

For passaging, cells were washed with PBS and incubated with 5ml of 0.5 g/L trypsin/0.2 g/L EDTA (Sigma-Aldrich®, T4174) (5 minutes, 37°C, 5%CO2) in PBS. An equal volume of medium was then added to the flask, a cell count was performed and cells were reseeded in an appropriate number of culture flasks at a density of 20,000 cells/cm² for further expansion.

3.4.4 Cell counts

For counting cells, 10µL of cell suspension was pipetted into a haemocytometer and an average cell count from the four girds used was calculated (cells touching the upper and left edges were included in the count). The average cell count was multiplied by 10,000 to give the number of cells per mL.
3.4.5 Freezing of cells

When appropriate, a proportion of cultured cells were frozen for future use. For freezing, cells were trypsinised and counted, as described above. The volume of cell solution containing the appropriate number of cells to freeze was then transferred to a 50mL centrifuge tube (BD Biosciences) and centrifuged (21°C, 400G, 5 minutes) (Universal 320, Hettich®). The supernatant was discarded and the cell pellet re-suspended in Freezing Media containing DMSO (Gibco®, 12648-010) at a density of 1x10^6 cells/mL. The cell solution was then split into fully labelled Cryo MCT Vials (Nunc A/S®) (1mL each), which were placed in controlled rate freezing containers and stored at -80°C.

3.4.6 Development of a method to extract RNA from fixed, permeabilised, labelled and sorted cells

3.4.6.1 Labelling of cells

All experiments were performed in triplicate and all reagents and instruments used were certified RNAse/DNAse-free. To mimic the small numbers of cells expected from a single foetal spine, 5x10^4 MCF-7 cells were used in each experiment.

In preliminary experiments it was found that, after fixation and permeabilisation and independently of the fixation and permeabilisation method utilised, cells were
difficult to pellet at 400G and, for that reason, a higher speed (2000G) was used to pellet cells after this step. This speed was found to allow for easier pelleting without lysing the cells – which happened at higher speeds (3000G and 4000G).

Monolayer-expanded MCF-7 cells were trypsinised, counted and appropriate numbers of cells transferred to a MCT and centrifuged. Pelleted cells were washed in 1mL of PBS, centrifuged (5 minutes at 4ºC at 400G) and PBS was aspirated to waste. Then, cells were re-suspended in the appropriate fixation and permeabilisation agent and incubated for 10 minutes at 4ºC, as described in Table 3.1 (PBS was used as control).

**Table 3.1** Fixation and permeabilisation agents tested. PBS was used as control

<table>
<thead>
<tr>
<th>Fixation/ Permeabilisation agents</th>
</tr>
</thead>
<tbody>
<tr>
<td>95% Ethanol / 5% acetic acid (Fisher Scientific, BP-1185-500)</td>
</tr>
<tr>
<td>100% RNAlater (Sigma-Aldrich®, R0901)</td>
</tr>
<tr>
<td>50% RNAlater® / 50% PBS</td>
</tr>
<tr>
<td>100% ethanol</td>
</tr>
<tr>
<td>UM-Fix® (90% methanol (Sigma Aldrich®, 494437) + 10% polyethylene glycol (Sigma Aldrich®, 81253) (Yamada et al., 2010)</td>
</tr>
</tbody>
</table>

Fixed and permeabilised cells were subsequently washed in PBS and incubated in the dark with FITC-conjugated mouse monoclonal anti-KRT18 antibody (2µg/mL in PBS, Abcam®, Ab52459) for 30 minutes at 4ºC. After antibody incubation, cells were re-suspended in ice-cold PBS and immediately used for FACS. Equal number of cells were stained with FITC-conjugated mouse anti-IgG1 isotype control (2µg/mL in PBS, Sigma-Aldrich®, F6397) and used as negative control.
To determine the optimal antibody concentration a range of dilutions were tested (1/10 = 10µg/mL, 1/50 = 2µg/mL, 1/100 = 1µg/mL and 1/200 = 0.5µg/mL). Equivalent concentrations of the FITC-conjugated anti-IgG1 isotype control were used as negative controls. Optimal antibody concentration was not dependent on the fixation method utilised.

To exclude non-specific staining and to assess whether this methodology would allow accurate separation of KRT18\textsuperscript{+ve} from KRT18\textsuperscript{-ve} cells, MCF-7 cells were combined with T/C28a2 (a cell line that doesn’t express KRT18) in varying proportions (Table 3.2), the cell mixture was labelled with FITC-conjugated anti-KRT18 antibody and the proportion of stained cells was analysed.

<table>
<thead>
<tr>
<th>Table 3.2 MCF-7 and T/C28a2 cell mixtures used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell mixture</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>Cell mixture A</td>
</tr>
<tr>
<td>Cell mixture B</td>
</tr>
<tr>
<td>Cell mixture C</td>
</tr>
<tr>
<td>Cell mixture D</td>
</tr>
<tr>
<td>Cell mixture E</td>
</tr>
</tbody>
</table>

### 3.4.6.2 Sorting of KRT18\textsuperscript{+ve} and KRT18\textsuperscript{-ve} cells

Labelled cells were analysed using a CyAn ADP analyser flow cytometer (Beckman Coulter\textsuperscript{®}) or analysed and sorted using a FACS Aria II (BD Biosciences\textsuperscript{®}) cell sorter. Summit software v. 4.3 (Beckman Coulter\textsuperscript{®}) was used
to analyse flow cytometry data. In both machines, excitation was performed with the blue laser (488nm) and detection was performed with the 530/30 filter.

To sort KRT18<sup>+</sup>ve and KRT18<sup>-</sup>ve cells, single viable cells were firstly identified and cell aggregates (Figure 3.3. A), dead cells and cell debris (Figure 3.3. B) were gated out. Then, a sample containing cells labelled with FITC-conjugated anti-IgG1 isotype control was analysed and a gate was drawn to the right of the identifiable events (care was taken to exclude at least 98% of all events) – this gate would correspond to KRT18<sup>+</sup>ve events. The KRT18<sup>-</sup>ve gate was drawn to include the events with less intense FITC-conjugated anti-IgG1 isotype control staining (Figure 3.3. C). In order to avoid contamination, the intermediate population of cells was discarded. After identification of positive and negative gates, cells labelled with FITC-conjugated anti-KRT18 antibody were analysed and, after gating cell aggregates, dead cells and cell debris out, positive and negative gates were applied to define the KRT18<sup>+</sup>ve and KRT18<sup>-</sup>ve cell populations (Figure 3.3 D). KRT18<sup>+</sup>ve and KRT18<sup>-</sup>ve cells were collected into separate MCTs containing 350µL of RLT lysis buffer each (RNeasy micro plus kit, Qiagen®).
Figure 3.3 Scatter plots depicting the gating utilised to identify and isolate KRT18^{+ve} and KRT18^{-ve} cells (example taken from sample M659 (15 WPC)). (A) A gate (P1) was drawn around single events (low forward scatter height vs. side scatter area). (B) A second gate (P2) was drawn that excluded cell debris and dead cells (low forward scatter height vs. side scatter height). (C) Cells labelled with FITC-conjugated anti-IgG1 isotype control were detected with the 530/30 filter and analysed: a positive gate (P3) was drawn to the right of at least 98% of all detectable events and a negative gate (P4) was drawn to include the events with the least detected fluorescence. (D) The gates drawn in C) were applied to a sample labelled with FITC-conjugated anti-KRT18 antibody. Events in the P3 area were considered KRT18^{+ve} and those in the P4 area were considered KRT18^{-ve}. 
3.4.6.3 RNA extraction, quantity and quality assessment and amplification to cDNA

RNA was extracted using the RNeasy Micro Plus kit (Qiagen®, 74034). Sorted cells were lysed with RLT lysis buffer. As cells were sorted into the lysis buffer within a sorting sheath fluid (PBS), whose quantity varied depending on the number of cells within the original cell mixture, additional RLT lysis buffer was added until a final RLT buffer/ sorting sheath fluid ratio of 3/1 was obtained. Cell lysate was then warmed to 37°C before homogenisation. For homogenisation, the cell lysate was pipetted into a QIAshredder spin column (Qiagen®, 79654) placed in a 2mL MCT and centrifuged for 2 minutes at 19,700G in a centrifuge (Thermo Electron Corporation®, IEC CL31R). The homogenised lysate was then transferred (750µL at a time) to a gDNA Eliminator spin column placed in a 2mL collection MCT and centrifuged for 30 seconds at 10,000G and the flow-through reserved; this step was repeated until all lysate was passed through the gDNA Eliminator spin column. Next, 1 volume of freshly prepared molecular grade 70% v/v ethanol was added to the DNA-free lysate, mixed well by pipetting and transferred (750µL at a time) to a RNeasy MinElute spin column placed in a 2mL collection MCT. The column was then centrifuged for 30 seconds at 10,000G and the flow-through was discarded; this step was repeated using the same column until passing all cell homogenate/ ethanol mixture through the membrane. Using the same column and collection MCT, 700µL of Buffer RW1 was added to the column, centrifuged for 30 seconds at 10,000G and the flow-through was discarded. Next, 500µL of Buffer RPE was added to the column and centrifuged for 30 seconds at 10,000G to wash the spin column; the flow-through was
discarded. Next, and to ensure complete ethanol removal from the membrane, 500µL of buffer RPE was added to the column and centrifuged for 2 minutes at 10,000G. After centrifugation, the column was carefully removed from the collection MCT, placed in a new 2mL collection MCT and centrifuged for 2 minutes at 19,700G. Finally, the column was carefully removed from the collection MCT and placed in a 1.5mL collection MCT and 15µL of molecular grade water (warmed to 60ºC) was added to the column; the column was then centrifuged at 19,700G for 4 minutes, to elute the RNA. To allow for higher RNA recovery, the eluent was re-pipetted through the same column and centrifuged again at 19,700G for 4 minutes. Eluted RNA was stored at -80ºC.

3.4.6.4 Assessment of RNA quantity, quality and quality

RNA quantity, quality, purity and integrity were analysed using the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific®) and the Agilent 2100 Bioanalyser (Agilent Technologies®). Parameters used for RNA assessment were concentration, 260/280 and 260/230 ratios (values close to 2 represent non-contaminated samples) and RNA integrity number (RIN, ranges from 0-10, with 10 being the highest degree of integrity).
3.4.7 RNA extraction from KRT18<sup>+</sup> and KRT18<sup>-</sup> foetal spine cells

3.4.7.1 Extraction, KRT18-labelling and sorting of foetal spine cells

Foetal sample M644 (9 WPC) was acquired as described in 2.3.1, dissected to obtain spinal tissue (containing IVDs and VBs) as described in 2.3.2 and placed in PBS.

Immediately after dissection, sample M644 was cut into small fragments and digested overnight in medium (alpha-MEM (Sigma-Aldrich®, M4526) containing 1% (v/v) antibiotic/antimycotic solution, 2.5mg/mL type II collagenase (Gibco®, 17101-015) and 0.1% (w/v) hyaluronidase (Sigma-Aldrich®, H3506) in an orbital shaker at 37°C. After digestion, the cell solution was re-suspended for 1 hour at 37°C in non-enzymatic cell dissociation solution (Sigma-Aldrich®, C5789) to dissociate notochordal cell clusters, as previously described (Guehring et al., 2009). After dissociation, cells were washed in RNase-free PBS and sieved through a 40µm cell filter. Cells were visualised under an inverted microscope (Leitz® Diavert).

Single foetal spine cells were fixed/permeabilised, labelled and sorted using the methodology developed in 3.4.6.1 and 3.4.6.2. KRT18<sup>+</sup> and KRT18<sup>-</sup> cells were collected into different MCT containing 350µL of RLT buffer®. RNA was extracted as detailed in 3.4.6.3 and its quantity, quality, purity and integrity were analysed as described in 3.4.6.4.
3.4.7.2 RNA amplification and purification

Due to the limited quantity of RNA obtained, RNA was amplified to SPIA® (Single Primer Isothermal Amplification) cDNA using the Ovation Pico WTA v2 kit (NuGen Technologies®, 3302). This is comprised of 4 steps: 1) primer annealing and generation of first strand cDNA, 2) generation of second strand cDNA and 3) purification of cDNA and 4) amplification to SPIA® cDNA.

For primer annealing, 2µL of first strand primer were added to 5µL of RNA in a 0.2mL MCT, mixed by pipetting and placed in a thermal cycler on program 1 (Table 3.3); resulting mixture was placed on ice. For first strand cDNA synthesis, a mixture consisting of 2.5µL first strand buffer mix and 0.5µL first strand enzyme mix was added to each MCT containing the first strand primer/RNA mixture, mixed by pipetting 5 times and placed in a pre-cooled thermal cycler on program 2 (Table 3.3).

For second strand cDNA synthesis a mixture comprising 9.7µL of second strand buffer mix and 0.3µL of second strand enzyme mix was added to each first strand reaction MCT, mixed by pipetting 5 times and placed in a pre-cooled thermal cycler on program 3 (Table 3.3).

For cDNA purification, 32µL of Agencourt RNAClean XP beads (Beckman Coulter®, A32782) were added to each second strand reaction MCT, mixed by pipetting 10 times, transferred to a 96-well PCR plate (Star Lab®, E1403-6200) and incubated at room temperature for 10 minutes. After incubation, the MCTs were transferred to a 96-well magnet ring stand (Life Technologies®, AM10050) and allowed to stand for 5 minutes to completely clear the solution of beads. Next, keeping the beads on the magnet, 45µL of binding buffer were pipetted off and
discarded, after which 200µL of freshly prepared 70% v/v ethanol were added to each well, allowed to stand for 30 seconds and then removed and discarded. This wash step was repeated twice. After pipetting the ethanol off, the beads were left to dry on the magnet for 15 minutes.

For SPIA® amplification a mixture of 50µL of SPIA® buffer mix, 25µL of SPIA® primer mix and 25µL of SPIA® enzyme mix was added to each MCT containing the double-stranded cDNA bound to the dried beads, mixed well by pipetting and placed in a pre-cooled thermal cycler programmed to run program 4 (Table 3.3.).

After incubation, the MCTs were removed from the cycler and placed on ice. To avoid introducing previously amplified cDNA into the MCTs, this final step was performed outside the pre-amplification workspace and using dedicated consumables and equipment. The plate was transferred to the magnet stand and allowed to stand for 5 minutes to completely clear the solution of beads and the supernatant containing the amplified cDNA was transferred to 1.5mL MCTs; the beads were discarded.
Table 3.3 Thermal cycling programming for RNA amplification.

<table>
<thead>
<tr>
<th>First strand cDNA synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program 1: Primer Annealing</strong></td>
</tr>
<tr>
<td><strong>Program 2: First Strand Synthesis</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second strand cDNA synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program 3: Second Strand Synthesis</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPIA® amplification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program 4: SPIA® Amplification</strong></td>
</tr>
</tbody>
</table>

Following amplification, cDNA was immediately purified using the RNeasy MinElute kit (Qiagen®, 74204). Three hundred microliters of Buffer ERC was added to each SPIA® cDNA sample (100 µL), MCTs were vortexed briefly to mix, centrifuged briefly, transferred to a MinElute spin column placed in a 2mL collection MCT and centrifuged for 1 minute at 19,700G at room temperature. The flow-through was discarded. Then, 750 µL of Buffer PE was added to the column and centrifuged for 1 minute at 19,700G. After centrifugation, the flow-through was discarded and the column was centrifuged for 2 additional minutes at 19,700G to remove any residual buffer. Finally, the column was carefully placed in a 1.5mL MCT, 15 µL of nuclease-free water (previously warmed to 60°C) was added, the column was allowed to stand for 1 minute and then centrifuged at 19,700G for 4 minutes. To allow for higher cDNA recovery, the eluent was re-pipetted through the column, allowed to stand for 1 additional minute and centrifuged at 19,700G for 4 minutes. cDNA quantity and quality after
amplification and purification was analysed using the Nanodrop ND-1000 Spectrophotometer. Samples were diluted to 5ng/µL and stored at -20°C.

3.4.7.3 Real-time quantitative polymerase chain reaction analysis

Real-time quantitative PCR (qRT-PCR) using gene specific primers was performed on cDNA from KRT18^ve and KRT18^-ve cells using TaqMan Universal PCR Master Mix (Applied Biosystems®, 4304437) (Minogue et al., 2010b, Minogue et al., 2010a). Each reaction was performed in triplicate.

For qRT-PCR 2µl of 5ng/µL cDNA from each sample was added to 8µL of a pre-prepared master mix (Table 3.4.) in a 0.2mL 96-well PCR plate. The plate was sealed with optical adhesive film (ABI®, 4311971), vortexed for 10 seconds and centrifuged for 1 minute at 2,000G. The plate was then run on an Applied Biosystems StepOnePlus® instrument using the settings described in table 3.5. Total foetal cDNA was used as positive control and water as negative control. Primers and probes used are detailed in Table 3.6.

Table 3.4 TaqMan qRT-PCR assay mastermix. Volumes are per reaction.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x mastermix (Sigma Aldrich®, LuminoCt qPCR Ready Mix, L66669-2000RXN)</td>
<td>5</td>
</tr>
<tr>
<td>Forward primer (concentration detailed in table 3.6)</td>
<td>1</td>
</tr>
<tr>
<td>Reverse primer (concentration detailed in table 3.6)</td>
<td>1</td>
</tr>
<tr>
<td>Probe (250nM)</td>
<td>0.5</td>
</tr>
<tr>
<td>40x ROX (Sigma Aldrich®, LuminoCt qPCR Ready Mix, L66669-2000RXN)</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Table 3.5 Thermal cycler programming for qRT-PCR.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature (°C)</th>
<th>Length of time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95.0</td>
<td>20 seconds</td>
</tr>
<tr>
<td>2</td>
<td>95.0</td>
<td>1 second</td>
</tr>
<tr>
<td>3</td>
<td>60.0</td>
<td>20 seconds</td>
</tr>
<tr>
<td>4</td>
<td>Repeat steps (2-3) 40 times</td>
<td>N/A</td>
</tr>
<tr>
<td>Gene name</td>
<td>Gene symbol</td>
<td>NCBI RefSeq</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate</td>
<td>GAPDH</td>
<td>NM_001256799</td>
</tr>
<tr>
<td>dehydrogenase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keratin 18</td>
<td>KRT18</td>
<td>NM_000224.2</td>
</tr>
<tr>
<td>Keratin 19</td>
<td>KRT19</td>
<td>NM_002276.4</td>
</tr>
<tr>
<td>Galectin 3</td>
<td>GAL3</td>
<td>NM_002306.3</td>
</tr>
<tr>
<td>Connective tissue growth factor</td>
<td>CTGF</td>
<td>NM_001901.2</td>
</tr>
</tbody>
</table>
Relative gene expression of each target gene was calculated by normalising it to the expression of the housekeeping gene (GAPDH) using the $2^{-\Delta\Delta Ct}$ method, as previously described (Livak and Schmittgen, 2001).

Statistical analysis was performed using GraphPad InStat software (GraphPad Software Inc.) using the Mann-Whitney U-test; p values < 0.05 were defined as being representative of a significant difference.

### 3.4.8 Complimentary DNA (cDNA) microarrays

SPIA® cDNA was fragmented and labelled for analysis using the Encore Biotin Module (NuGen Technologies®, 4200). Fragmentation and labelling was performed by staff within the Genomic Technologies Facility (Faculty of Life Sciences, The University of Manchester), according to manufacturer’s instructions.

For each microarray hybridisation, 2.5µg input of cDNA was used with the gene chip Human Genome U133 Plus 2.0 microarray (Affymetrix®, 900466); this was performed by staff within the Genomic Technologies Facility (Faculty of Life Sciences, The University of Manchester), according to manufacturer’s instructions. Technical quality control was performed with dChip software (Li and Wong, 2001).
3.4.9 Microarray analysis

Microarray analysis was performed by Dr Leo Zeef, a bioinformatician, within the Genomic technologies facility. Analysis of differentially expressed genes was performed with the PUMA package (Pearson et al., 2009) and differentially expressed genes were defined as those with a minimum fold change of 3 and a probability of positive log ratio (PPLR) <0.05 or >0.95.

Interactive pathway analysis software (IPA, Ingenuity Systems®) was used to identify upstream regulators of differentially expressed genes. Upstream regulators are molecules that can affect the expression of molecules in the dataset. To identify such molecules, IPA analyses the differentially expressed genes (dataset molecules) and identifies molecules that have an expected effect on them; this information is stored in the Ingenuity Knowledge Database®, a database detailing known molecular interactions (curated from published scientific literature). Upstream regulators can be predicted to be active (positive upstream regulators) or to be inhibited (negative upstream regulators) and this information is given by the z score. The z score is positive when the regulator is predicted to be active (i.e., when it is predicted to upregulate molecules that are upregulated in the dataset and to downregulate molecules that are downregulated in the dataset) and it is negative when the regulator is predicted to be inhibited (i.e., when it is predicted to upregulate molecules that are downregulated in the dataset and to downregulate molecules that are upregulated in the dataset). Upstream regulator molecule list was filtered to include only those with a predictable activating state (Figure 3.4).
Figure 3.4 Graphical explanation of IPA upstream regulator analysis. (A) Microarrays provide a list of differentially expressed genes (up or downregulated). (B) IPA identifies molecules that are predicted to induce the gene expression changes seen in the dataset. Those molecules can be predicted to be active or inhibited.

To identify KRT18$^{+ve}$ cell surface markers, the upregulated genes were screened using the European Bioinformatics Database® (http://www.ebi.ac.uk), the Uniprot Knowledge base® (http://www.uniprot.org/uniprot) and the NCBI Protein® database (http://www.ncbi.nlm.nih.gov/protein) to identify those that code for cell membrane-bound proteins with an extracellular domain.
3.4.10 Immunolocalisation of HGF and c-Met

To assess if notochordal cells express hepatocyte growth factor (HGF) and its receptor c-Met, the expression of these proteins was analysed in a cohort of foetal samples. Samples M234 (3.5 WPC), M320 (5.5 WPC), M741 (7 WPC), M779 (7.5 WPC), M636 (9.5 WPC), M685 (10 WPC), M404 (11.5 WPC), M742 (12 WPC), M776 (13 WPC), M739 (14 WPC) and M777 (17 WPC) were used (detailed in Table 2.1).

Optimal antibody concentration and antigen retrieval methods were identified as described in 2.3.8. Table 3.7 details the optimised antigen retrieval methods and concentrations for these two proteins. Antibody stock concentrations were diluted in 2% BSA in TBS.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Antibodies</th>
<th>Antigen retrieval method</th>
</tr>
</thead>
<tbody>
<tr>
<td>HGF</td>
<td>1.33 µg/mL rabbit polyclonal anti-HGF IgG1 (Abcam®, ab83760)</td>
<td>Heat citrate</td>
</tr>
<tr>
<td>c-Met</td>
<td>2.5µg/mL mouse monoclonal anti-Met IgG1 (Abcam®, ab59884)</td>
<td>Heat citrate</td>
</tr>
</tbody>
</table>
3.5 RESULTS

3.5.1 Development of a method to extract RNA from fixed, permeabilised, labelled and sorted cells

The optimal FITC-conjugated anti-KRT18 antibody concentration was empirically determined as 2µg/mL. Lower antibody concentrations showed a double fluorescence peak indicative of only a proportion of cells being labelled and higher antibody concentrations did not allow for an improved discrimination between stained and unstained cells (Figure 3.5).

**Figure 3.5** Determination of the ideal FITC-conjugated anti-KRT18 antibody concentration. Flow cytometry overlay histograms depicting MCF-7 cells labelled with FITC-conjugated anti-KRT18 antibody (blue) and FITC-conjugated anti-IgG1 isotype control (red) at different concentrations: (A) 2µg/mL, (B) 1µg/mL and (C) 0.5µg/mL.

Using the optimised antibody concentration and when different mixtures of MCF-7 (KRT18<sup>+</sup>) and T/C28a2 (KRT18<sup>-</sup>) cells were analysed, there was a clear separation between both cell types, indicating that, using this methodology,
KRT18<sup>ve</sup> and KRT18<sup>-ve</sup> cells could be clearly identified and separated (Figure 3.6).

**Figure 3.6** Analysis of FITC-conjugated anti-KRT18 antibody staining in different mixtures of MCF-7 and T/C28a2 cells. **A-E**: Flow cytometry overlay histograms showing FITC-conjugated anti-KRT18 antibody (blue) and FITC-conjugated anti-IgG1 isotype control (red) staining mixtures (**A**) 100% MCF-7 cells, (**B**) 75% MCF-7, 25% T/C28a2 cells, (**C**) 50% MCF-7, 50% T/C28a2 cells, (**D**) 25% MCF-7, 75% T/C28a2 cells and (**E**) 100% T/C28a2 cells. (**F**) Scatter plot of the labelling, detected with the 530/30 filter, of mixture C, showing good discrimination between MCF-7 and T/C28a2 cells.

RNA analysis following incubation of MCF-7 cells with different fixation and permeabilisation agents (Figure 3.2.B, Step 1) showed that there was a significant decrease in RNA concentration with all agents when compared with the control. The mixture containing 95% ethanol/5% acetic acid and the reagent 100% RNALater® were the agents that showed the least detrimental effect on the RNA purity (higher 260/280 and 260/230 ratios) and higher integrity (higher RIN
values) and, for that reason, were chosen for the subsequent steps of optimisation (Figure 3.7)

**Figure 3.7** Effect of different fixation and permeabilisation agents on the RNA quantity, quality and integrity. (A) Concentration, (B) 260/280 ratio, (C) 260/230 ratio and (D) RIN of RNA extracted from 50,000 MCF-7 cells subjected to different fixation and permeabilisation agents. PBS was used as control. *p<0.05

To assess the effect of antibody labelling, cells were fixed with either 95% ethanol/5% acetic acid or 100% RNAlater® followed by incubation with PBS or FITC-conjugated anti-KRT18 antibody (Figure 3.2B, Step 2). Labelling with KRT18 antibody was found to have no detrimental effect on the RNA quantity and purity, but to decrease the 260/230 ratio in cells fixed with RNAlater® (Figure 3.8).
Figure 3.8 Effect of antibody labelling on the RNA quantity, quality and integrity. (A) Concentration, (B) 260/280 ratio, (C) 260/230 ratio and (D) RIN of RNA extracted from 50,000 MCF-7 cells fixed with either 95% ethanol/5% acetic acid or 100% RNAlater and incubated with PBS or FITC-anti-KRT18. * represents p<0.05.

However, after fixing, permeabilising, labelling and sorting (Figure 3.2.B, Step 3), the RNA concentration and RIN values were significantly lower when RNAlater® was used (Figure 3.9).
Figure 3.9 Effect of cell sorting on the RNA quantity, quality and integrity. (A) Concentration, (B) 260/280 ratio, (C) 260/230 ratio and (D) RIN of RNA extracted from 50,000 fixed, permeabilised, labelled and sorted MCF-7 cells. * represents p<0.05.

These results demonstrated that a protocol using 95% ethanol/5% acetic acid, followed by labelling with FITC-anti-KRT18 and cell sorting allowed for the purification of RNA from low numbers of cells, with sufficient quality to be used for microarray analysis.
3.5.2 Isolation and RNA extraction from notochordal and sclerotomal foetal spine cells

Following optimisation of the cell fixation, permeabilisation and labelling method using MCF-7 cells, this method was utilised to isolate KRT18<sup>+</sup> and KRT18<sup>-</sup> human foetal spine cells.

Microdissected spinal tissue from specimen M644 (9 WPC) was enzymatically digested and cells were dissociated. Microscopic analysis of extracted cells showed a heterogeneous population of small round cells and larger vacuolated cells; the proportion of vacuolated cells was smaller than that of small round cells; clustered cells were still present although the morphology of the clustered cells was not evident (Figure 3.10).

**Figure 3.10** Microscopic view of foetal spine cells extracted from sample M644. A smaller proportion of large vacuolated cells (arrow) were seen within a larger population of small round cells (square). Some small round cells were clustered (circle). * depicts cell debris. (Magnification: 10x).
Cells were then fixed and permeabilised with 95% ethanol/5% acetic acid and labelled with FITC-conjugated anti-KRT18 antibody as described above.

Analysis of labelled cells revealed a distinct FITC-positive population (detected with the 530/30 filter) and sorting yielded 3,800 KRT18$^{+ve}$ and 40,000 KRT18$^{-ve}$ events (Figure 3.11).

**Figure 3.11** Sorting of sample M644. Scatter plot depicting the fluorescence (detected with the 530/30 filter) vs. granularity (side scatter) of M644 spine cells labelled with FITC-conjugated anti-KRT18 antibody, after gating out debris, dead cells and cell aggregates. P3 represents events gated as being KRT18$^{+ve}$. Due to the small numbers of sorted cells all remaining events were considered KRT18$^{-ve}$. P3 was drawn to exclude more than 98% of the same cells when labelled with FITC-conjugated IgG1 isotype control.
After RNA extraction, 13.7ng/µL of KRT18\(^{+ve}\) and 12.4ng/µL of KRT18\(^{-ve}\) RNA was obtained. This RNA had high 260/280 ratio and RIN values but low 260/230 ratio (Figure 3.12).

**Figure 3.12** Analysis of the RNA extracted from KRT18\(^{+ve}\) and KRT18\(^{-ve}\) sorted foetal spine cells. (A) Concentration, (B) 260/280 ratio, (C) 260/230 ratio and (D) RIN. As RNA was extracted from only one specimen, no statistical analysis was performed.

RNA was amplified by 25-fold to 340.2ng/µL of KRT18\(^{+ve}\) cDNA and by 45-fold to 556.4ng/µL of KRT18\(^{-ve}\) cDNA; this amplification substantially improved the 260/230 ratio for each sample (Figure 3.13).
Figure 3.13 Analysis of the cDNA amplified from RNA extracted from KRT18⁺ve and KRT18⁻ve sorted foetal spine cells. (A) Concentration, (B) 260/280 ratio, (C) 260/230 ratio. As RNA was extracted from only one specimen, no statistical analysis was performed.

3.5.3 Real time quantitative PCR validation of notochordal cell separation

To validate the accuracy of the methodology employed to separate notochordal (KRT18⁺ve) from sclerotomal foetal spine cells (KRT18⁻ve) the gene expression of KRT18, KRT19, T, GAL3 and CTGF was analysed by qRT-PCR. Comparison between both cell types showed that KRT18⁺ve sorted cells had higher relative expression of all genes than KRT18⁻ve cells (Figure 3.14).
Figure 3.14 qRT-PCR validation of KRT18 cell separation. Gene expression of KRT18, KRT19, GAL3 and CTGF was first normalised to GAPDH expression; relative gene expression values for KRT18<sup>ve</sup> cells were normalised to the relative expression of KRT18<sup>ve</sup> cells and presented on a log scale. As RNA was extracted from only one specimen, no statistical analysis was performed. Error bars represent standard error of the mean of the technical repeats.

3.5.4 Microarray analysis

Microarray analysis identified 1020 up-regulated and 859 down-regulated genes with a fold-change difference of at least 3 between both cell types. Table 3.8 depicts the top 10 positive markers (top genes with increased differential expression in KRT18<sup>ve</sup> cells compared to KRT18<sup>ve</sup> cells) and the top 10 negative markers (genes with decreased differential expression in KRT18<sup>ve</sup> cells compared to KRT18<sup>ve</sup> cells).
Within the top 125 upregulated KRT18\textsuperscript{ve} genes, 10 coded for membrane-bound proteins with an extracellular domain and that can, therefore, be used as notochordal cell surface markers. Those markers are described in table 3.9.
Table 3.8 Top differentially expressed (increased and decreased expression) genes between KRT18\textsuperscript{+}ve and KRT18\textsuperscript{-}ve sorted foetal spine cells. Fold change, p-values and function of the encoded protein are also presented.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Fold change</th>
<th>Log\textsubscript{2} ratio</th>
<th>p-value</th>
<th>Function of the encoded protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>GENES WITH INCREASED EXPRESSION IN KRT18\textsuperscript{+}VE CELLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRB14</td>
<td>Growth factor receptor-bound protein 14</td>
<td>195.155</td>
<td>7.608</td>
<td>3.13 x 10(^{-3})</td>
<td>Binds to, and suppresses signals from the activated insulin receptor (Hemming et al., 2001).</td>
</tr>
<tr>
<td>HLA-DQB1</td>
<td>Major histocompatibility complex, class II, DQ beta 1</td>
<td>122.640</td>
<td>6.938</td>
<td>4.16 x 10(^{-3})</td>
<td>HLA class II molecules are usually expressed in the cell surface of antigen presenting cells (such as lymphocytes, dendritic cells and macrophages), which are involved in the immune system (Prifti-Kurti et al., 2014).</td>
</tr>
<tr>
<td>SLC19A1</td>
<td>Solute carrier family 19 (folate transporter), member 1</td>
<td>94.061</td>
<td>6.556</td>
<td>1.49 x 10(^{-2})</td>
<td>Involved in cellular folate uptake (Stanislawska-Sachadyn et al., 2009).</td>
</tr>
<tr>
<td>FGF10</td>
<td>Fibroblast growth factor 10</td>
<td>90.925</td>
<td>6.507</td>
<td>3.47 x 10(^{-2})</td>
<td>Mitogenic and cell survival activities. Involved in embryonic development and tissue repair (Emoto et al., 1997).</td>
</tr>
<tr>
<td>SAMD9L</td>
<td>Sterile alpha motif domain-containing protein 9-like</td>
<td>88.772</td>
<td>6.472</td>
<td>6.11 x 10(^{-4})</td>
<td>Protein that is widely expressed in adult and foetal tissues. Unknown function (Li et al., 2007).</td>
</tr>
<tr>
<td>SEL1L2</td>
<td>Sel-1 suppressor of lin-12-like 2</td>
<td>69.851</td>
<td>6.126</td>
<td>1.76 x 10(^{-2})</td>
<td>Unknown.</td>
</tr>
<tr>
<td>ADORA3</td>
<td>Adenosine A3 receptor</td>
<td>62.815</td>
<td>5.973</td>
<td>8.20 x 10(^{-3})</td>
<td>G-coupled receptor with cardioprotective functions during cardiac ischemia and that has been reported to have both neuroprotective and neurodegenerative effects (Bouma et al., 1997, Chen et al., 2006a).</td>
</tr>
<tr>
<td>MYOZ2</td>
<td>Myozenin 2</td>
<td>62.738</td>
<td>5.971</td>
<td>5.57 x 10(^{-3})</td>
<td>Sarcomeric protein involved in calcineurin signalling (Frey et al., 2000).</td>
</tr>
<tr>
<td>CADPS</td>
<td>Ca\textsuperscript{++}-dependent secretion activator</td>
<td>62.590</td>
<td>5.968</td>
<td>6.07 x 10(^{-4})</td>
<td>Neural/endocrine-specific cytosolic and peripheral membrane protein (Loyet et al., 1998).</td>
</tr>
<tr>
<td>CALCRL</td>
<td>Calcitonin receptor-like</td>
<td>61.014</td>
<td>5.931</td>
<td>1.84 x 10(^{-3})</td>
<td>G protein-coupled receptor that is regulated by receptor activity-modifying proteins (RAMPs). Deficiencies in the latter have been associated with vertebrae and intervertebral disc abnormalities (Kadmiel et al., 2011).</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Fold Change</td>
<td>Log2 Fold Change</td>
<td>p-Value</td>
<td>Function and Notes</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>NOL12</td>
<td>Nucleolar protein 12</td>
<td>-146.585</td>
<td>-7.196</td>
<td>1.20E x 10^2</td>
<td>Unknown</td>
</tr>
<tr>
<td>C5AR1</td>
<td>Complement component 5a receptor 1</td>
<td>-138.632</td>
<td>-7.115</td>
<td>1.40E x 10^2</td>
<td>G protein-coupled receptor for the complement component C5a. Receptor activation stimulates chemotaxis, granule enzyme release, intracellular calcium release and superoxide anion production (Gerard and Gerard, 1994).</td>
</tr>
<tr>
<td>SERPINB6</td>
<td>Serpin peptidase inhibitor, clade B (ovalbumin), member 6</td>
<td>-132.476</td>
<td>-7.050</td>
<td>3.79 x 10^4</td>
<td>Placental thrombin inhibitor. Involved in gonad development (Hayashi et al., 2011).</td>
</tr>
<tr>
<td>DAPK1</td>
<td>Death-associated protein kinase 1</td>
<td>-100.758</td>
<td>-6.655</td>
<td>1.37 x 10^2</td>
<td>Positive mediator of interferon-γ induced cell death and regulator of autophagy (Gade et al., 2014).</td>
</tr>
<tr>
<td>ADAMTS1</td>
<td>ADAMTS-like 1</td>
<td>-70.340</td>
<td>-6.136</td>
<td>1.75 x 10^2</td>
<td>ECM protein that resembles members of the ADAMTS family but lacking the propeptide region and the metalloproteinase and Disintegrin-like domains (Hirohata et al., 2002).</td>
</tr>
<tr>
<td>TMCO5A</td>
<td>Transmembrane and coiled-coil domains 5A</td>
<td>-62.241</td>
<td>-5.960</td>
<td>2.10 x 10^2</td>
<td>Unknown</td>
</tr>
<tr>
<td>MTMR3</td>
<td>Myotubularin related protein 3</td>
<td>-58.115</td>
<td>-5.861</td>
<td>1.03 x 10^2</td>
<td>Protein that binds to phophoinositide lipids and is involved in autophagy (Taguchi-Atarashi et al., 2010).</td>
</tr>
<tr>
<td>MYH3</td>
<td>Myosin, heavy chain 3, skeletal muscle, embryonic</td>
<td>-54.039</td>
<td>-5.756</td>
<td>1.35 x 10^2</td>
<td>Contractile protein that converts chemical energy into mechanical energy through the hydrolysis of ATP. Mutations in its gene have been associated with two congenital contracture (arthrogryposis) syndromes: Freeman-Sheldon syndrome and Sheldon-Hall syndrome (Beck et al., 2014, Alvarado et al., 2011).</td>
</tr>
<tr>
<td>HAS3</td>
<td>Hyaluronan synthase 3</td>
<td>-53.359</td>
<td>-5.738</td>
<td>5.04 x 10^2</td>
<td>Protein involved in the synthesis of hyaluronic acid, a major constituent of the ECM (Jacobson et al., 2000).</td>
</tr>
<tr>
<td>WNT4</td>
<td>Wingless-type MMTV integration site family, member 4</td>
<td>-50.455</td>
<td>-5.657</td>
<td>5.11 x 10^3</td>
<td>Canonical Wnt receptor signalling pathway protein involved in the regulation of cell fate and patterning during embryogenesis, in response to hydrostatic pressure, in epithelial to mesenchymal transition and cellular response to TGF-β signalling (Colwell et al., 2006).</td>
</tr>
</tbody>
</table>
Table 3.9 Top 10 cell surface KRT18$^{+}$ve markers.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Fold change</th>
<th>Log$_2$ ratio</th>
<th>p-value</th>
<th>Function of the encoded protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADORA3</td>
<td>Adenosine A3 receptor</td>
<td>62.815</td>
<td>5.973</td>
<td>8.20x10$^{-03}$</td>
<td>See table 3.8.</td>
</tr>
<tr>
<td>TBXA2R</td>
<td>Thromboxane A2 receptor</td>
<td>39.849</td>
<td>5.316</td>
<td>1.72x10$^{-02}$</td>
<td>Member of the G protein-coupled receptor family that interacts with thromboxane A2 to induce platelet aggregation and regulate haemostasis (Huang et al., 2004).</td>
</tr>
<tr>
<td>CLDN12</td>
<td>Claudin 12</td>
<td>36.087</td>
<td>5.173</td>
<td>6.53x10$^{-03}$</td>
<td>Integral membrane protein and component of tight junction strands. Tight junctions act as physical barriers controlling solutes and water transport (Gunzel and Fromm, 2012).</td>
</tr>
<tr>
<td>CDH6</td>
<td>Cadherin 6, type 2, K-cadherin</td>
<td>31.137</td>
<td>4.961</td>
<td>2.08x10$^{-03}$</td>
<td>Calcium dependent cell-cell adhesion glycoprotein, involved in kidney development and endometrium and placenta formation (MacCalman et al., 1998, Shimazui et al., 2000)</td>
</tr>
<tr>
<td>ANPEP</td>
<td>Alanyl (membrane) aminopeptidase</td>
<td>26.912</td>
<td>4.750</td>
<td>2.11x10$^{-02}$</td>
<td>Enzyme located in several plasma membranes (small-intestine, renal microvilli). It has been proposed to be a biomarker of prostate cancer progression (Sorensen et al., 2013)</td>
</tr>
<tr>
<td>SELE</td>
<td>Selectin E</td>
<td>25.367</td>
<td>4.665</td>
<td>1.21x10$^{-05}$</td>
<td>Cell adhesion molecule expressed by cytokine-stimulated endothelial cells. It is involved in recruiting leucocytes to the sites of inflammation (Daxecker et al., 2002)</td>
</tr>
<tr>
<td>CD69</td>
<td>CD69 molecule</td>
<td>22.544</td>
<td>4.495</td>
<td>1.28x10$^{-02}$</td>
<td>Transmembrane C-type lectin protein induced by the activation of T lymphocytes and natural killer cells. It is involved in lymphocyte proliferation (Santis et al., 1994).</td>
</tr>
<tr>
<td>EPHA5</td>
<td>EPH receptor A5</td>
<td>22.106</td>
<td>4.466</td>
<td>1.47x10$^{-02}$</td>
<td>Receptor involved in nervous system development and modulation (Murai and Pasquale, 2002).</td>
</tr>
<tr>
<td>HLA-DQA1</td>
<td>Major histocompatibility complex, class II, DQ α1</td>
<td>17.115</td>
<td>4.097</td>
<td>2.12x10$^{-02}$</td>
<td>Cell surface receptor involved in the immune system by presenting peptides derived from extracellular proteins (He et al., 2014).</td>
</tr>
<tr>
<td>STAB1</td>
<td>Stabilin 1</td>
<td>16.858</td>
<td>4.075</td>
<td>5.03x10$^{-03}$</td>
<td>Transmembrane receptor involved in cell adhesion, angiogenesis, and lymphocyte homing (Politz et al., 2002).</td>
</tr>
</tbody>
</table>
IPA analysis was used to identify molecules predicted to be upstream of differentially expressed genes. There were 34 upstream regulators, with a predictable activation state (29 were predicted to be active and 5 to be inhibited) (Table 3.10).
Table 3.10 Upstream regulators of differentially expressed genes between KRT18<sup>ve</sup> and KRT18<sup>ve</sup> cells. Log<sub>2</sub> ratio represents differential expression of the upstream regulator between KRT18<sup>ve</sup> and KRT18<sup>ve</sup> cells (missing values correspond to molecules that were not differentially expressed between KRT18<sup>ve</sup> and KRT18<sup>ve</sup> cells); activation z-score represents the probability of being active, with positive values representing positive regulators (activated) and negative values representing negative regulators (inhibited). Molecule type is also presented.

<table>
<thead>
<tr>
<th>Upstream regulator</th>
<th>Molecule Type</th>
<th>Activation z-score</th>
<th>Log&lt;sub&gt;2&lt;/sub&gt; Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidermal growth factor (EGF)</td>
<td>Growth factor</td>
<td>3.736</td>
<td>-0.455</td>
</tr>
<tr>
<td>Carbon tetrachloride</td>
<td>Chemical toxicant</td>
<td>3.124</td>
<td></td>
</tr>
<tr>
<td>Extracellular signal-regulated kinases 1/2 (ERK1/2)</td>
<td>Group</td>
<td>2.818</td>
<td></td>
</tr>
<tr>
<td>Endothelin 1 (EDN1)</td>
<td>Cytokine</td>
<td>2.764</td>
<td>1.338</td>
</tr>
<tr>
<td>Interleukin 5 (IL5)</td>
<td>Cytokine</td>
<td>2.750</td>
<td>-4.643</td>
</tr>
<tr>
<td>Calcium ion (Ca&lt;sup&gt;2+&lt;/sup&gt;)</td>
<td>Chemical - endogenous</td>
<td>2.715</td>
<td></td>
</tr>
<tr>
<td>Wilms tumor 1 (WT1)</td>
<td>Transcription regulator</td>
<td>2.568</td>
<td>-2.521</td>
</tr>
<tr>
<td>Interleukin 3 (IL3)</td>
<td>Cytokine</td>
<td>2.466</td>
<td>-0.177</td>
</tr>
<tr>
<td>Interleukin 11 (IL11)</td>
<td>Cytokine</td>
<td>2.408</td>
<td>0.252</td>
</tr>
<tr>
<td>FBJ murine osteosarcoma viral oncogene homolog (FOS)</td>
<td>Transcription regulator</td>
<td>2.368</td>
<td>0.178</td>
</tr>
<tr>
<td>Hepatocyte growth factor (HGF)</td>
<td>Growth factor</td>
<td>2.353</td>
<td>3.451</td>
</tr>
<tr>
<td>Jun proto-oncogene (JUN)</td>
<td>Transcription regulator</td>
<td>2.317</td>
<td>-0.301</td>
</tr>
<tr>
<td>Chemokine (C-C motif) ligand 2 (Ccl2)</td>
<td>Cytokine</td>
<td>2.236</td>
<td></td>
</tr>
<tr>
<td>5-azacytidine</td>
<td>Chemical drug</td>
<td>2.236</td>
<td></td>
</tr>
<tr>
<td>Coagulation factor II (thrombin) receptor (F2R)</td>
<td>G-protein coupled receptor</td>
<td>2.213</td>
<td>-1.379</td>
</tr>
<tr>
<td>Runx-related transcription factor 1 (RUNX1)</td>
<td>Transcription regulator</td>
<td>2.207</td>
<td></td>
</tr>
<tr>
<td>Indian hedgehog (IHH)</td>
<td>Enzyme</td>
<td>2.200</td>
<td>3.695</td>
</tr>
<tr>
<td>Thioacetamide</td>
<td>Chemical toxicant</td>
<td>2.156</td>
<td></td>
</tr>
<tr>
<td>Interferon, lambda 1 (IFNL1)</td>
<td>Cytokine</td>
<td>2.138</td>
<td>0.621</td>
</tr>
<tr>
<td>Triggering receptor expressed on myeloid cells 1 (TREM1)</td>
<td>Transmembrane receptor</td>
<td>2.102</td>
<td>1.224</td>
</tr>
<tr>
<td>Early growth response 1 (EGR1)</td>
<td>Transcription regulator</td>
<td>2.092</td>
<td>0.509</td>
</tr>
<tr>
<td>Fibroblast growth factor receptor 2 (FGFR2)</td>
<td>Kinase</td>
<td>2.082</td>
<td>-3.336</td>
</tr>
<tr>
<td>Interleukin 17A (IL17A)</td>
<td>Cytokine</td>
<td>2.032</td>
<td>3.191</td>
</tr>
<tr>
<td>Hexachlorobenzene</td>
<td>Chemical toxicant</td>
<td>2.000</td>
<td></td>
</tr>
<tr>
<td>Hypermethylated in cancer 1 (HIC1)</td>
<td>Other</td>
<td>2.000</td>
<td>2.769</td>
</tr>
<tr>
<td>HNF1 homeobox A (HNF1A)</td>
<td>Transcription regulator</td>
<td>2.000</td>
<td>-3.676</td>
</tr>
<tr>
<td>Jagged 1 (JAG1)</td>
<td>Growth factor</td>
<td>2.000</td>
<td>1.515</td>
</tr>
<tr>
<td>Gentamicin C</td>
<td>Chemical drug</td>
<td>2.000</td>
<td></td>
</tr>
<tr>
<td>K-252</td>
<td>Chemical - kinase inhibitor</td>
<td>2.000</td>
<td></td>
</tr>
<tr>
<td>Hemin</td>
<td>Chemical - endogenous</td>
<td>2.000</td>
<td></td>
</tr>
</tbody>
</table>
### MOLECULES PREDICTED TO BE INHIBITED

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Type</th>
<th>Log2 Fold Change</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aldehyde dehydrogenase 1 family, member A2 (ALDH1A2)</td>
<td>Enzyme</td>
<td>-2.000</td>
<td>2.308</td>
</tr>
<tr>
<td>Pulmonary surfactant-associated protein A1 (SFTPA1)</td>
<td>Transporter</td>
<td>-2.000</td>
<td></td>
</tr>
<tr>
<td>Atorvastatin</td>
<td>Chemical drug</td>
<td>-2.019</td>
<td></td>
</tr>
<tr>
<td>Pirinixic acid</td>
<td>Chemical toxicant</td>
<td>-2.150</td>
<td></td>
</tr>
<tr>
<td>DNA (cytosine-5)-methyltransferase 3 beta (DNMT3B)</td>
<td>Enzyme</td>
<td>-2.236</td>
<td>-1.401</td>
</tr>
<tr>
<td>Recombination signal binding protein for immunoglobulin kappa J region (RBPJ)</td>
<td>Other</td>
<td>-2.359</td>
<td>0.420</td>
</tr>
<tr>
<td>TATA box binding protein (TBP)-associated factor (TAF4)</td>
<td>Transcription regulator</td>
<td>-2.376</td>
<td>-1.289</td>
</tr>
<tr>
<td>Protein kinase, cAMP-dependent, catalytic, alpha (PRKACA)</td>
<td>Kinase</td>
<td>-2.396</td>
<td>0.308</td>
</tr>
<tr>
<td>Clofibrate</td>
<td>Chemical drug</td>
<td>-2.401</td>
<td></td>
</tr>
<tr>
<td>Glucagon</td>
<td>Biologic drug</td>
<td>-2.433</td>
<td></td>
</tr>
<tr>
<td>Collagen, type XVIII, alpha 1 (COL18A1)</td>
<td>Other</td>
<td>-2.538</td>
<td>-0.373</td>
</tr>
<tr>
<td>Metribolone</td>
<td>Chemical reagent</td>
<td>-2.753</td>
<td></td>
</tr>
</tbody>
</table>

Within the upstream regulators identified, there were 3 growth factors and all were predicted to be active: EGF, HGF and JAG1. Each of these growth factors is predicted to interact with (by up or downregulating) a given number of dataset molecules. Figure 3.15 is a schematic representation of the dataset molecules and their regulation by each growth factor.
3.5.5 Expression of HGF and c-Met in the developing spine

Since HGF was one of the top growth factors positively regulating the genes in the dataset and microarray results suggested a higher differential expression of this gene in KRT18\textsuperscript{ive} than in KRT18\textsuperscript{ve} cells, its expression was analysed at the protein level in a cohort of foetal spines. HGF was found to be expressed by all notochordal cells in all developmental stages analysed (Fig 3.16 A, B2-I2). While in the 3.5 WPC specimen HGF was notochord-specific (Fig 3.16 A) it became co-
expressed by sclerotomal cells after that stage. In the 5.5 WPC HGF was co-expressed by notochordal and sclerotomal embryonic spine cells (Fig 3.16 B2 and C2). After 6-7 WPC, when sclerotomal cells had adopted their segmented pattern with the developing AF and VB being discernible, and until the later stage analysed (18 WPC), HGF was co-expressed by all notochordal cells (Fig 3.16 C2-I2) and by sclerotomal cells in the developing VB and in the inner AF (Fig 3.16 C2, D2 and E-I, panels 3 and 4); outer AF cells did not express HGF (Fig 3.16 C2, D2 and E-I, panel 4).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 3.16 HGF immunostaining of a cohort of developing spines. HGF was localised to all notochordal (arrows) and somite/sclerotomal cells in the developing VB; sclerotomal AF cells did not express HGF in the stages analysed. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M779 (7.5 WPC); E: M685 (10 WPC); F: M404 (11.5 WPC); G: M739 (12 WPC); H: M777 (17 WPC); I: M784 (18 WPC); J: M404 (11.5 WPC) rabbit isotype IgG control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP, panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
Since HGF interacts with cells through its receptor, c-Met, the expression of this marker was also analysed by immunohistochemistry. Similar to HGF, c-Met was expressed by all notochordal cells in all developmental stages analysed (Fig 3.17 A and B-I panels 2), being notochord-specific at 3.5 WPC (Fig 3.17 A). Between 5.5-7 WPC, it was expressed by all notochordal and sclerotomal (VB and AF anlagen) cells (Figure 3.17 B-C panels 2). Between 7-8 WPC and until the last stage analysed (17 WPC), c-Met was localised to the cell membrane of all notochordal cells (Figure 3.17 D-I panels 2) and of the sclerotomal cells in developing VB and in the inner AF (Figure 3.17 C-D panels 2 and E-I, panels 3 and 4); outer AF cells did not express c-Met (Fig 3.17 C-D panels 2 and E-I, panels 4).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO

251
Isolation and phenotypic characterisation of human notochordal cells
RICARDO RODRIGUES PINTO
Figure 3.17  c-Met immunostaining of a cohort of developing spines. c-Met was localised to all notochordal (arrows) and somite/sclerotomal cells in the developing VB and inner AF; sclerotomal cells in the outer AF cells did not express HGF in the stages analysed. A: M234 (3.5 WPC); B: M320 (5.5 WPC); C: M741 (7 WPC); D: M779 (7.5 WPC); E: M685 (10 WPC); F: M742 (12 WPC); G: M776 (13 WPC); H: M739 (14 WPC); I: M777 (17 WPC); J: M636 (9.5 WPC) rabbit isotype IgG control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or three higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP; panels 3 are centred to the developing sclerotomal AF and panels 4 are centred to the developing VB. No staining was seen when the antibody was substituted by an equal concentration of its isotype control (-ve control).
3.6 DISCUSSION

Notochordal cells or factors they produce have been proposed to have an anabolic role in the IVD and to protect against IVD degeneration; their loss with maturity in humans has been suggested to initiate the degenerative process (Hunter et al., 2003). As such, understanding their phenotype and the molecules that regulate them may provide a better understanding of the factors that regulate IVD homeostasis and aid in the development of new biological and cell-based strategies to address IVD degeneration.

While several attempts have been made to characterise the phenotype of these cells in animal models (Chen et al., 2006b, Kim et al., 2009, Minogue et al., 2010b, Oguz et al., 2007, Poirraudeau et al., 1999, Rastogi et al., 2009, Tang et al., 2012) results from these studies may not be directly translatable to human research as the NP cell (which derives from the notochordal cell) phenotype varies considerably between species (Minogue et al., 2010b, Minogue et al., 2010a, Rodrigues-Pinto et al., 2013). This may be due to differences in tissue and cell organisation, IVD and spine size and shape and biomechanics and posture in different mammals (Alini et al., 2008, Minogue et al., 2010b, Minogue et al., 2010a, Rodrigues-Pinto et al., 2013).

The implementation of studies in humans has, however, been hindered by ethical and logistical difficulties associated with the acquisition of a suitable human source of notochordal cells and with the technical difficulties associated with separating cells within mixed cell populations. Since this gap in knowledge is a major limitation in the field, the aims of this work were to develop an appropriate
method to isolate human notochordal cells from their adjacent sclerotomal spine cells and to then use microarrays to phenotype notochordal cells, identify notochordal cell surface markers and identify factors involved in the biological function of these cells.

Due to the small size and the lack of a clear demarcation between the foetal vertebrae and IVD, and particularly between the AF and the notochordal NP, a strategy was devised which involved the identification of a notochord-specific marker within the foetal spine (chapter 2) and subsequent labelling of foetal spine cells with that marker to allow separation of a pure notochordal cell population using FACS.

KRT8, KRT18 and KRT19 were found to be notochord-specific markers between the developmental stages of 3.5-18WPC. Since there were no differences in the staining intensity or pattern between the 3 KRTs and KRT18 has previously been identified as an adult NP marker (Minogue et al., 2010a), it was chosen to label notochordal cells. Being an intracellular marker, KRT18 required cells to be fixed and permeabilised prior to antibody labelling to allow for antibody penetration into the cell, a technique that, due to RNA degradation, is challenging.

To isolate KRT18^ve and KRT18^ve foetal spine cells from which to extract RNA with sufficient quality for microarrays, a fixation, permeabilisation, labelling and sorting method was developed. This was initially developed using MCF-7 cells and only then applied to foetal spine cells. To minimise RNA degradation, all procedures were performed in an RNAse-free environment and the total number of steps was reduced by minimising the number of washes and by using a directly conjugated antibody. Initially, the effect on the RNA quality of several alcohol-based (95% ethanol/ 5% acetic acid, 100% and 50% ethanol and UM-Fix®) or
non-alcoholic (100% RNAlater® and 50% RNAlater®) fixation and permeabilisation agents, which were selected or modified from previous studies (Barrett et al., 2002, Diez et al., 1999, Esser et al., 1995, Lamoreaux et al., 2006), was analysed. Two agents (95% ethanol/5% acetic acid and 100% RNAlater) were found to have the least detrimental effect and were used in the subsequent experiments. Next, the effect of intracellular labelling on the RNA quality using selected fixation and permeabilisation agents was analysed and was found to have little detrimental effect with either fixation and permeabilisation agents (although more detrimental with RNAlater®). Finally, the RNA quality after cell sorting was assessed and found to be significantly lower with RNAlater® than with 95% ethanol/ acetic acid.

It has recently been shown that the electric field radiation and mechanical stress cells are subjected to during FACS may cause cell disruption (Li et al., 2013) and thus it is possible that this would impact on the RNA quality. It is however, not known why 95% ethanol/ acetic acid does not have the same detrimental effect but this is likely due to the nature of the fixation/ permeabilisation method.

The methodology developed here (95%ethanol/acetic acid fixation and permeabilisation, followed by FITC-conjugated anti-KRT18 antibody labelling and cell sorting) provides a new strategy to isolate high quality RNA from limited numbers of cells labelled with an intracellular marker. This was the strategy used to phenotype human notochordal cells but is also a novel methodology whose application extends to other areas of research.

Using the optimised labelling protocol, human foetal notochordal and sclerotomal cells were isolated from one sample. To minimise the number of cell types to be sorted, the foetal spine was dissected from all its adjacent tissues. Hence, the
dissected spine contained only two populations of cells: notochordal (NP anlagens) and sclerotomal (AF and VB anlagens) cells that were separated and compared between each other.

As expected from the histological analysis of the foetal spine, the proportion of sclerotomal cells was much higher than that of the notochordal cells. Despite the limited number of cells in this sample, RNA was extracted from both cell types and, after amplification, it was in high quantity and quality. The accuracy of the developed methodology to separate two populations of cells was confirmed using qRT-PCR by the higher differential expression of KRT18, KRT19, GAL3 and CTGF (markers that have been shown to be notochord-specific at this developmental stage (9WPC), as detailed in chapter 2) in KRT18\textsuperscript{+ve} than in KRT18\textsuperscript{−ve} cells.

Microarray comparison between human foetal notochordal and sclerotomal cells identified a list of markers (positive and negative) with high differential expression between both cell types. Due to depletion of the samples used for the microarrays, validation of the differentially expressed genes identified in this microarray analysis could not be performed with these samples and will only be described at a later stage of this thesis (detailed in chapter 5).

The identified markers provide a comprehensive list of genes that, due to their high/low differential expression in notochordal cells may have pivotal roles in their biology. GRB14, a molecule that inhibits insulin growth factor (IGF) signalling via its receptor was the top differentially expressed KRT18\textsuperscript{+ve} marker. IGF1 has been to show to increase the proteoglycan production by bovine (Osada et al., 1996) and rat NP cells (Okuda et al., 2001), to stimulate the proliferation of bovine (Pratsinis and Kletsas, 2007) and human NP cells \textit{in vitro} (Zhang et al.,
and to inhibit IL-1-induced cell death in rabbit NP cells (Zhang et al., 2013). It is possible that GRB14 expression by notochordal cells at this developmental stage acts as a modulator of IGF signalling, controlling insulin production and preventing uncontrolled cell growth.

Interestingly, and confirming the importance of analysing human tissue, although some of the markers identified in this analysis have been linked to functions that may be relevant to their role in the notochordal tissue, neither of them had previously been associated with the notochord or with notochordal cells.

Another aim of this study was to identify notochordal cell surface markers. The identification of such markers would allow the isolation of notochordal cells without the need for cell fixation and permeabilisation, which would possibly allow for a more reproducible technique that could be used in a larger number of specimens to characterise the phenotype of notochordal cells. This study has identified 10 putative notochordal cell surface markers all with a differential expression of at least 16 fold between KRT18$^{+ve}$ and KRT18$^{-ve}$ cells. ADORA3 was the top differentially expressed cell surface marker. This G-coupled protein has recently been found to have higher differential expression in the AF from degenerated discs than in the AF of discs from adolescents with idiopathic scoliosis, having been proposed as biomarkers of IVD degeneration (Ye et al., 2015). However, and as aforementioned, the use of discs from patients with scoliosis as control samples, particularly for studying AF degeneration as in this study, should be criticised. AF cells from scoliotic discs have morphological abnormalities and exhibit cell necrosis (Sitte et al., 2013, Chen et al., 2005) that should preclude from classifying them as healthy cells – the differential expression of ADORA3 between the cells in the study may therefore be related to
the pathogenesis of the two different clinical conditions (disc degeneration and scoliosis). In fact, ADORA3 has been shown to have cardioprotective (Bouma et al., 1997) and neuroprotective (Chen et al., 2006a) functions. It is possible that the role of ADORA3 in notochordal cells is related to its anti-inflammatory properties, such as has been shown in inflammatory bowel disease (Odashima et al., 2005). Notochordal cells have been shown to protect NP cells from pro-inflammatory cytokine secretion induced by IL-1ß and FasL (Erwin et al., 2011). Interestingly, CD69, another regulator of inflammation which has recently been proposed as a novel therapeutic agent for inflammatory bowel disease (Radulovic and Niess, 2015), was also identified as a top cell surface notochordal cell marker. Future studies investigating the role of these markers may provide more clues on notochordal cell biology.

Besides highlighting potential functions of notochordal cells, the list of differentially expressed genes provides a phenotypic signature for such cells. This phenotypic signature can potentially be used to identify notochord-derived cells in the adult disc. Furthermore, and for regenerative medicine purposes, this phenotypic profile will enable assessment of accurate differentiation of progenitor cells to notochordal cells.

Another benefit of microarrays is that they allow identification of upstream regulators that can explain the observed gene expression differences seen in the dataset. This can potentially highlight the biological activities occurring in the cells being studied. Growth factor regulation of notochordal cells is of particular interest as those growth factors can potentially be used to induce some of the notochordal cell functions. For this reason, the list of upstream regulators was filtered to identify growth factors and EGF1, HGF and JAG1 were identified as
regulating differentially expressed dataset genes. This indicates that notochordal cells have receptors that interact with these growth factors and that from this interaction a number of dataset molecules are up or down-regulated.

HGF is of particular relevance as IPA predicted it to up-regulate the expression of GRB14 (the top differentially expressed gene) and also FoxA2 and NRP1, genes that are relevant to intervertebral disc development and function. FoxA2 is a gene that, together with FoxA1 is crucial for the development of the nucleus pulposus (Maier et al., 2013). NRP-1 is the receptor for semaphorin 3A (Sema3A), which is involved in neuronal repulsion from the notochord (Masuda et al., 2004) and Sema3A has been proposed to act as a barrier to neuronal ingrowth into the healthy IVD (Tolofari et al., 2010). The potential for HGF to interact with notochordal cells (through its receptor c-Met) and their presence in the developing notochord was confirmed by immunohistochemistry. These findings suggest an important role for HGF as a notochordal cell regulator.
3.7 IMPLICATIONS OF THIS INVESTIGATION

This is, to the author’s knowledge, the first study to develop a methodology to allow for the isolation of notochordal cells from surrounding sclerotomal tissues of human developing spines and to characterise the phenotype of these cells and their potential regulatory networks and pathways. To do so, a methodology to extract high-quality RNA from low numbers of fixed, permeabilised, labelled and sorted cells was devised, a strategy that will have applications beyond the IVD field.

The genes identified here, albeit requiring further validation, are a set of positive and negative notochordal markers that represent the human notochordal genetic signature and can now be used to characterise notochordal cell development but also to understand the fate of these cells in the maturing NP. Furthermore, such markers can also be used to assess end-stage differentiation of progenitor cells towards notochordal cells.

Importantly, HGF was identified as an upstream regulator of notochordal cell genes and may alone, or in combination with other growth factors, drive progenitor cell differentiation to notochordal cells.
CHAPTER 4

*Isolation and characterisation of human notochordal cells using the specific cell surface marker CD24*
4.1 INTRODUCTION

To enable the development of biological or cell-based therapies for disc degeneration it is essential to characterise the molecular phenotype of the target cell. As discussed in previous chapters, and as notochordal cells have been proposed to play a fundamental role in IVD homeostasis, it is essential to characterise their molecular phenotype. Importantly since, in humans, notochordal cells are restricted to the embryonic, foetal and juvenile stages of IVD development, isolation and characterisation of such cells is needed.

In chapter 2, KRT8, KRT18 and KRT19 were identified as notochord-specific markers in all developmental stages analysed (3.5-18 WPC) and CD24 between 5.5-18WPC and, in chapter 3, KRT18 was used to label and separate notochordal from sclerotomal cells. However, due to the need for fixation and permeabilisation prior to cell labelling with an intracellular marker, the isolation of good quality RNA from the few cells obtained was technically challenging. While this task was pursued and eventually accomplished with a novel methodology being described (chapter 3), an alternative method was simultaneously pursued.

CD24 was identified as a cell surface notochord-specific marker in the developing spine between 5.5-18WPC (chapter 2) that could potentially be used to isolate notochordal cells during these stages. Importantly, being localised to the outer cell membrane, this marker would offer two important advantages over KRT18: 1) cells wouldn’t need to be fixed and permeabilised prior to labelling, minimising the limitations associated with RNA quality obtained using this method; 2) by
labelling with a cell surface marker, cells would stay viable and therefore the same methodology could be used to isolate and culture notochordal cells. This would permit not only the characterisation of the notochordal cell phenotype but would also offer the potential to subsequently culture these cells and study their biological function.

In addition to providing an essential phenotype for human notochordal cells, the information from phenotyping studies can also be used to understand the networks, canonical pathways, functions and upstream regulators of the genes expressed by those cells. This information would allow understanding of what drives notochordal cell differentiation (IPA Upstream Regulator and Master Regulator Analyses), the most relevant pathways and molecular interactions within or between notochordal cells and/ or adjacent cells within the developing spine (IPA Canonical Pathways and IPA Networks) and the functions and diseases downstream of notochordal cells (IPA Diseases and Functions). This knowledge is important as it will improve our understanding of how notochordal cells exert their putative anabolic and protective role in the IVD (Aguiar et al., 1999, Cappello et al., 2006, Cornejo et al., 2015, Erwin and Inman, 2006, Erwin et al., 2011, Erwin et al., 2009, Miyazaki et al., 2009, Purmessur et al., 2015, Purmessur et al., 2011), how to drive stem cell differentiation to notochordal cells or NP cell progenitors and whether there are molecules produced by notochordal cells that are responsible for maintaining IVD homeostasis.
4.2 HYPOTHESES AND AIMS

The hypotheses for this study were that:

1. CD24 is a unique notochordal cell surface marker during the developmental stages of 5.5-18 WPC that can be used to identify and isolate viable notochordal from sclerotomal cells in the developing human spine.

2. Isolation of human notochordal cells will allow a characterisation of their phenotype and molecular pathways, networks and upstream regulators; this knowledge may allow a better understanding of their function and role in development of the IVD and in protecting the IVD from degeneration.

The aims for this study were to:

1. Develop methodologies to allow for the isolation of viable notochordal cells from surrounding sclerotomal tissues of the human developing spine using the notochordal unique expression of CD24.

2. Characterise the human notochordal cell phenotype using microarrays.

3. Identify the molecular factors pathways, networks, upstream regulators and master regulators of notochordal cells.
4.3 SUMMARY OF THE EXPERIMENTAL DESIGN

To isolate notochordal cells from developing spinal tissues, a methodology similar to that described in Figure 3.1 (chapter 3) was used but, here, using CD24 as a marker. However, since CD24 is cell surface marker, cells did not need to be fixed and permeabilised as they were for isolation utilising KRT18.

In chapter 3 enzymatic methods were used to obtain cells from spinal tissues. The use of collagenase for tissue digestion, however, may have an effect on cell surface marker expression (Abuzakouk et al., 1996). Accumax®, a commercially available product containing proteolytic and collagenolytic enzymes is an alternative to collagenase digestion. This reagent is stated to replace all applications of trypsin and collagenase in tissue dissociation and its usefulness in obtaining single cells from brain tissue for FACS analysis and cell sorting has previously been demonstrated (Buehler et al., 2012).

To identify the ideal method to isolate CD24⁺ve and CD24 negative (CD24⁻ve) cells from developing spinal tissues with minimal cell surface marker and viability impact, Accumax® and collagenase digestions were assessed. Due to limitations in acquiring sufficient human samples for methodology development, MCF-7 cells (CD24⁺ve (Kim et al., 2007)) were used. To mimic the 3D collagenous microenvironment, in which IVD cells reside, MCF-7 cells were encapsulated in type I collagen gels.

After identification of the ideal digestion method, cells from human developing spines were obtained and notochordal (CD24⁺ve) and sclerotomal (CD24⁻ve) cells
were separated. RNA from both cell types was extracted and the differential gene expression analysed using microarrays to identify the notochordal cell phenotype.

Finally, differentially expressed genes were subjected to IPA analysis to identify the networks, pathways, functions and regulators of human notochordal cells.
4.4 MATERIALS AND METHODS

4.4.1 Cell labelling optimisation

All experiments were performed in triplicate and all reagents and instruments used were certified RNase/DNase-free.

MCF-7 cells were acquired as described in 3.4.1, thawed as described in 3.4.2, expanded as described in 3.4.3 and counted as described in 3.4.4.

For CD24 labelling optimisation, 1x10^6 MCF-7 cells were washed in cold PBS, centrifuged (400G, 5 minutes, 4°C) and re-suspended in 150µL of PE-conjugated-anti-CD24 antibody (Beckman Coulter®, PN IM1428U). To identify the optimal antibody concentration, the following dilutions in FACS buffer (0.5% BSA in 2mM EDTA in PBS) were tested: 1/10 (1µg/mL), 1/50 (0.2µg/mL) and 1/100 (0.1µg/mL). Cells were incubated in antibody solution in the dark for 10 minutes at 4°C. After cell labelling, 800µL of FACS buffer were added to the cell suspension, mixed by pipetting and centrifuged at 400G for 5 minutes at 4°C. Finally, the supernatant was aspirated and cells were re-suspended in 300µL of cold PBS for flow cytometry analysis. One million MCF-7 cells labelled with equivalent concentrations of PE-conjugated-IgG1 isotype control (R&D Systems®, IC002A) were used as negative controls as described in 3.4.6.1. Cell aggregates and cell debris were gated out as described in 3.4.6.2. CD24^+ve and CD24^-ve gates were drawn using IgG1 isotype control labelling, as described for KRT18 in 3.4.6.2. The 488nm laser was used for excitation and the emission detected using the 575/25nm filter.
Draq7® (0.3mM, Biostatus®, DR71000) was used to assess cell viability. This is a far-red fluorescent DNA dye that only stains the nuclei in dead and permeabilised cells. For Draq7® labelling optimisation, 1x10⁶ MCF-7 cells were washed in FACS buffer, centrifuged (400G, 5 minutes, 4ºC) and re-suspended in 150µL of Draq7® in FACS buffer. To identify the optimal Draq7® concentration, the following dilutions in FACS buffer were tested (1/100, 1/500, 1/1000). Cells in Draq7® were incubated in the dark for 10 minutes at 4ºC. After cell labelling, 800µL of FACS buffer were added to the cell suspension, mixed by pipetting and centrifuged at 400G for 5 minutes at 4ºC. Finally, the supernatant was aspirated and cells were re-suspended in 300µL of cold PBS for flow cytometry analysis. An optimal Draq7® concentration was pursued that would allow discrimination between labelled and unlabelled cells. Cell aggregates and cell debris were gated as described in 3.4.6.2. The 633nm laser was used for excitation and the emission detected using the 660/20nm filter.

4.4.2 Development of a methodology to isolate viable CD24⁺ve cells from collagenous tissues

To encapsulate MCF-7 cells in type I collagen gels, 4 x 10⁶/ml cells in MCF-7 expansion medium were centrifuged at 400G at room temperature for 5 minutes, the supernatant was aspirated and the cells re-suspended in 100µL of MCF-7 cell expansion medium. Then, 800µL of atelosoluble purified bovine type I collagen (2.9mg/ml, Devro Medical®, 10DM001004523) and 100µL of neutralisation buffer (0.2M sodium phosphate dibasic anhydrous (Sigma Aldrich®, 127H0168),
1.3M sodium chloride (Lab Suppliers®, K32104233323, pH 11.2) were added to the cell suspension. The solution was vortexed to distribute the cells evenly and then pipetted into 0.4µm high density translucent membrane cell culture inserts (BD Biosciences®, 353493), within a 24 well plate (BD Biosciences®, 353504), at a volume of 100µl per insert. For gelation, 750µL of MCF-7 cell medium was added to the well of the plate and 500µL of medium to the top of the hydrogel and plates were incubated for 45 minutes at 37°C. This methodology has previously been described in the host laboratory where this work was conducted (Clarke et al., 2014).

MCF-7 cells encapsulated in collagen gels were incubated in an orbital shaker with either 0.1% type II collagenase solution in DMEM (Sigma Aldrich®, D0819) or Accumax® for 1h, 2h and 16h. After incubation, cell solution was centrifuged at 400G for 5 minutes at 4°C, re-suspended in 200µL of Draq7® and PE-conjugated-anti-CD24 antibody and incubated in the dark for 10 minutes at 4°C. After incubation, 800µL of FACS buffer was added, mixed by pipetting and centrifuged at 400G for 5 minutes at 4°C. Supernatant was aspirated and discarded and cells were re-suspended in 300µL of cold PBS for analysis. Cell viability and CD24 expression were analysed by flow cytometry. All experiments were performed in triplicate (Figure 4.1).
Figure 4.1 Schematic illustration of the methodology used to assess the ideal method to obtain viable CD24^ve cells. MCF-7 cells in suspension were encapsulated in collagen gels and incubated for 1, 2 or 16 hours in the presence of Accumax® or 0.1% type II collagenase in medium. After incubation, cell viability and CD24 expression were analysed in a flow cytometer.

4.4.3 RNA extraction from viable CD24^ve and CD24^ve foetal spine cells

Embryonic and foetal samples were acquired and staged as described in 2.3.1 and dissected to obtain whole spines containing IVD and VB as detailed in 2.3.2. Table 4.1 details the samples used in the CD24 microarray analysis.
Table 4.1 Identification and age of the samples used in the CD24 microarray analysis.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Stage (WPC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M728</td>
<td>7.5</td>
</tr>
<tr>
<td>M708</td>
<td>8</td>
</tr>
<tr>
<td>M753</td>
<td>8.5</td>
</tr>
<tr>
<td>M729</td>
<td>12</td>
</tr>
<tr>
<td>M739</td>
<td>14</td>
</tr>
</tbody>
</table>

Immediately after dissection, spines were cut into small fragments and digested for 2 hours in medium (alpha-MEM (Sigma-Aldrich®, M4526)) containing 1% (v/v) antibiotic/antimycotic solution and 0.1% (w/v) type II collagenase (Gibco®, 17101-015) in an orbital shaker at 37°C. After digestion, the cell solution was re-suspended in non-enzymatic cell dissociation solution for 10 minutes at 37°C. This incubation was reduced in comparison with that used for KRT18 isolation as it was found to be sufficient to dissociate notochordal cell clusters. After cell dissociation, cells were washed in FACS buffer and sieved through a 40 µm cell filter.

Cells were then incubated for 10 minutes in a solution containing 0.3 µM Draq7® and 0.5 µg/mL of PE-conjugated anti-CD24 antibody in FACS buffer in the dark, for 10 minutes, at 4°C. After incubation, 800 µL of FACS buffer was added to each tube, mixed well by pipetting, centrifuged (400G, 5 minutes, 4°C) and cells were re-suspended in 300 µL of cold PBS for cell sorting. A small aliquot of each sample was stained with anti-IgG1 isotype control antibody (0.2 µg/mL IgG1, BD Pharminogen®, 550617) and Draq7® viability dye (0.3 µM) and used as negative control and to draw viable CD24^ve and CD24^-ve gates. After gating non-viable cells (Draq7 positive (Draq7^ve)) (Figure 4.2 A), cell aggregates (Figure 4.2 B) and cell debris (Figure 4.2 C) out, viable (Draq7 negative (Draq7^-ve)) CD24^ve
and CD24\(^{\text{ve}}\) cells were identified and collected separately. CD24 positivity and negativity were determined by analysing a sample stained with PE-isotype control (Figure 4.2. D). To ensure accuracy in gating positive and negative events, only those with strongest PE labelling were considered as CD24\(^{\text{ve}}\) and those with weakest labelling were considered CD24\(^{\text{ve}}\), the intermediate population of cells was discarded (Figures 4.2. D and E). Cells were sorted into MCTs containing 350\(\mu\)L of RLT lysis buffer and used immediately for RNA extraction.
Figure 4.2 Example of the gating to isolate viable CD24^{+ve} and CD24^{−ve} cells. (A) Drag7^{−ve} events (P1) were selected to exclude non-viable cells; (B) Single events were selected (P2) by excluding those with higher forward scatter (area and height), which represent cell aggregates; (C) A third gate (P3) was drawn that excluded events with small size (FSC-H) and granularity (SSC-H), indicative of cell debris. (D) A sample stained with PE-conjugated isotype control antibody was initially analysed and positive (P4) and negative (P5) gates were drawn. (E) The same gates (P4 and P5) were applied to a sample labelled with PE-conjugated anti-CD24 antibody to identify CD24^{+ve} (P4) and CD24^{−ve} (P5) events; to avoid contamination, the intermediate population of events was excluded.

RNA was extracted as detailed in 3.4.6.3 and its quantity, quality, purity and integrity assessed as described in 3.4.6.4. RNA was amplified to cDNA and purified as described in 3.4.7.2.
4.4.4 qRT-PCR

qRT-PCR using gene specific primers was performed using the TaqMan method as described in 3.4.7.3. All reactions were performed in triplicate. The expression of each gene was normalised to the housekeeping gene GAPDH and data was analysed using the previously described $2^{\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), in which the expression of CD24$^{\text{+ve}}$ cells was normalised to the expression of CD24$^{\text{-ve}}$ cells. qRT-PCR primers and probes for GAPDH are detailed in Table 3.6 and those for CD24 and T are detailed on Table 4.2.

Table 4.2 Human oligonucleotide primers and probes used for qRT-PCR analysis.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>NCBI RefSeq</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Probe sequence</th>
<th>Optimal primer concentration (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD24</td>
<td>NM_001291737</td>
<td>GCTCCTACCACGCAGATTAT</td>
<td>CCTGTTGGTGGTGCTTTAGTTG</td>
<td>CCAGTGAACACAACACAC</td>
<td>900</td>
</tr>
<tr>
<td>Brachyury (T)</td>
<td>NM_080646.1</td>
<td>CTCAACCAACTGCTCAATTTAT</td>
<td>CTGTGAATCTCTCTTTGTTCTCTT</td>
<td>CCAGATCTGCGCTGAACCTTGTGC</td>
<td>900</td>
</tr>
</tbody>
</table>

4.4.5 cDNA microarrays

Only high quality RNA (RNA integrity number $>7$) was used for microarrays. RNA was fragmented and labelled using the Encore Biotin Module (NuGen Technologies®), as described in 3.4.8.
Four micrograms of SPIA® cDNA from each sample was hybridized to the GeneChip human genome U133 Plus 2.0 array (Affymetrix®, 900466); this was performed by staff within the Genomic Technologies Facility (Faculty of Life Sciences, The University of Manchester), according to manufacturer’s instructions. Technical quality control was performed with dChip software (Li and Wong, 2001); an array was considered to be an outlier when more than 5% of the probe sets for that array were judged as outliers by the dChip outlier detection algorithm. Background correction, quantile normalization, and gene expression analysis were performed using the robust multiarray average (RMA) analysis in the Bioconductor software package (Bolstad et al., 2003). Differential expression analysis was performed using routine analytical methods (Smyth, 2004). Lists of differentially expressed genes were controlled for false discovery rate (FDR) errors using the QValue method (Smyth, 2004).

4.4.6 Microarray analysis

Differentially expressed genes were defined as those with an FDR-corrected $p$ value (called the q value) of $\leq 0.05$ and a minimum normalized expression level of $\geq 2$ fold.

Principal component analysis (PCA) and hierarchical clustering were analysed using Partek Genomics Suite (Partek Inc.®, St. Louis, MO, USA).

IPA was used to identify networks, biological functions, pathways, upstream regulators and master regulators of differentially expressed genes. $p$-values were
calculated using the right-tailed Fisher Exact Test; Benjamini-Hochberg test was used for p-value multiple testing correction.

IPA Network Analysis identifies regulatory network between genes in a dataset to help understand how those genes are biologically related. Networks represent diagrams of protein-protein (gene-product) interactions. The higher the number of focus genes and the number of connections between themselves in each network, the higher will be its rank.

IPA Functions analyses which biological functions and disease processes are most relevant to the dataset genes. It identifies biological functions expected to be increased or decreased given the gene expression changes in the dataset. If the changes seen in the dataset are consistent with those expected from the literature, then a prediction is made about that activation state (z score). If the z score is <0 the function is predicted to be decreased; if the score is >0, the function is predicted to be increased (Table 4.3). However, the larger the absolute number is, the less likely it would be obtained by chance and, therefore, z-scores greater than 2 or smaller than -2 were considered significant. p-values were calculated to measure the likelihood of the association between the dataset genes and a function being due to random association.
Table 4.3 IPA biological function prediction. The gene state in the dataset is compared with the gene-function interaction (according to published literature, curated in the IPA database) to predict the activation state of that function or disease state downstream of the dataset genes.

<table>
<thead>
<tr>
<th>Gene state in the dataset</th>
<th>Gene-function effect (literature)</th>
<th>Predicted activation state of biological function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upregulated</td>
<td>Activates</td>
<td>Increased</td>
</tr>
<tr>
<td>Upregulated</td>
<td>Inhibits</td>
<td>Decreased</td>
</tr>
<tr>
<td>Downregulated</td>
<td>Activates</td>
<td>Decreased</td>
</tr>
<tr>
<td>Downregulated</td>
<td>Inhibits</td>
<td>Increased</td>
</tr>
</tbody>
</table>

Canonical Pathway Analysis investigates which cell signalling pathways are more relevant to the genes in the dataset. Canonical pathways are well-characterised pathways that have been curated and stored in the Ingenuity Knowledge Base®. IPA will calculates a ratio between the number of database genes that belong to a given pathway and the total number of genes in that pathway. This allows inference about the possible biological and molecular roles of the genes in the dataset. A corrected p-value is also calculated to infer the strength of the association and the likelihood of it occurring due to random chance.

Upstream regulators were analysed as described in 3.4.9. The list of regulators was filtered to identify those with a predictable activation state and that were either cytokines or growth factors.

Master regulators are molecules that can affect the expression of regulators of the molecules in the dataset. Similar to upstream regulators, master regulators can be predicted to be active (positive master regulators) or to be inhibited (negative master regulators) and this information is given by the z score. The list of regulators was filtered to identify those with a predictable activation state (z-
scores greater than 2 or smaller than -2) and that were either cytokines or growth factors.
4.5 RESULTS

4.5.1 Optimisation of Draq7® and CD24 labelling

The optimal PE-conjugated anti-CD24 antibody concentration was empirically determined as 0.2µg/mL. Lower and higher antibody concentrations showed a second, less intense PE-CD24 peak, that overlapped with PE-isotype labelling, indicative of a suboptimal discrimination between positive and negative cells (Figure 4.3).

Figure 4.3 Flow cytometry overlay histograms depicting MCF-7 cells labelled with different concentrations of PE-conjugated anti-CD24 antibody (blue) and PE-conjugated anti-IgG1 isotype control (red) at different concentrations: (A) 1µg/mL, (B) 0.2µg/mL and (C) 0.1µg/mL.

A Draq7® dilution of 1/100 was empirically found to allow for discrimination between viable and non-viable cells. Non-viable cells were Draq7"ve and had lower forward scatter intensity values (Figure 4.4).
Figure 4.4 Histograms (A) and scatter plots (B) used for Draq7® optimisation at 1/100 (A1 and B1), 1/500 (A2 and B2) and 1/1000 (A3 and B3) dilutions. In A1 two peaks of Draq7 fluorescence labelling (detected with the 660/20 filter) were seen, corresponding to viable (lower fluorescence intensity) and non-viable (higher fluorescence intensity) events, whereas in A2 and A3, Draq7® positive staining was not detected, indicating that Draq7 concentration was not sufficiently high to penetrate into dead cells. In plots B1-B3, Draq7 fluorescence was plotted against forward scatter, showing that Draq7⁻ve cells (R1) were also those with higher forward scatter intensity (a surrogate for cell size), whereas dead cells had lower forward scatter and higher Draq7 labelling.

4.5.2 Methodology to isolate viable CD24⁺ve cells from collagenous tissues

To identify the optimal method to extract cells from collagenous tissues with minimal effect on CD24 expression and cell viability, MCF-7 cells encapsulated in collagen gels were incubated for 1, 2 and 16 hours in either 0.1% type II collagenase in DMEM or Accumax® and CD24 and Draq7 expression analysed by flow cytometry.
CD24 labelling was not affected by Accumax® incubation with a good discrimination between CD24<sup>+</sup> and CD24<sup>-</sup> cells; this was independent of the time spent in Accumax® (Figure 4.5 A1-A3). Cell viability, however, was affected by Accumax® incubation, with the proportion of viable cells decreasing from 1 to 2 and to 16 hours of incubation (Figure 4.5 B1-B3 and C1-C3 and Table 4.4).

CD24 labelling was not affected by incubation in collagenase for 1 or 2 hours, but after 16 hours of incubation, there was an overlap between CD24<sup>+</sup> and CD24<sup>-</sup> cells (Figure 4.6 A1-A3). Cell viability was affected by collagenase incubation, with the proportion of viable cells decreasing with the length of time spent in collagenase (Figure 4.6 B1-B3 and C1-C3 and Table 4.4).
Figure 4.5 MCF-7 cells were encapsulated in collagen gels, incubated in Accumax® for different time periods and then labelled with Draq7® and PE-conjugated-anti-CD24 antibody or Draq7® and PE-IgG isotype control. (A) Overlay histograms showing good discrimination between CD24⁺ve cells (MCF-7 cells labelled with PE-CD24 (blue)) and CD24⁻ve cells (PE-IgG isotype control (red)) incubated with Accumax® for one (A1), two (A2) and sixteen (A3) hours. (B) Histograms and (C) Scatter plots showing viable (V) and non-viable (NV) cells (Draq7 staining) after incubation with Accumax® for one (B1 and C1), two (B2 and C2) and sixteen (B3 and C3) hours.
Figure 4.6 MCF-7 cells were encapsulated in collagen gels, incubated in type II collagenase for different time periods and then labelled with Draq7® and PE-conjugated-anti-CD24 antibody or Draq7® and PE-IgG isotype control. (A) Overlay histograms showing good discrimination between CD24+ve cells (MCF-7 cells labelled with PE-CD24 (blue)) and CD24-ve cells (PE-IgG isotype control (red)) incubated with type II collagenase for one (A1) and two (A2) but not for sixteen (A3) hours. (B) Histograms and (C) Scatter plots showing viable (V) and non-viable (NV) cells (Draq7 staining) after incubation with type II collagenase for one (B1 and C1), two (B2 and C2) and sixteen (B3 and C3) hours.

Table 4.4 Proportion of viable cells (Draq7+ve) after incubation for 1, 2 or 16 hours in Accumax® or type II collagenase. Values represent the average of n=3.

<table>
<thead>
<tr>
<th></th>
<th>Proportion of viable cells (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-hour incubation</td>
</tr>
<tr>
<td>Accumax®</td>
<td>57.18</td>
</tr>
<tr>
<td>Type II collagenase</td>
<td>66.46</td>
</tr>
</tbody>
</table>
Since cell viability was highest for collagenase digestion for 1 or 2 hours (61-66%) (Table 4.4) and CD24 expression was not affected in the 1 or 2 hour-incubation periods, a 2 hour-incubation in type II collagenase was used to digest spinal tissues. The 2-hour incubation was chosen in detriment of the 1-hour because, theoretically, higher total number of cells would be obtained after 2 hours of digestion.

### 4.5.3 Separation of CD24$^{+ve}$ and CD24$^{-ve}$ foetal spine cells and qRT-PCR validation of cell separation

A small population (5.0-19.5%) of CD24$^{+ve}$ cells was identified within a larger population (42.1-89.9%) of CD24$^{-ve}$ cells in all samples (Table 4.5 and Figure 4.7). No differences were detected in the forward and side scatter distribution of CD24$^{+ve}$ and CD24$^{-ve}$ events within each sample.

<table>
<thead>
<tr>
<th>ID</th>
<th>Separation</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>M728</td>
<td>CD24$^{+ve}$</td>
<td>19.5%</td>
</tr>
<tr>
<td></td>
<td>CD24$^{-ve}$</td>
<td>55.5%</td>
</tr>
<tr>
<td>M708</td>
<td>CD24$^{+ve}$</td>
<td>5.5%</td>
</tr>
<tr>
<td></td>
<td>CD24$^{-ve}$</td>
<td>42.1%</td>
</tr>
<tr>
<td>M753</td>
<td>CD24$^{+ve}$</td>
<td>5.0%</td>
</tr>
<tr>
<td></td>
<td>CD24$^{-ve}$</td>
<td>89.9%</td>
</tr>
<tr>
<td>M729</td>
<td>CD24$^{+ve}$</td>
<td>11.5%</td>
</tr>
<tr>
<td></td>
<td>CD24$^{-ve}$</td>
<td>73.6%</td>
</tr>
<tr>
<td>M739</td>
<td>CD24$^{+ve}$</td>
<td>10.5%</td>
</tr>
<tr>
<td></td>
<td>CD24$^{-ve}$</td>
<td>43.3%</td>
</tr>
</tbody>
</table>
Figure 4.7 FACS plots of the samples used in the microarray analysis. (A) M728 (7.5 weeks), (B) M708 (8 weeks), (C) M753 (8.5 weeks), (D) M729 (12 weeks), (E) M739 (14 weeks). Gates were drawn as described in figure 4.2. P4 represents viable CD24+ve events and P5 represents viable CD24-ve events.

To validate the accuracy of separation between notochordal (CD24+ve) and sclerotomal (CD24-ve) cells the expression of CD24 and also of the transcription factor T, a transcription factor involved in the morphogenesis of the notochord, was analysed using qRT-PCR. CD24 and T had a significantly higher differential expression in CD24+ve than in CD24-ve cells, suggesting a good separation between CD24+ve and CD24-ve cells (Figure 4.8).
4.5.4 cDNA microarrays

4.5.4.1 Microarray quality control

Three CD24^ve samples failed to pass microarray quality control and were therefore excluded from the analysis. For this reason microarray analysis was performed by comparing the expression of CD24^ve cells from 5 specimens with that of CD24^-ve cells from 2 specimens (Table 4.6).
Table 4.6 dchip analysis of samples used in the microarray analysis. Samples marked with an asterisk (*) were considered as outliers and eliminated from analysis.

<table>
<thead>
<tr>
<th>Array</th>
<th>% array outlier</th>
</tr>
</thead>
<tbody>
<tr>
<td>M728 CD24+ve</td>
<td>1.224</td>
</tr>
<tr>
<td>M728 CD24-ve</td>
<td>20.433*</td>
</tr>
<tr>
<td>M708 CD24+ve</td>
<td>0.26</td>
</tr>
<tr>
<td>M708 CD24-ve</td>
<td>29.866*</td>
</tr>
<tr>
<td>M753 CD24+ve</td>
<td>1.712</td>
</tr>
<tr>
<td>M753 CD24-ve</td>
<td>0.81</td>
</tr>
<tr>
<td>M729 CD24+ve</td>
<td>1.783</td>
</tr>
<tr>
<td>M729 CD24-ve</td>
<td>17.344*</td>
</tr>
<tr>
<td>M739 CD24+ve</td>
<td>0.487</td>
</tr>
<tr>
<td>M739 CD24-ve</td>
<td>0.132</td>
</tr>
</tbody>
</table>

4.5.4.2 Hierarchical clustering and principal clustering analysis

Hierarchical clustering showed segregation between CD24+ve and CD24-ve genes, with a similar gene expression profile between the cells from the 5 CD24+ve samples, which differed from that of the cells from the 2 CD24-ve samples (Figure 4.9). PCA of microarray data showed that the genes from CD24+ve cells clustered together and away from those of CD24-ve cells (Figure 4.10). Together, these results indicate that the gene expression profile of CD24+ve cells is distinct from that of CD24-ve cells, suggesting a good separation between cell types.
Figure 4.9 Hierarchical clustering of the genes in the 5 CD24<sup>+</sup> and 2 CD24<sup>-</sup> samples analysed. Each row represents one sample and each line represents one gene. Boxes in red indicate upregulation and boxes in green indicate downregulation. The dendogram identifies two main groups corresponding to CD24<sup>+</sup> and CD24<sup>-</sup> samples.
Figure 4.10 PCA of the genes expressed by CD24^{+} (dark blue) and CD24^{-} (light blue) samples. As CD24^{+} samples clustered in close proximity to each other and distant to CD24^{-} samples, this indicated a clear separation between two different cell types with a gene expression profile that differed between them.

4.5.4.3 Identification of CD24^{+} and CD24^{-} markers

Array analysis identified 884 upregulated (log ratio >2, p<0.05) and 1460 downregulated (log ratio <2, p<0.05) genes in CD24^{+} versus CD24^{-} cells (Figure 4.11).
Figure 4.11 Volcano plot depicting the differential gene expression signals between CD24<sup>ve</sup> and CD24<sup>ve</sup> cells. Only genes with a fold change < -2 or >2 and with a p-value < 0.05 were considered to be differentially expressed and were used for further analysis.

Top notochordal markers were CD24, STMN2, RTN1, PRPH, CXCL12, IGF1, MAP1B, ISL1, CLDN1 and THBS2 (>19 fold change differential expression, p<0.05) and top sclerotomal markers were WISP3, CHST11, SERPINA3 and CHAD (>21 fold change differential expression, p<0.05) (Table 4.7).
Table 4.7 Top (upregulated) and bottom (downregulated) differentially expressed genes between CD24^{+ve} and CD24^{−ve} cells.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Gene name</th>
<th>Fold Change</th>
<th>Log2 Ratio</th>
<th>p-value</th>
<th>Function of the encoded protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD24</td>
<td>CD24 molecule</td>
<td>68.102</td>
<td>6.090</td>
<td>1.29x10^{-06}</td>
<td>Sialoglycoprotein expressed on the surface of mature granulocytes and in B cells. Involved in axon guidance, B cell receptor transport into membrane raft, cell-cell adhesion and migration. Regulates TGFβ3 production, epithelial cell differentiation, MAPK cascade, response to hypoxia and Wnt signalling pathway.</td>
</tr>
<tr>
<td>STMN2</td>
<td>Stathmin-like 2</td>
<td>48.064</td>
<td>5.587</td>
<td>2.34x10^{-02}</td>
<td>Protein involved in microtubule dynamics, neuronal growth and osteogenesis (Liu et al., 2002, Okazaki et al., 1995, Chiellini et al., 2008).</td>
</tr>
<tr>
<td>RTN1</td>
<td>Reticulon 1</td>
<td>45.294</td>
<td>5.501</td>
<td>2.25x10^{-03}</td>
<td>Belongs to a family of proteins (reticulons) involved in neuroendocrine secretion or in membrane trafficking in neuroendocrine cells (Nepravishta et al., 2012).</td>
</tr>
<tr>
<td>PRPH</td>
<td>Peripherin</td>
<td>45.207</td>
<td>5.498</td>
<td>2.62x10^{-04}</td>
<td>Cytoskeletal protein found in neurons of the peripheral nervous system (Eriksson et al., 2008).</td>
</tr>
<tr>
<td>CXCL12</td>
<td>Chemokine (C-X-C motif) ligand 12</td>
<td>33.187</td>
<td>5.053</td>
<td>1.51x10^{-03}</td>
<td>Ligand for the G-protein coupled receptor involved in embryogenesis, immune surveillance, inflammation response, tissue homeostasis, and tumour growth and metastasis (Kryczek et al., 2007, Muller et al., 2001).</td>
</tr>
<tr>
<td>IGF1</td>
<td>Insulin-like growth factor 1 (somatomedin C)</td>
<td>28.073</td>
<td>4.811</td>
<td>4.65x10^{-04}</td>
<td>Protein with functional and structural similarity to insulin. It is involved in blood vessel remodelling, bone mineralization and maturation, epithelial cell proliferation, regulation of glycolysis and osteoblast differentiation (Junnila et al., 2013).</td>
</tr>
<tr>
<td>MAP1B</td>
<td>Microtubule-associated protein 1B</td>
<td>22.298</td>
<td>4.479</td>
<td>4.32x10^{-04}</td>
<td>Microtubule protein involved in neurogenesis and microtubule formation (Del Rio et al., 2004).</td>
</tr>
<tr>
<td>ISL1</td>
<td>ISL LIM homeobox 1</td>
<td>21.659</td>
<td>4.437</td>
<td>3.40x10^{-02}</td>
<td>Binds to the enhancer region of the insulin gene and regulates insulin gene expression (Peng et al., 2005).</td>
</tr>
<tr>
<td>CLDN1</td>
<td>Claudin 1</td>
<td>21.351</td>
<td>4.416</td>
<td>1.64x10^{-05}</td>
<td>Membrane protein involved in cell-cell adhesion and that serves as a physical barrier to prevent solutes and water from passing freely through the peri-cellular space (Morita et al., 1999).</td>
</tr>
<tr>
<td>THBS2</td>
<td>Thrombospondin 2</td>
<td>19.325</td>
<td>4.272</td>
<td>3.86x10^{-03}</td>
<td>Glycoprotein involved in cell adhesion. Negatively regulates angiogenesis (Iruela-Arispe et al., 2004).</td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Fold Change</td>
<td>Log2</td>
<td>p-value</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>------</td>
<td>------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>WISP3</td>
<td>WNT1 inducible signalling pathway protein 3</td>
<td>-26.830</td>
<td>-4.746</td>
<td>1.41x10^-04</td>
<td>Member of the WNT1 inducible signalling pathway (WISP) protein subfamily, which belongs to the CTGF/CCN family. Mutations of this gene are associated with progressive pseudo-rheumatoid dysplasia, an autosomal recessive skeletal disorder, indicating that the gene is essential for normal postnatal skeletal growth and cartilage homeostasis (Hurvitz et al., 1999).</td>
</tr>
<tr>
<td>CHST11</td>
<td>Carbohydrate (chondroitin 4) sulfotransferase 11</td>
<td>-26.430</td>
<td>-4.724</td>
<td>5.69x10^-04</td>
<td>Sulfotransferase 2 family member that catalyses the transfer of sulfate to position 4 of the N-acetylgalactosamine residue of chondroitin, forming chondroitin sulfate, the predominant proteoglycan present in cartilage (Kluppel et al., 2012).</td>
</tr>
<tr>
<td>SERPINA3</td>
<td>Serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antiproteinase), member 3</td>
<td>-24.647</td>
<td>-4.623</td>
<td>1.27x10^-03</td>
<td>Gene that encodes for alpha 1-antichymotrypsin, a protease inhibitor. This protein is produced in the liver (Kalsheker et al., 2002) during inflammation and has been found protect the lower respiratory tract from proteolytic degradation (Kalsheker, 1996).</td>
</tr>
<tr>
<td>CHAD</td>
<td>Chondroadherin</td>
<td>-21.463</td>
<td>-4.424</td>
<td>2.44x10^-03</td>
<td>Cartilage ECM protein thought to mediate adhesion of isolated chondrocytes (Hessle et al., 2014).</td>
</tr>
<tr>
<td>ZNF385B</td>
<td>Zinc finger protein 35B</td>
<td>-20.667</td>
<td>-4.369</td>
<td>2.33x10^-04</td>
<td>Zinc-finger protein expressed by B cells and that has both pro- and anti-apoptotic activities (Iijima et al., 2012).</td>
</tr>
<tr>
<td>LIF</td>
<td>Leukaemia inhibitory factor</td>
<td>-20.373</td>
<td>-4.349</td>
<td>8.86x10^-04</td>
<td>Interleukin 6 class cytokine that affects cell growth by inhibiting differentiation. It is widely used in stem cell culture to inhibit spontaneous differentiation (Kawahara et al., 2009).</td>
</tr>
<tr>
<td>PLA2G2A</td>
<td>Phospholipase A2, group IIA</td>
<td>-18.677</td>
<td>-4.223</td>
<td>1.58x10^-03</td>
<td>Extracellular enzyme involved in the regulation of the phospholipid metabolism in membranes.</td>
</tr>
<tr>
<td>PRSS33</td>
<td>Protease, serine, 33</td>
<td>-17.432</td>
<td>-4.124</td>
<td>9.16x10^-03</td>
<td>Proteolytic protein expressed by macrophages (Chen et al., 2003).</td>
</tr>
<tr>
<td>FOSL1</td>
<td>FOS-like antigen 1</td>
<td>-16.776</td>
<td>-4.068</td>
<td>5.05x10^-02</td>
<td>Regulator of cell proliferation, differentiation, and transformation (Galvagni et al., 2013).</td>
</tr>
<tr>
<td>COL11A2</td>
<td>Collagen, type XI, α2</td>
<td>-16.233</td>
<td>-4.021</td>
<td>5.44x10^-03</td>
<td>Fibrillar collagen involved in cartilage development, chondrocyte differentiation, collagen catabolic process, collagen fibril organization, ECM disassembly, ECM organization and ossification (Fernandes et al., 2007).</td>
</tr>
</tbody>
</table>
4.5.4.4 Gene pathway analysis

To identify the biological functions, pathways, networks and to reveal putative regulators of notochordal cells, dataset molecules (differentially expressed genes with fold change $<-2$ or $>2$ and $p < 0.05$) were subjected to IPA analysis. Within the top networks identified, those with highest known physiological relevance were connective tissue development and function (#1, score 66, comprising 64 dataset molecules), nervous system development and function (#2 and #3, scores 64, comprising 63 dataset molecules) and developmental, skeletal and muscular disorders (#6, score 60, comprising 61 dataset molecules) (Table 4.8).

To understand which diseases or functions may be affected by the genes in the dataset top diseases and functions were analysed; only those with a predictable activated state ($z$ score $<2$ for inhibition and $>2$ for activation) were considered. There were 503 dataset molecules associated with organism survival, with this function being predicted to be active. Within the top functions predicted to be inhibited downstream of the dataset molecules, several were associated with vasculogenesis: development of blood vessels (194, $z$ score -2.704, $p=1.89 \times 10^{-15}$), vasculogenesis (176 molecules, $z$ score -2.652, $p=5.18 \times 10^{-15}$) and angiogenesis (160 molecules, $z$ score -2.624, $p=5.63 \times 10^{-12}$) (Table 4.9).
Table 4.8 High-scoring networks (Score $>$40) identified by IPA. Each network is scored according to the number of focus molecules in the dataset and comprises a set of diseases and functions.

<table>
<thead>
<tr>
<th>ID</th>
<th>Score</th>
<th>Focus Molecules</th>
<th>Top Diseases and Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>66</td>
<td>64</td>
<td>Connective Tissue Development and Function, Tissue Development, Cellular Development</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>63</td>
<td>Post-Translational Modification, Nervous System Development and Function, Organ Morphology</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>63</td>
<td>Cellular Assembly and Organization, Nervous System Development and Function, Cell Morphology</td>
</tr>
<tr>
<td>4</td>
<td>63</td>
<td>64</td>
<td>Gene Expression, Cellular Assembly and Organization, Cellular Compromise</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>62</td>
<td>Reproductive System Development and Function, Cancer, Cellular Development</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>61</td>
<td>Developmental Disorder, Skeletal and Muscular Disorders, Cellular Assembly and Organization</td>
</tr>
<tr>
<td>7</td>
<td>58</td>
<td>61</td>
<td>Cell Cycle, Post-Translational Modification, Cellular Assembly and Organization</td>
</tr>
<tr>
<td>8</td>
<td>57</td>
<td>60</td>
<td>Cellular Movement, Cellular Development, Cellular Growth and Proliferation</td>
</tr>
<tr>
<td>9</td>
<td>54</td>
<td>59</td>
<td>Embryonic Development, Organ Development, Organ Morphology</td>
</tr>
<tr>
<td>10</td>
<td>52</td>
<td>59</td>
<td>Cellular Assembly and Organization, Cell Morphology, Infectious Disease</td>
</tr>
<tr>
<td>11</td>
<td>51</td>
<td>57</td>
<td>Behaviour, Nervous System Development and Function, Embryonic Development</td>
</tr>
<tr>
<td>12</td>
<td>51</td>
<td>56</td>
<td>Post-Translational Modification, Hereditary Disorder, Neurological Disease</td>
</tr>
<tr>
<td>13</td>
<td>49</td>
<td>55</td>
<td>Developmental Disorder, Hereditary Disorder, Organismal Injury and Abnormalities</td>
</tr>
<tr>
<td>14</td>
<td>49</td>
<td>55</td>
<td>Endocrine System Development and Function, Molecular Transport, Protein Synthesis</td>
</tr>
<tr>
<td>15</td>
<td>49</td>
<td>55</td>
<td>Cell-mediated Immune Response, Cellular Development, Cellular Function and Maintenance</td>
</tr>
<tr>
<td>16</td>
<td>47</td>
<td>54</td>
<td>Cellular Assembly and Organization, Nervous System Development and Function, Cellular Function and Maintenance</td>
</tr>
<tr>
<td>17</td>
<td>44</td>
<td>52</td>
<td>Cell Cycle, Gene Expression, Cardiovascular Disease</td>
</tr>
<tr>
<td>18</td>
<td>41</td>
<td>50</td>
<td>Lipid Metabolism, Small Molecule Biochemistry, Molecular Transport</td>
</tr>
<tr>
<td>19</td>
<td>41</td>
<td>50</td>
<td>Connective Tissue Disorders, Dermatological Diseases and Conditions, Gastrointestinal Disease</td>
</tr>
<tr>
<td>20</td>
<td>39</td>
<td>49</td>
<td>Molecular Transport, Haematological Disease, Cellular Movement</td>
</tr>
<tr>
<td>21</td>
<td>38</td>
<td>48</td>
<td>Gene Expression, Protein Synthesis, RNA Post-Transcriptional Modification</td>
</tr>
<tr>
<td>22</td>
<td>37</td>
<td>48</td>
<td>Cell Signalling, Cellular Growth and Proliferation, Cellular Movement</td>
</tr>
<tr>
<td>23</td>
<td>36</td>
<td>48</td>
<td>Developmental Disorder, Skeletal and Muscular Disorders, Connective Tissue Disorders</td>
</tr>
<tr>
<td>24</td>
<td>36</td>
<td>47</td>
<td>Cellular Assembly and Organization, Cellular Movement, Cell-To-Cell Signalling and Interaction</td>
</tr>
<tr>
<td>25</td>
<td>35</td>
<td>47</td>
<td>Connective Tissue Disorders, DNA Replication, Recombination, and Repair, Gene Expression</td>
</tr>
</tbody>
</table>
Table 4.9 Top 20 diseases and functions associated with the genes in the dataset. Only diseases and functions with a predictable activation score (z-score ≤-2 and >2) and a p value <0.05 were included.

<table>
<thead>
<tr>
<th>Categories</th>
<th>Diseases or Functions Annotation</th>
<th>p-Value</th>
<th>Predicted activation State</th>
<th>Activation z-score</th>
<th>Number of dataset molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cellular growth and proliferation</td>
<td>Proliferation of cells</td>
<td>1.39x10^{-28}</td>
<td>Decreased</td>
<td>-2.299</td>
<td>733</td>
</tr>
<tr>
<td>(Cellular movement)</td>
<td>Cell movement</td>
<td>2.35x10^{-23}</td>
<td>Decreased</td>
<td>-2.647</td>
<td>446</td>
</tr>
<tr>
<td>(Cellular movement)</td>
<td>Migration of cells</td>
<td>3.63x10^{-22}</td>
<td>Decreased</td>
<td>-2.479</td>
<td>405</td>
</tr>
<tr>
<td>Organismal survival</td>
<td>Organismal death</td>
<td>7.25x10^{-21}</td>
<td>Increased</td>
<td>3.156</td>
<td>503</td>
</tr>
<tr>
<td>Cancer</td>
<td>Malignant neoplasm of abdomen</td>
<td>2.84x10^{-18}</td>
<td>Decreased</td>
<td>-2.558</td>
<td>812</td>
</tr>
<tr>
<td>Gene expression</td>
<td>Expression of RNA</td>
<td>1.47x10^{-16}</td>
<td>Decreased</td>
<td>-3.309</td>
<td>422</td>
</tr>
<tr>
<td>Cardiovascular system development and function</td>
<td>Development of cardiovascular system</td>
<td>1.06x10^{-15}</td>
<td>Decreased</td>
<td>-2.709</td>
<td>238</td>
</tr>
<tr>
<td>(Cardiovascular system development and function, organismal development)</td>
<td>Development of blood vessel</td>
<td>1.89x10^{-15}</td>
<td>Decreased</td>
<td>-2.704</td>
<td>194</td>
</tr>
<tr>
<td>Gene expression</td>
<td>Transcription</td>
<td>3.60x10^{-15}</td>
<td>Decreased</td>
<td>-3.354</td>
<td>379</td>
</tr>
<tr>
<td>Gene expression</td>
<td>Transcription of RNA</td>
<td>3.99x10^{-15}</td>
<td>Decreased</td>
<td>-3.294</td>
<td>373</td>
</tr>
<tr>
<td>Cardiovascular system development and function, organismal development</td>
<td>Vasculogenesis</td>
<td>5.18x10^{-15}</td>
<td>Decreased</td>
<td>-2.652</td>
<td>176</td>
</tr>
<tr>
<td>Gene expression</td>
<td>Transcription of DNA</td>
<td>1.68x10^{-13}</td>
<td>Decreased</td>
<td>-2.011</td>
<td>298</td>
</tr>
<tr>
<td>Cellular movement</td>
<td>Cell movement of tumour cell lines</td>
<td>3.25x10^{-13}</td>
<td>Decreased</td>
<td>-2.189</td>
<td>183</td>
</tr>
<tr>
<td>(Cellular movement)</td>
<td>Migration of tumour cell lines</td>
<td>2.40x10^{-12}</td>
<td>Decreased</td>
<td>-2.058</td>
<td>151</td>
</tr>
<tr>
<td>Cardiovascular system development and function</td>
<td>Angiogenesis</td>
<td>5.63x10^{-12}</td>
<td>Decreased</td>
<td>-2.624</td>
<td>160</td>
</tr>
<tr>
<td>Developmental disorder, skeletal and muscular disorders</td>
<td>Congenital anomaly of musculoskeletal system</td>
<td>5.18x10^{-11}</td>
<td>Increased</td>
<td>2.298</td>
<td>147</td>
</tr>
<tr>
<td>Cardiovascular system development and function, cellular movement</td>
<td>Cell movement of endothelial cells</td>
<td>6.18x10^{-11}</td>
<td>Decreased</td>
<td>-2.067</td>
<td>91</td>
</tr>
<tr>
<td>Cardiovascular system development and function, tissue development</td>
<td>Development of cardiovascular tissue</td>
<td>6.31x10^{-11}</td>
<td>Decreased</td>
<td>-2.013</td>
<td>103</td>
</tr>
<tr>
<td>Cellular growth and proliferation</td>
<td>Proliferation of epithelial cells</td>
<td>2.19x10^{-10}</td>
<td>Decreased</td>
<td>-2.259</td>
<td>102</td>
</tr>
<tr>
<td>Organismal development</td>
<td>Size of body</td>
<td>2.79x10^{-10}</td>
<td>Decreased</td>
<td>-2.293</td>
<td>193</td>
</tr>
</tbody>
</table>
The list of differentially expressed genes was found to be associated with several canonical pathways. The top pathways were #1 hepatic fibrosis/ hepatic stellate cell activation, #2 roles of macrophages fibroblasts and endothelial cells in rheumatoid arthritis, #3 axonal guidance and signalling, #4 ILK signalling, #5 factors protecting cardiogenesis in vertebrates and #6 Wnt/ β catenin signalling. Filtering for growth factor signalling, 8 pathways were identified, including IGF-1, HGF, renin-angiotensin and FGF signalling (Figure 4.12).

**Figure 4.12** Top growth factor signalling pathways that were significantly differentially expressed in CD24^+ve cells. On the x-axis the growth factor pathways with the most significant association with the dataset are displayed and on the left y-axis the –log (Benjamini-Hochberg corrected p –value) is presented. p -value was considered significant (threshold) if it was <0.05. The jagged orange line represents the ratio (right y-axis), which is calculated by dividing the number of genes in a given pathway that meet the cut-off criteria by the total number of genes that make up that pathway.

To identify which molecules regulate notochordal cell gene expression, upstream regulator analysis was performed. Nine growth factors were identified upstream
of the genes in the dataset which had a predicted activation state – all were predicted to be inhibited. The top growth factors negatively regulating notochordal cells were TGFβ1 and fibroblast growth factor 2 (FGF2). Eighteen cytokines were predicted to be regulating notochordal cell gene expression, two of which were positive regulators: IL-1RA and wingless-type MMTV integration site family, member 5A (WNT5A). Among the negative regulators that were cytokines were TNF, IL-1β and IFN-γ (Table 4.10).
Table 4.10 Upstream regulators of differentially expressed genes. Log$_2$ ratio represents differential expression of the regulator in the dataset; activation z-score represents the activation state (inhibited if < -2 and active if > 2), and p-values represent the likelihood of the regulator regulating the genes in the dataset not being due to chance.

<table>
<thead>
<tr>
<th>Upstream Regulator</th>
<th>z-score</th>
<th>Log$_2$ Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth factors predicted to be inhibited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFβ1</td>
<td>-3.403</td>
<td>-0.700</td>
<td>7.00x10$^{-36}$</td>
</tr>
<tr>
<td>FGF2</td>
<td>-2.707</td>
<td>-2.092</td>
<td>2.89x10$^{-17}$</td>
</tr>
<tr>
<td>HGF</td>
<td>-2.796</td>
<td>0.936</td>
<td>6.72x10$^{-16}$</td>
</tr>
<tr>
<td>Angiotensinogen (AGT)</td>
<td>-2.241</td>
<td>-0.880</td>
<td>1.09x10$^{-13}$</td>
</tr>
<tr>
<td>EGF</td>
<td>-4.081</td>
<td>0.503</td>
<td>3.43x10$^{-13}$</td>
</tr>
<tr>
<td>TGFβ3</td>
<td>-2.222</td>
<td>1.538</td>
<td>2.92x10$^{-11}$</td>
</tr>
<tr>
<td>BMP7</td>
<td>-2.154</td>
<td>-0.792</td>
<td>2.60x10$^{-05}$</td>
</tr>
<tr>
<td>BMP6</td>
<td>-2.361</td>
<td>-1.841</td>
<td>9.98x10$^{-05}$</td>
</tr>
<tr>
<td>Neuregulin 1 (NRG1)</td>
<td>-2.735</td>
<td>1.170</td>
<td>3.68x10$^{-04}$</td>
</tr>
<tr>
<td><strong>Cytokines predicted to be active</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-1RA</td>
<td>2.312</td>
<td>1.530</td>
<td>2.50x10$^{-04}$</td>
</tr>
<tr>
<td>WNT5A</td>
<td>2.058</td>
<td>1.425</td>
<td>4.37x10$^{-03}$</td>
</tr>
<tr>
<td><strong>Cytokines predicted to be inhibited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNF</td>
<td>-4.728</td>
<td>0.425</td>
<td>1.21x10$^{-25}$</td>
</tr>
<tr>
<td>Oncostatin M (OSM)</td>
<td>-3.620</td>
<td>-0.131</td>
<td>8.27x10$^{-14}$</td>
</tr>
<tr>
<td>IL-1β</td>
<td>-4.721</td>
<td>3.213</td>
<td>1.66x10$^{-11}$</td>
</tr>
<tr>
<td>IL-5</td>
<td>-3.880</td>
<td>0.046</td>
<td>3.83x10$^{-09}$</td>
</tr>
<tr>
<td>IFNy</td>
<td>-4.659</td>
<td>0.025</td>
<td>1.58x10$^{-07}$</td>
</tr>
<tr>
<td>Ciliary neurotrophic factor (CNTF)</td>
<td>-2.263</td>
<td>-0.234</td>
<td>6.81x10$^{-07}$</td>
</tr>
<tr>
<td>CD40 ligand (CD40LG)</td>
<td>-2.192</td>
<td>-0.174</td>
<td>7.95x10$^{-06}$</td>
</tr>
<tr>
<td>Leukemia inhibitory factor (LIF)</td>
<td>-3.052</td>
<td>-4.349</td>
<td>1.36x10$^{-04}$</td>
</tr>
<tr>
<td>IL-6</td>
<td>-4.157</td>
<td>-0.042</td>
<td>1.68x10$^{-04}$</td>
</tr>
<tr>
<td>Tumour necrosis factor ligand superfamily member</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 (TNFSF11)</td>
<td>-2.372</td>
<td>-0.344</td>
<td>3.67x10$^{-04}$</td>
</tr>
<tr>
<td>IFNα2</td>
<td>-3.773</td>
<td>0.111</td>
<td>6.69x10$^{-04}$</td>
</tr>
<tr>
<td>Interferon lambda 1 (IFNL1)</td>
<td>-2.263</td>
<td>-0.329</td>
<td>8.16x10$^{-04}$</td>
</tr>
<tr>
<td>Erythropoietin (EPO)</td>
<td>-2.489</td>
<td>-0.365</td>
<td>1.42x10$^{-03}$</td>
</tr>
<tr>
<td>IL-15</td>
<td>-2.502</td>
<td>0.443</td>
<td>1.58x10$^{-03}$</td>
</tr>
<tr>
<td>IL-2</td>
<td>-3.410</td>
<td>0.510</td>
<td>2.35x10$^{-03}$</td>
</tr>
<tr>
<td>IL-1α</td>
<td>-2.671</td>
<td>0.891</td>
<td>2.84x10$^{-02}$</td>
</tr>
</tbody>
</table>
Finally, IPA was used to identify master regulators of notochordal cell biology. The list of master regulators was filtered to identify those which had a predictable activation state and that were either growth factors or cytokines. Again, and similar to the upstream regulators, several pro-inflammatory cytokines (IL-2, TNF, IL-5 and IL-20) were predicted to be negative upstream regulators of notochordal cells, while IL-1RA was predicted to be active. Among the growth factors, CCN2, Noggin and EDN2 were predicted to be positive regulators of notochordal cells and VEGF and TGFβ1 were predicted to be negative regulators. Table 4.11 lists the regulators with the highest z-score.
Table 4.11 Master regulators of differentially expressed genes in the CD24 microarray analysis. Log$_2$ ratio represents differential expression of the regulator in the dataset; activation z-score represents the activation state (inhibited if $<-2$ and active if $>2$), and p-values represent the likelihood of the regulator regulating the genes in the dataset not being due to chance.

<table>
<thead>
<tr>
<th>Master Regulator</th>
<th>z-score</th>
<th>Log$_2$ Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Growth factors predicted to be active</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCN2</td>
<td>3.253</td>
<td>-0.143</td>
<td>2.12 x 10^{-24}</td>
</tr>
<tr>
<td>Noggin</td>
<td>3.103</td>
<td>-2.306</td>
<td>1.19 x 10^{-14}</td>
</tr>
<tr>
<td>Endothelin 2 (EDN2)</td>
<td>2.142</td>
<td>0.059</td>
<td>5.24 x 10^{-38}</td>
</tr>
<tr>
<td><strong>Growth factors predicted to be inhibited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brain-derived neurotrophic factor (BDNF)</td>
<td>-4.507</td>
<td>0.93</td>
<td>2.73 x 10^{-22}</td>
</tr>
<tr>
<td>Neuregulin 1 (NRG1)</td>
<td>-4.415</td>
<td>1.17</td>
<td>1.64 x 10^{-28}</td>
</tr>
<tr>
<td>C-type lectin domain family 11, member A (CLEC11A)</td>
<td>-4.32</td>
<td>-0.498</td>
<td>8.39 x 10^{-17}</td>
</tr>
<tr>
<td>VEGFA</td>
<td>-3.967</td>
<td>-1.631</td>
<td>3.49 x 10^{-24}</td>
</tr>
<tr>
<td>IGF1</td>
<td>-3.768</td>
<td>4.811</td>
<td>1.32 x 10^{-27}</td>
</tr>
<tr>
<td>TGFβ1</td>
<td>-3.762</td>
<td>-0.7</td>
<td>6.41 x 10^{-17}</td>
</tr>
<tr>
<td>Epidermal growth factor (EGF)</td>
<td>-3.755</td>
<td>0.503</td>
<td>2.79 x 10^{-11}</td>
</tr>
<tr>
<td>Gastrin-releasing peptide (GRP)</td>
<td>-3.53</td>
<td>0.034</td>
<td>1.21 x 10^{-22}</td>
</tr>
<tr>
<td>Jagged 1 (JAG1)</td>
<td>-3.516</td>
<td>1.344</td>
<td>9.11 x 10^{-25}</td>
</tr>
<tr>
<td>Fibroblast growth factor 1 (FGF1)</td>
<td>-3.316</td>
<td>-0.513</td>
<td>1.60 x 10^{-26}</td>
</tr>
<tr>
<td><strong>Cytokines predicted to be active</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL-1RA</td>
<td>5.317</td>
<td>1.53</td>
<td>1.59 x 10^{-22}</td>
</tr>
<tr>
<td>IL37</td>
<td>4.73</td>
<td>0.068</td>
<td>2.05 x 10^{-21}</td>
</tr>
<tr>
<td>Sprouty-related, EVH1 domain containing 2 (SPRED2)</td>
<td>2.209</td>
<td>-0.888</td>
<td>3.77 x 10^{-13}</td>
</tr>
<tr>
<td><strong>Cytokines predicted to be inhibited</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL2</td>
<td>-5.607</td>
<td>0.51</td>
<td>5.71E x 10^{-23}</td>
</tr>
<tr>
<td>Oncostatin M (OSM)</td>
<td>-5.088</td>
<td>-0.131</td>
<td>4.76 x 10^{-20}</td>
</tr>
<tr>
<td>Chemokine (C-C motif) ligand 5 (CCL5)</td>
<td>-4.899</td>
<td>0.207</td>
<td>2.36 x 10^{-20}</td>
</tr>
<tr>
<td>TNF</td>
<td>-4.807</td>
<td>0.425</td>
<td>4.55 x 10^{-21}</td>
</tr>
<tr>
<td>TNF (ligand) superfamily, member 13b (TNFSF13B)</td>
<td>-4.741</td>
<td>0.236</td>
<td>5.89 x 10^{-17}</td>
</tr>
<tr>
<td>CXCL12</td>
<td>-4.434</td>
<td>5.053</td>
<td>8.25 x 10^{-27}</td>
</tr>
<tr>
<td>Thyroid hormone receptor interactor 6 (TRIP6)</td>
<td>-4.426</td>
<td>-0.882</td>
<td>6.35 x 10^{-12}</td>
</tr>
<tr>
<td>Vav3 guanine nucleotide exchange factor</td>
<td>-4.33</td>
<td>1.178</td>
<td>8.42 x 10^{-17}</td>
</tr>
<tr>
<td>IL5</td>
<td>-4.202</td>
<td>0.046</td>
<td>2.08 x 10^{-20}</td>
</tr>
<tr>
<td>IL20</td>
<td>-4.122</td>
<td>-0.199</td>
<td>5.84 x 10^{-20}</td>
</tr>
</tbody>
</table>
4.6 DISCUSSION

Results presented here show, for the first time, the isolation of viable human notochordal cells and their molecular characterisation. To do so, CD24, a cell surface marker that is notochord-specific between the developmental stages of 5.5-18 WPC, was used.

Due to the cell membrane localisation of this marker, a methodology was used that allowed the isolation of RNA from notochordal and sclerotomal cells that had fewer steps and was more reproducible than the methodology used in chapter 3, in which labelling with KRT18 required prior cell fixation and permeabilisation. This allowed isolation of notochordal and sclerotomal cells from 5 different specimens, adding robustness to the data and reducing the differences that could be inherent to the biology of a specific specimen.

There was a large variability in the number of events and in the proportion of CD24<sup>ve</sup> events in each of the specimens analysed. This was somewhat expected as samples varied in age, size, integrity before dissection and method of pregnancy termination (medical/ surgical); surgical samples tended to be more disrupted than samples from medical terminations of pregnancy.

The specimen age varied between 7.5-14 WPC. During these stages notochordal cells display their characteristic vacuolated morphology and the notochord changes from being a continuous rod along the embryo axis to becoming localised to the developing NP in the centre of the IVD (explained in detail in chapter 2). PCA and hierarchical clustering showed a good discrimination between both cell types (CD24<sup>ve</sup> and CD24<sup>ve</sup>) in all specimens, suggesting that...
molecular differences between each individual CD24^{+ve} sample were smaller than those between CD24^{+ve} and CD24^{-ve} samples. Furthermore, a pattern of gene expression variation was not found in the 5 CD24^{+ve} samples, indicating a relatively homogenous gene expression between these developmental stages. Finally, CD24 and T expression was significantly higher in CD24^{+ve} samples validating the methodology as a good option to separate between notochordal (CD24^{+ve}) and sclerotomal (CD24^{-ve}) cells.

Microarray analysis identified a list of differentially expressed notochordal and sclerotomal markers. Validation of these markers at the gene and protein level will be detailed, together with validation of the markers identified in the KRT18 microarray analysis, in chapter 5. CD24 was the top differentially expressed gene. As discussed in chapter 2, this molecule has previously been identified in notochordal rat NP and in the juvenile NP, suggesting its association with a notochordal ontogeny in other species and in the younger and less degenerated human NP. While its notochord-specificity in the fully matured human NP is not yet known, this marker, together with the stable expression of HIF-1alpha, GLUT-1, Shh, Brachyury, KRT18, KRT19, CA12 and an aggrecan/collagen ratio >20 has been recently proposed to define the phenotype of a healthy NP cell (Risbud et al., 2014).

The top notochordal cell markers identified here have had various attributed functions in other tissues where they have been found (embryogenesis, nervous system development, epithelial cell proliferation, cell adhesion, insulin regulation). It is unclear if these genes perform similar functions in the human spine development or whether they have different mechanisms of action. Interestingly, IGF-1, CLDN1 and THBS2 have previously been associated with the IVD.
IGF-1 has been proposed to stimulate NP cell proliferation and proteoglycan synthesis by rat, bovine and human NP cells (Okuda et al., 2001, Osada et al., 1996, Pratsinis and Kletsas, 2007, Zhang et al., 2006) and, for this reason, has been suggested to play a fundamental role in NP cell anabolism and in regulating its homeostasis. IGF-1 exerts its role by binding to the receptor IGF-1R and initiating its intracellular signalling cascade; IGF-1 binding proteins (IGFBP), which are also present in the IVD competitively bind to IGF-1R decreasing IGF-1/IGF-1R binding and its downstream effects (Asfour et al., 2015). The decrease in NP proteoglycan content seen during ageing and degeneration has been shown to be mediated by an increase in IGFBP and by a downregulation of IGF-1R (Okuda et al., 2001). Adding to this anabolic role in NP cells, IGF-1 (together with TGFβ1) has been implicated in inducing proteoglycan production and in a switch from type I to type II collagen by outer AF cells during post-natal growth and ageing (Hayes and Ralphs, 2011, Gruber et al., 2004). When combined with PDGF, IGF-1 has also been shown to have anti-apoptotic effect on AF cells (Gruber et al., 2000) and to rescue them from induced premature senescence (Gruber et al., 2008). IGF-1 could potentially reach the IVD exogenously, from the vasculature surrounding the IVD or, endogenously, being produced by IVD cells. However, and since the IVD is the largest avascular structure in the human body, IGF-1 signalling to the IVD, and particularly to its inner core, the NP, is derived by an autocrine/paracrine mechanism (Osada et al., 1996). The identification of IGF-1 as one of the top notochordal cell markers further highlights this aspect – notochordal cells secrete IGF-1, which will act autocrinally on NP cells and paracrinally on AF cells. IGF-1 secreted by notochordal cells will have an important role in maintaining IVD homeostasis,
ECM anabolism and preventing cell death and senescence, all relevant features of IVD degeneration.

CLDN1, another top notochordal cell marker, has also been previously investigated in the IVD (Gruber et al., 2007). Immunopositivity to CLDN1 (together with CLDN11 and PAR3) was found in all outer AF cells while inner AF and NP cells showed progressively less positive immunoreactivity for that protein. CLDN1, which is involved in cell polarity in epithelial cells, is hypothesised to be involved in the highly organised AF lamellar structure (Gruber et al., 2007). Expression of CLDN1 by notochordal cells in this study may be explained by the fact that these cells also display epithelioid characteristics such as cell polarity and cell-cell contact (Hay, 2005).

The identification of THBS2 as a notochordal cell marker is in agreement with previous findings highlighting the importance of this gene in the maintenance of a normal healthy disc's ECM. THBS2 (and THBS1) are IVD ECM proteins that regulate the effective levels of the catabolic proteins MMP2 and MMP9 (Yang et al., 2000, Bein and Simons, 2000). THBS2 expression has been identified in the sand rat and human IVD and has been hypothesised to be important for the maintenance of the avascular state of the IVD (Gruber et al., 2006). Kryiakides and colleagues have analysed THBS2-null mice and identified several connective tissue abnormalities including tail malformations (although the authors did not analyse this tissue in detail) (Kyriakides et al., 1998). Finally, the THBS2 rs9406328 polymorphism was found to be associated with disc herniation (Hirose et al., 2008). All these studies highlight the importance of THBS2 in regulating, possibly through MMP2 and MMP9, IVD ECM degradation. Its expression by
notochordal cells further stresses the importance of these cells in maintaining IVD ECM integrity.

Although STMN2, RTN1, PRPH and MAP1B have not been directly associated with notochordal cells and/ or the IVD field their expression is frequently associated with the nervous system and neural tissues. It should be noted that during embryonic and foetal development, the notochord displays a fundamental role in neural tube patterning, development and growth (Wilson and Maden, 2005) and it is, therefore, not unexpected that notochordal cells express markers associated with the nervous system.

Most of the top negative notochordal markers have biological functions involved in cartilage and skeletal growth and development, which is in agreement with the cartilaginous nature of sclerotomal cells, being the origin of AF and VB cells.

WISP3 is a member of the CTGF family, which is essential for normal skeletal growth and cartilage homeostasis (Hurvitz et al., 1999). The importance of WISP3 as a sclerotomal marker and in the normal development and homeostasis of the sclerotome-derived VB is highlighted by the fact that mutations in the WISP3 gene give rise to pseudorheumatoid dysplasia, a rare autosomal-recessive disorder characterised by platyspondyly (widening and flattening of the vertebral bodies), Scheuermann-like lesions (uneven growth of the vertebral bodies), and peripheral arthropathy at an unusually early stage (Mampaey et al., 2000, Yang et al., 2013).

CHST11 is an enzyme involved in the formation of chondroitin sulphate. Chondroitin sulphate, the predominant proteoglycan found in cartilage (Kluppel et al., 2012), is also synthesised by the notochord (Masuda et al., 2004). It is possible that CHST11 synthesised by sclerotomal cells is involved in the synthesis
of chondroitin sulphate by notochordal cells. It has recently been suggested that chondroitin sulphate, together with other notochordal soluble factors, may be capable of inhibiting neo-angio and neurogenesis seen with IVD degeneration, therefore being involved in the notochord-induced resistance to IVD degeneration (Purmessur et al., 2015, Cornejo et al., 2015).

CHAD is a cartilage ECM protein that is involved in cell-cell and cell-matrix interactions (Hessle et al., 2014). It provides communication between chondrocytes and their surrounding ECM and regulates collagen fibril assembly in articular cartilage (Mansson et al., 2001). In the non-degenerate IVD, CHAD has been identified in the NP and in the AF independently of patient age (Akhatib et al., 2013). Its fragmentation, however, has been found in the IVD of patients with adolescent idiopathic scoliosis (Haglund et al., 2009) and with IVD degeneration (Akhatib et al., 2013) with the number of CHAD fragments being positively correlated with the degree of degeneration. Altogether, these findings indicate that CHAD integrity may be a biomarker of a healthy IVD and that its fragmentation may be a biomarker of adolescent idiopathic scoliosis and of IVD degeneration (Akhatib et al., 2013, Tsai et al., 2007).

COL11A2 encodes one of the two alpha chains of type XI collagen, which is essential for normal embryonic skeletal development. It is proposed to be in the core of type II collagen fibrils, acting as a template for type II collagen fibrilogenesis and as a regulator of fibril diameter in cartilage (Fernandes et al., 2007). Polymorphisms in COL11A2 gene have been found to be associated with intervertebral disc degeneration (Videman et al., 2009) and with ossification of the posterior longitudinal ligament, leading to lumbar spinal stenosis (Noponen-Hietala et al., 2003) and to cervical spine myelopathy (Wilson et al., 2013).
The notochordal and sclerotomal markers identified here provide not only a phenotypic characterisation of these developing tissues, but also highlight genes that are relevant to the IVD physiology and homeostasis, some of which have already been proposed as biomarkers of a healthy IVD. Further investigations into the role of these genes in the adult normal and degenerate IVD may highlight their relevance in this tissue. Interestingly, and confirming the importance of analysing human tissue, although some of these markers have been linked to functions that may be relevant to their role in the notochordal tissue, from all the markers identified in this analysis, only CD24 had previously been associated with the notochord or with the notochordal NP (Fujita et al., 2005).

IPA was used to analyse the top biological networks associated with notochordal cells and identified “connective tissue development and function” as the top and “nervous system development and function” as the second top network. Interestingly, these two networks reflect two of the main functions of the developing notochord: providing structural support and induction of the development of the neural tube. It is the notochord elongation that provides longitudinal support to the embryo (Andre et al., 2015, Greene and Copp, 2009). Furthermore, notochordal signalling (particularly through Shh, Wnt and BMP) is fundamental for neural tube patterning, differentiation and development (Andre et al., 2015, Choudhry et al., 2014, Dahia et al., 2012, Kozmikova et al., 2013).

One of the main functions attributed to the notochord during development is the maintenance of an avascular midline region in the embryonic tissue. Studies in quail embryos have shown that notochord ablation results in vascular plexus formation in the midline and that its implantation in the paraxial or lateral mesoderm inhibits vasculogenesis in those regions (Reese et al., 2004).
negative regulatory function has been suggested to be induced by Noggin and Chordin (BMP antagonists) suppression of endothelial cell differentiation and maturation (Bressan et al., 2009, Cornejo et al., 2015). This avascular, and also aneural, state of the midline is seen in the fully matured healthy IVD. Degeneration of the IVD coincides with the migration of neo-nerve and -vessels to the IVD (Freemont et al., 1997). In fact, several functions involved in vasculogenesis were identified by IPA analysis of differentially expressed genes (development of blood vessels, vasculogenesis and angiogenesis). Confirming the inhibitory role of the notochord on vascularization, all these functions were predicted to be inhibited by the genes in the dataset. It is possible that the loss of this inhibition with notochordal cell disappearance can, itself, be in the genesis of IVD degeneration.

Many attempts have been performed to identify growth factors that could aid in the regeneration or repair of the degenerated IVD. Given that notochordal cells have a fundamental role in protecting NP cells from degeneration it was hypothesised that pathways involved in their development, particularly growth factor signalling pathways, could be responsible for this anabolic role. IPA identified IGF-1, FGF, PDGF and renin-angiotensin signalling pathways within the top growth factor pathways associated with the dataset. IGF-1, FGF2 and PDGF have been proposed to differentiate human MSC to cells with a NP phenotype (Ehlicke et al., 2010). In another study, Liu and colleagues proposed that a combination of the growth factors FGF and IGF-1 (together with EGF and VEGF) and NP ECM could induce iPSC differentiation into notochordal cells, suggesting that this methodology would generate notochord-like cells for IVD regeneration (Liu et al., 2014).
Interestingly, HGF, a molecule that was identified as a positive upstream regulator of notochordal cells isolated using their unique KRT18 expression (chapter 3) was also one of the top growth factor signalling pathways associated with notochordal cells isolated using their unique CD24 expression. This suggests that, although the list of differentially expressed notochordal genes obtained using KRT18 and CD24 methodologies differs, both methodologies allowed for the separation of notochordal (KRT18^{+ve} and CD24^{+ve}) from sclerotomal cells (KRT18^{−ve} and CD24^{−ve}).

Several factors can explain why different molecules occupy the top and bottom lists in both methodologies. First, when using the KRT18 methodology only one sample was utilised and the differentially expressed genes reflect the differences in the gene expression between notochordal and sclerotomal cells in that specific 9 WPC specimen; the genes obtained with the CD24 methodology, however, represent an average of the differentially expressed genes of a set of samples spanning from 7.5-14 WPC which may ignore specific differences inherent to the biology of a specific sample and to its developmental stage and reflect more accurately the general molecular differences between cell types. Second, it is not possible to exclude that the methodologies used, particularly the use of a fixation and permeabilisation agent, have had an impact on the detection of some specific genes. Only the analysis of the genes identified in these studies in notochordal and sclerotomal cells isolated using a different methodology would allow understanding if those gene expression changes were inherent to the specific methodology utilised. This was performed and will be detailed in chapter 5.

Finally, upstream analysis was performed to identify molecules that regulate notochordal gene expression; here, the focus was on growth factors and on
cytokines. The reason to choose these specific regulators was i) to fulfil a need for a growth factor or combination of growth factors that could drive progenitor cell differentiation to notochordal cells; ii) since cytokine regulation is paramount for the maintenance of the healthy adult NP state, it would be interesting to understand the cytokine microenvironment in which notochordal cells reside.

Two different upstream regulator analyses were performed: i) upstream regulators, which directly influence the dataset genes and ii) master regulators, which, by connecting to other regulators, could reach dataset genes, regulating notochordal, cell gene expression.

No positive growth factors were found to directly upregulate notochordal cell gene expression. Within the negative upstream regulators, however, TGFß1 and FGF2 were identified. The fact that TGFß1 is a negative upstream regulator indicates that this molecule is reported to upregulate genes that are downregulated in the dataset and to downregulate genes that are upregulated in the dataset, i.e., to induce sclerotomal cell gene expression. TGFß1 is a member of the transforming growth factor beta superfamily of cytokines that has a fundamental role in bone formation and resorption (Tang et al., 2009, Bonewald and Mundy, 1990). Indeed, at the developmental stages analysed, the sclerotomal chondrocytes populating the VB anlagen were undergoing hypertrophy to later become bone cells. Similarly, FGF2 has recently been shown to possess chondro-inductive properties in AF cells from degenerated tissues (Hegewald et al., 2013). Therefore, TGFß1 and FGF2, negative regulators of the notochordal cell gene expression, can potentially be used to drive VB and AF differentiation. Further studies confirming this are needed but fall beyond the scope of this project.
Further suggesting a protective role of notochordal cells, the pro-inflammatory cytokines TNF, IL-1β and IFNγ were found to be negative upstream regulators, while IL-1RA was found to be a positive regulator of notochordal cells. The role of pro-inflammatory cytokines in the degeneration of the NP has been extensively studied, with their expression being upregulated in degenerated discs (Le Maitre et al., 2007b, Weiler et al., 2005); while TNFα has been implicated in radicular pain occurring in patients with herniated discs, an imbalance between IL-1β and its receptor (IL-1RA) has been shown to be involved in the pathogenesis of IVD degeneration with the authors proposing IL-1RA as a potential therapy for disc degeneration (Hoyland et al., 2008, Le Maitre et al., 2007c). Taken together, these data suggests that notochordal cells reside in an environment that is relatively cytokine-free and anti-osteogenic and anti-chondrogenic.

When analysing for master regulators, three growth factors were predicted to be active: CCN5, NOG and EDN2. CCN5 is a member of the CTGF/CCN family of secreted, EMC-associated proteins (Russo and Castellot, 2010) that has been shown to be inhibitor of angiogenesis (Lake et al., 2003). Noggin is a BMP antagonist that is produced by the notochord and that is involved in somite patterning (Hirsinger et al., 1997) and, as aforementioned, in the maintenance of an avascular state in the embryo’s midline. A recent study has suggested that Noggin, together with chondroitin sulphate were notochordal soluble factors responsible for inhibiting endothelial cell invasion and blood vessel formation in the IVD (Cornejo et al., 2015). END2 is a peptide implicated in female reproduction, gastrointestinal function, immunology, cancer pathogenesis, and mainly in vascular homeostasis. A recent study has found that END2 is a potent angiogenic inhibitor in the developing retina (Rattner et al., 2013). These three
master regulators of notochordal cells have a common aspect, which is relevant to notochordal cell biology and, possibly to the notochordal cell role in protecting against IVD degeneration – they inhibit angiogenesis. Interestingly, among the growth factor master regulators that were predicted to be inhibited was VEGFA, again suggesting that notochordal cell function is widely regulated by anti-angiogenic mechanisms. Again TGFβ1 was found to be a negative master regulator of notochordal cell genes, indicating that TGFβ1 inhibits notochordal cell gene expression directly (upstream negative regulator) and indirectly through other regulators (master negative regulator).

Finally, and further stressing the protective and anti-inflammatory signalling in notochordal cells, the pro-inflammatory cytokines IL-2 TNF, TNFSF13B, IL-5 and IL-20 were all master regulators predicted to be inhibited and IL-1RA was predicted to be activated. Furthermore, and since TNF and IL-1RA were also upstream regulators predicted to be inhibited and active, respectively, this indicates that these two molecules have a direct (upstream regulators) and indirect effect through other regulators (master regulators) on notochordal cell biology.

Altogether, the data on upstream and master regulators suggests that anti-inflammatory and anti-angiogenic mechanisms are upstream of and regulate notochordal cell biology.
4.7 IMPLICATIONS OF THIS INVESTIGATION

This study reports, for the first time, a methodology to isolate viable human notochordal cells and describes their molecular profile, as well as their regulatory networks and pathways, functions and regulators. Importantly, the methodology developed here can be used to isolate and subsequently culture human notochordal cells, thereby providing a tool to study their biological function and differentiation towards a more mature NP cell phenotype. The knowledge gained from this and future studies will allow a more detailed understanding of the developmental pathways involved in IVD formation, maturation and degeneration.
CHAPTER 5

Validation of identified markers in a notochordal cell population and expression analysis in the adult NP
5.1 INTRODUCTION

The search for cell-based therapies for IVD degeneration and the observation that notochordal cells, or factors they produce, may be protective and exert an anabolic effect on small NP cells has prompted the investigation of the notochordal cell phenotype. Elucidation of the notochordal cell phenotype would allow a better understanding of these cells and of their role in the IVD. While studies suggest that notochordal cells are retained in the human adult NP, they are reported to undergo a morphological change to small round and non-vacuolated cells (Yang et al., 2009). As such, and since in humans, this morphological change occurs a few years after birth, notochordal cells used here were isolated from the developing spines of human embryos and foetuses. Due to the small specimen size, and to difficulties in obtaining a pure notochordal cell population without contamination from adjacent sclerotomal cells using conventional microdissection techniques, a methodology was developed to isolate human notochordal cells using their unique protein expression of KRT18 (chapter 3) and CD24 (chapter 4).

In chapter 3, and due to technical difficulties in obtaining RNA from cells that had been fixed, permeabilised, labelled and sorted, and also to delays in sample acquisition, very little RNA was obtained (with only one specimen (9 WPC) being acquired). The isolated RNA was used to compare the gene expression of notochordal and sclerotomal cells and to identify notochordal cell surface markers. Due to sample depletion, however, there was insufficient RNA for validation.
In chapter 4, RNA was obtained from sclerotomal and notochordal cells from 5 specimens (7.5-14 WPC), compared used microarrays and utilised to identify notochordal cell markers and to understand notochordal cell biology.

Microarrays constitute a powerful tool to measure the activity of genes being expressed by a given cell in a specific moment on a genome scale, allowing an understanding of the interactions between cells and the environment in which they reside. While this valuable information has been used to characterise disease states, predict disease progression and develop new therapies (Karlsson et al., 2010, Saito et al., 2010, Yang et al., 2010), some authors have highlighted a few caveats to this technique, particularly the occurrence of type I errors (false positive detection). As such, qRT-PCR, is often used as a reliable independent technique to validate genes identified using microarrays (Wang et al., 2006).

Furthermore, and although care was taken to minimise all detrimental effects of the methodology used for cell isolation, it is also possible that 1) enzymatic tissue digestion (Autengruber et al., 2012), 2) cell fixation, permeabilisation and labelling (Diez et al., 1999, Esser et al., 1995) and 3) the mechanical forces applied during FACS (Li et al., 2013) could induce small but important changes to the cell phenotype. As such, it is important to assess/validate the expression of the genes identified in chapters 3 and 4 in notochordal and sclerotomal cells isolated from spinal tissues using non-enzymatic methods and separated without the need for labelling and cell sorting.

In molecular biology it is well accepted that information is transcribed from DNA to mRNA and mRNA is translated into amino acids, which are subsequently folded into functional proteins in the ribosome. However, due to factors such as rate of transcription versus translation, protein stability, RNA processing,
alternative splicing and post-transcriptional regulation, the transcriptome (mRNA) often doesn’t correlate with the translatome (protein) (Vogel and Marcotte, 2012, Maier et al., 2009). Thus, a more in-depth understanding of the notochordal cell function requires the validation of the identified genes at the protein level.

Finally, while mice fate mapping studies have shown that notochord-derived cells persist in the adult NP (Choi et al., 2008, McCann et al., 2012), their presence in the human adult NP using human notochordal cell markers has not yet been assessed and it is not known if the expression of such markers varies with IVD degeneration.
5.2 HYPOTHESES AND AIMS

The hypotheses for this study were that:

1. The notochordal cell genes (identified in chapters 3 and 4 using microarrays) could be validated using different gene expression analysis techniques in the same samples, in different samples and in a notochordal cell population isolated using a different isolation procedure and that, at least a subset of genes, would be translated to proteins, constituting the “true” notochordal cell gene and protein phenotype.

2. The analysis of the expression of such genes in adult disc samples would allow the understanding of whether notochord-derived cells persist in the human adult IVD and how their expression changes with degeneration.

The aims for this study were to:

1. Validate, at the gene level, the markers identified in chapters 3 and 4 in:
   b. Notochordal and sclerotomal cells isolated without the need for enzymatic tissue digestion, cell labelling and sorting.

2. Analyse, at the protein level, the expression of the validated notochordal genes in embryonic and foetal spine tissue.

3. Analyse the gene expression of the identified markers in adult NP samples and correlate their expression with IVD degeneration.
4. Analyse the protein expression of a panel of notochordal markers in human adult NP tissue.
5.3 MATERIALS AND METHODS

5.3.1 Sample acquisition and processing

5.3.1.1 Embryonic and foetal samples

Embryonic and foetal samples were acquired and staged as described in 2.3.1 and dissected to obtain whole spines containing IVD and VB as detailed in 2.3.2. For qRT-PCR three sets of samples were used:

1. RNA from CD24^{+ve} and CD24^{-ve} cells extracted from the samples used in the CD24 microarray in which the RNA had not been depleted (M728 CD24^{+ve}, M753 CD24^{+ve}, M729 CD24^{+ve}, M739 CD24^{+ve} and M739 CD24^{+ve}) (Table 5.1).
2. RNA from CD24^{+ve} and CD24^{-ve} cells sorted from 2 additional samples (M792 and M801) (Table 5.1).
3. RNA extracted from notochordal and sclerotomal tissue microdissected from an older (M784, 18 WPC) sample (Table 5.1).
Table 5.1 Samples used for microarray validation. RNA for microarray validation was obtained from developing notochordal and sclerotomal cells using two separation methods: 1) FACS of CD24⁺ve and CD24⁻ve cells and 2) microdissection of notochordal and sclerotomal tissues from an 18 WPC sample.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stage (WPC)</th>
<th>Method of separation</th>
<th>Cell/ tissue type</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>M728</td>
<td>7.5</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>Microarray analysis &amp; qRT-PCR</td>
</tr>
<tr>
<td>M753</td>
<td>8.5</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>Microarray analysis &amp; qRT-PCR</td>
</tr>
<tr>
<td>M729</td>
<td>12</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>Microarray analysis &amp; qRT-PCR</td>
</tr>
<tr>
<td>M739</td>
<td>14</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>Microarray analysis &amp; qRT-PCR</td>
</tr>
<tr>
<td>M792</td>
<td>9</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>M801</td>
<td>11</td>
<td>FACS</td>
<td>CD24⁺ve</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>M784</td>
<td>18</td>
<td>Microdissection</td>
<td>Notochordal</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Microdissection</td>
<td>Sclerotomal</td>
<td>qRT-PCR</td>
</tr>
</tbody>
</table>

Samples used for immunohistochemistry ranged between 6-19 WPC. Sample ID and age is detailed in table 2.1 with the addition of specimen M809 (19 WPC), which was acquired at a later stage and, therefore only used for microarray validation.

To obtain notochordal and sclerotomal cells from embryonic (7.5 WPC) and younger foetal spines (8-14 WPC), the whole foetal spine was digested to obtain cells, which were labelled with the notochordal marker CD24 and sorted using FACS.

The older (18 WPC) specimen was much larger in size and the anatomical demarcations between sclerotomal (AF and VB anlagens) and notochordal regions were discernible. For this reason notochordal and sclerotomal regions were dissected as described in 2.3.2 (under sterile conditions, using microsurgical
instruments and a stereomicroscope) and notochordal and sclerotomal tissues digested to obtain RNA. To confirm that, at this stage, the NP is fully formed by large vacuolated notochordal cells IVD sections were cut, fixed, decalcified, processed to wax, mounted, sectioned and stained with H&E, as described in chapter 2.

5.3.1.2 Adult NP tissue

Adult IVD tissue was obtained from an Intervertebral Disc Tissue Bank, held at the University of Manchester. Samples used in the study were collected at the time of surgery from 17 patients with disc degeneration diagnosed with magnetic resonance imaging that underwent disc replacement or spinal fusion; informed consent was given by all patients and the study was approved by the local research ethics committee (08/H1010/36). After obtaining tissue from the patient, it was placed in a 50mL centrifuge tube containing DMEM supplemented with 1% (v/v) antibiotic/antimycotic solution, 1mM sodium pyruvate (Sigma Aldrich®, S8636) and 1mM ascorbate (0.25g ascorbic acid salts dissolved in 100mL of deionised water), stored at 4°C and transported to the laboratory for use within 24 hours.

From each specimen, a portion incorporating the AF and the NP was used for histology and immunohistochemistry. All samples were graded for histological changes of degeneration by an experienced histopathologist (Professor Anthony Freemont) and according to the features of degeneration described in Table 5.2. To do so, samples were fixed, decalcified, processed to wax, mounted, sectioned and stained with H&E, as described in chapter 2.
NP was dissected from AF tissue and used for gene expression analysis, as described below. Table 5.3 details the adult samples used for gene expression analysis and for immunohistochemistry.

### Table 5.2 Histological grading system for IVD degeneration (Sive et al., 2002)

#### Breakdown of the four groups of parameters used for the scoring system

<table>
<thead>
<tr>
<th>Loss of demarcation between the NP and AF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Loss of proteoglycan from the NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Presence and extent of fissures</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cell cluster formation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>
Table 5.3 Details of adult NP samples used for gene (qRT-PCR) and protein (immunohistochemistry (IHC)) analysis.

<table>
<thead>
<tr>
<th>ID</th>
<th>Anatomical region</th>
<th>Grade</th>
<th>Patient age</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH0520</td>
<td>L5/S1</td>
<td>2</td>
<td>31</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0512</td>
<td>L4/L5</td>
<td>3</td>
<td>43</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0402</td>
<td>C6/C7</td>
<td>4.5</td>
<td>49</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0529</td>
<td>C5/C6</td>
<td>4.5</td>
<td>53</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0268</td>
<td>L5/S1</td>
<td>4</td>
<td>22</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0362</td>
<td>C5/C6</td>
<td>4</td>
<td>43</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0369</td>
<td>C3/C4</td>
<td>4</td>
<td>71</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0269</td>
<td>L5/S1</td>
<td>5</td>
<td>35</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0368</td>
<td>C4/C5</td>
<td>6</td>
<td>71</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0384</td>
<td>C6/C7</td>
<td>6</td>
<td>47</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0420</td>
<td>C4/C5</td>
<td>6</td>
<td>55</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0422</td>
<td>L4/L5</td>
<td>6</td>
<td>25</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0427</td>
<td>C5/C6</td>
<td>6</td>
<td>63</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0524</td>
<td>C5/C6</td>
<td>6</td>
<td>41</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0380</td>
<td>C6/C7</td>
<td>7</td>
<td>49</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0418</td>
<td>L4/L5</td>
<td>7</td>
<td>25</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0513</td>
<td>L4/L5</td>
<td>7</td>
<td>54</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0518</td>
<td>C6/C7</td>
<td>7</td>
<td>48</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0383</td>
<td>C5/C6</td>
<td>8</td>
<td>49</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0416</td>
<td>L5/S1</td>
<td>8</td>
<td>46</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0422</td>
<td>L4/L5</td>
<td>8</td>
<td>22</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0526</td>
<td>C4/C5</td>
<td>8</td>
<td>49</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0440</td>
<td>L4/L5</td>
<td>9</td>
<td>31</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0408</td>
<td>L5/S1</td>
<td>10</td>
<td>33</td>
<td>qRT-PCR &amp; IHC</td>
</tr>
<tr>
<td>HH0415</td>
<td>L5/S1</td>
<td>11</td>
<td>42</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>HH0593</td>
<td>L4/L5</td>
<td>11</td>
<td>27</td>
<td>IHC</td>
</tr>
<tr>
<td>HH0386</td>
<td>L5/S1</td>
<td>12</td>
<td>38</td>
<td>IHC</td>
</tr>
</tbody>
</table>
5.3.2 RNA isolation from embryonic and foetal notochordal and sclerotomal cells

To obtain notochordal and sclerotomal cells from embryonic (7.5 WPC) and younger foetal spines (8-14 WPC), the whole foetal spine was digested to obtain cells, cells were labelled with the notochordal marker CD24 and sorted using FACS into a tube containing lysis buffer as described in 4.4.3. RNA from 7.5-14 WPC spines was obtained using the Cell Protocol from the RNeasy micro plus kit (Qiagen®), as described in 3.4.6.3. RNA quantity, quality, purity and integrity was analysed as described in 3.4.6.4. RNA was amplified and purified as described in 3.4.7.2.

5.3.3 RNA isolation and cDNA synthesis from foetal notochordal and sclerotomal tissue

5.3.3.1 Notochordal and sclerotomal tissue dissection

After dissecting the whole 18WPC spine, axial sections through the IVD region were cut; those sections were made through the IVD itself to avoid contamination with adjacent VB and sclerotomal regions. As notochordal regions were visible as gelatinous central regions encircled by the lamellar AF regions (Figure 5.1 A), a 0.8mm sample corer (Fine Science Tools®, 18035-80) was used to puncture and isolate the central notochordal region (Figure 5.1 B). Sclerotomal regions were obtained from the peripheral AF and VB anlagens. To obtain sufficient tissue,
nootchordal and sclerotomal samples were pooled from more than one spinal segment. Samples were immediately snap frozen in liquid nitrogen and stored at -80°C for RNA extraction.

Figure 5.1 Photomicrograph of a section through the IVD of sample M784 (18WPC). In (A) the central and gelatinous notochordal NP (arrow head) is clearly distinguishable from the peripheral and fibrillar sclerotomal AF (*). In (B) the same IVD is shown after notochordal region puncture.

5.3.3.2 RNA extraction, amplification, purification and cDNA synthesis

For RNA extraction from the 18 WPC spinal tissues the RNeasy micro plus kit (Qiagen®) Tissue Protocol was used. To disrupt notochordal and sclerotomal tissue a glass tissue grinder was used. All steps were performed on dry ice to avoid tissue from thawing. Frozen notochordal and sclerotomal tissue was quickly transferred to appropriately labelled glass 2mL dounce tissue grinder set (Sigma Aldrich®, D8938) containing liquid nitrogen and was thoroughly ground. After liquid nitrogen evaporation, 350µL of RLT lysis buffer was added to each tube and mixed well by pipetting; at this stage, samples were warmed to 37°C. Disrupted tissues were then pipetted into a QIAshredder spin column placed in a
2mL MCT and centrifuged for 2 minutes at 19,700G. The homogenised lysate was then centrifuged for another 3 minutes at 19,700G, the supernatant carefully removed by pipetting and transferred to a gDNA Eliminator spin column placed in a 2mL collection MCT. Samples in gDNA Eliminator spin columns were centrifuged for 30 seconds at 10,000G, the columns were discarded and 350µL of freshly prepared molecular grade 70% v/v ethanol was added to the DNA-free lysate, mixed well by pipetting and transferred to an RNeasy MinElute spin column placed in a 2mL collection MCT. The column was centrifuged for 30 seconds at 10,000G and the flow-through discarded. Then, 700µL of Buffer RW1 was added to the column, centrifuged for 30 seconds at 10,000G and the flow-through was discarded. Next, 500µL of Buffer RPE was added to the column and centrifuged for 30 seconds at 10,000G to wash the spin column; the flow-through was discarded. To ensure complete ethanol removal from the membrane, another 500µL of buffer RPE was added to the column and centrifuged for 2 minutes at 10,000G. After centrifugation, the column was carefully removed from the collection MCT, placed in a new 2mL collection MCT and centrifuged for 2 minutes at 19,700G. Finally, the column was carefully removed from the collection MCT and placed in a 1.5mL collection MCT and 30µL of molecular grade water (warmed to 60°C) was added to the column; the column was then centrifuged at 19,700G for 4 minutes, to elute the RNA. To allow for higher RNA recovery, the eluent was re-pipetted through the column and centrifuged again at 19,700G for 4 additional minutes. Eluted RNA was stored at -80°C. RNA quantity, quality, purity and integrity was analysed as described in 3.4.6.4. RNA was amplified and purified as described in 3.4.7.2.
5.3.4 RNA isolation and cDNA synthesis from adult NP cells

5.3.4.1 NP cell isolation

NP tissue was dissected from the AF using a sterile scalpel and macerated to small tissue pieces using sterile techniques. Dissected tissue was digested overnight at 37°C in a 50mL centrifuge tube containing 25mL of 0.1% (w/v) type II collagenase and 0.1% (w/v) hyaluronidase in DMEM (supplemented with 1% (v/v) antibiotic/antimycotic solution). Digestion solution was then passed through a 40µm cell sieve to remove any undigested tissue and centrifuged at 400G for 5 minutes to pellet cells.

NP cells were re-suspended in 5mL of NP cell medium (DMEM supplemented with 10% v/v FCS, 1mM sodium pyruvate and 1% (v/v) antibiotic/antimycotic solution), counted (as detailed in 3.4.4) and seeded into a T25 cell culture flask at a cell density of 20,000 cells/cm². Cells were allowed to adhere and grow as detailed in 3.4.2. When 80% confluence was reached, cells were trypsinised, trypsin was neutralised with cell culture medium and cells were pelleted for RNA extraction (Pockert et al., 2009).

5.3.4.2 RNA extraction

RNA from adult NP cells was extracted using TRIzol (Invitrogen®, 15596018). One millilitre of TRIzol was added to the cell pellet, pipetted well to mix and incubated for 10 minutes at room temperature. After incubation the cell
suspension was transferred to a 1.5mL MCT and stored at -80°C for RNA extraction.

For RNA extraction, MCT containing cells lysed in TRIzol were placed on ice and allowed to thaw. Then, MCT were centrifuged at 12,000g for 15 minutes at 4°C to remove cell debris and the supernatant transferred to a clean 1.5mL MCT. Next, 200µL of chloroform (Sigma Aldrich®, C2432) was added, the MCT vortexed for 20 seconds and incubated at room temperature for 3 minutes. After incubation, MCT were centrifuged at 12,000G for 15 minutes at 4°C, after which the upper aqueous phase (approximately 500µL) was transferred to a clean MCT – care was taken not to disturb the interface. Then, 2µL glycoblue (Ambion®, AM9515) and 500µL of isopropanol were added to the aqueous phase, the MCT were inverted 10 times to mix and incubated at room temperature for 10 minutes. After incubation, MCT were centrifuged at 12,000G for 20 minutes at 4°C. After this step, a small blue pellet was visible. Taking care not to disturb the pellet, the supernatant was removed by pipetting and 1mL of ice-cold 75% ethanol (v/v) was added to the pellet. After vortexing briefly, the MCT were centrifuged at 12,000G for 5 minutes at 4°C, and the ethanol was pipetted out, taking care not to disturb the pellet. The MCT were again briefly spun down to drawn down any excess ethanol, which was again pipetted off. MCT containing the pellet were allowed to air dry for approximately 10 minutes (while watching closely so as not to over-dry the pellet), after which 21.2µL of Tris-EDTA buffer was added. RNA concentration and purity were analysed as described in 3.4.6.4 using 1.2µL of the sample.
5.3.4.3 Genomic DNA elimination

To ensure removal of genomic DNA contamination, quantified RNA was treated with DNase. For such, 1µL of DNase I enzyme (Invitrogen®, 18068-015) and 2µL of DNase I reaction buffer (Invitrogen®, P/N-y02340) were added to 20µL of RNA and the mixture was incubated for 30 minutes at 37ºC. After incubation, 77µL of RNase-free water was added to the samples. To extract the RNA and remove DNase I, 100µL of Phenol:Chloroform:IAA (Sigma Aldrich®, P3803) was added and samples were vortexed to an emulsion. After this, samples were centrifuged at 12,000G for 5 minutes at 4ºC. The upper aqueous phase containing the RNA was transferred to a clean 1.5mL MCT and the RNA was again precipitated by adding 1µL of glycoblue, 10µl ammonium acetate and 330µl 100% ethanol and inverting gently to mix. MCT were then incubated for 30 minutes on ice and centrifuged at 12,000G for 5 minutes at 4ºC to form a blue pellet and the supernatant was removed by pipetting. Pellets were then washed in ice-cold 75% ethanol, the MCT centrifuged at 12,000G for 5 minutes at 4ºC and the supernatant removed. MCT were then briefly spun down to draw any remaining ethanol, which was pipetted out, after which the MCT were left with the lids open for approximately 10 minutes to allow any excess ethanol to evaporate, but taking care not to over dry the pellet. Pellets containing RNA were re-suspended in 21.2µl Tris-EDTA buffer, RNA was quantified as detailed 3.4.6.4 and stored at -80ºC.
5.3.4.4 cDNA synthesis

DNA-free RNA samples were reverse transcribed using the High Capacity RT Kit (Applied Biosystems®, 4368814). All samples with an RNA concentration above 200ng/µL were diluted to 200ng/µL. A Reverse Transcription mastermix (Table 5.4) was prepared on ice and 10µL added to 0.5mL MCT. Then, 10µL of RNA from each sample was added to each MCT and vortexed for 20 seconds to mix. MCT were then placed in a Thermal Cycler (MJ Research PTC-200, Peltier Thermal Cycler) to run the program described in Table 5.5. After reverse transcription, molecular biology grade water was added to the samples to dilute them to a concentration of 5ng/µL. Samples were stored at -20ºC.

Table 5.4 Reverse transcription mastermix. Volumes are per reaction.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x RT buffer</td>
<td>2</td>
</tr>
<tr>
<td>25x 100mM dNTP mix</td>
<td>0.8</td>
</tr>
<tr>
<td>10x random primers</td>
<td>2</td>
</tr>
<tr>
<td>Multiscribe Reverse Transcriptase</td>
<td>1</td>
</tr>
<tr>
<td>RNase inhibitor</td>
<td>1</td>
</tr>
<tr>
<td>Molecular biology-grade water</td>
<td>3.2</td>
</tr>
</tbody>
</table>

Table 5.5 Thermal cycler programming for cDNA synthesis.

<table>
<thead>
<tr>
<th>Temperature (ºC)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>10 minutes</td>
</tr>
<tr>
<td>37</td>
<td>120 minutes</td>
</tr>
<tr>
<td>85</td>
<td>5 seconds</td>
</tr>
<tr>
<td>4</td>
<td>∞</td>
</tr>
</tbody>
</table>
5.3.5 qRT-PCR

RNA quantity, quality, purity and integrity were assessed as described in 3.4.6.4. RNA from embryonic and foetal spines was amplified to cDNA and purified as described in 3.4.7.2.

qRT-PCR using gene specific primers was performed using the TaqMan method for the genes GAPDH, Mitochondrial ribosomal protein L19 (MRPL19), and CD24 as described in 3.4.7.3. Specific details of qRT-PCR assays are detailed in table 5.6. Gene expression of STMN2, RTN1, PRPH, CXCL12, IGF1, MAP1B, ISL1, CLDN1, THBS2, WISP3, CHST11, SERPINA3 and CHAD was analysed using the SYBR green method (rationale for the selection of genes to validate is given in the results sections). Pre-optimised SYBR green primers were purchased from PrimerDesign®. Specific details of qRT-PCR assays are detailed in table 5.6. An assay mastermix was prepared for each gene as detailed in Table 5.7. Two microliters of 5ng/µL cDNA from each sample were added to 8µL of the mastermix in a 0.2mL 96-well PCR plate. The plate was sealed with optical adhesive film, vortexed for 10 seconds and centrifuged for 1 minute at 2000G. The plate was then run on an Applied Biosystem StepOnePlus® instrument using the settings described in table 5.8. Total foetal cDNA was used as positive control and water as negative control.
### Table 5.6 Human oligonucleotide primers and probes used for qRT-PCR analysis

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>NCBI RefSeq</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Probe sequence</th>
<th>Optimal primer concentration (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>NM_001256799</td>
<td>CTCCTCTGACCTCAACAG</td>
<td>CGTTGCTATCCACAGAAA</td>
<td>CACCCACTCCACACCTTTGA</td>
<td>600</td>
</tr>
<tr>
<td>MRPL19</td>
<td>NM_014763</td>
<td>CACCCGCGCGTGGA</td>
<td>TCCCTCTGAGAAATGGAATTC</td>
<td>AAGCAGTTCTTGAGTC</td>
<td>900</td>
</tr>
<tr>
<td>GRB14</td>
<td>NM_004490</td>
<td>CACCGCCCCGTGGAA</td>
<td>TCCCCTTCGAGGAATGAATTC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>SLC19A1</td>
<td>NM_001205206</td>
<td>CGGCTCTACGTTCTCTCT</td>
<td>CGAAGACGAGGCCACAGA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>FGF10</td>
<td>NM_004465</td>
<td>TGGCTGAAAAACCCATTACA</td>
<td>CATTTTCTTCTATCTCTCTCTCA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>NOL12</td>
<td>NM_024813</td>
<td>AAGAAGAGAGCAGATGTTGAC</td>
<td>C1CCCTAATGCTGCTTC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>C5AR1</td>
<td>NM_001736</td>
<td>AGCCCAGGAGACCAGAAGT</td>
<td>CGCAGCGTGTTAGAAGTTT</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>SERPINB6</td>
<td>NM_001195291</td>
<td>CGCCGTAGAGAAGTCCAGAA</td>
<td>ACCAGAACCAGCTTTGCAA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>ADORA3</td>
<td>NM_020683</td>
<td>TTGACCAAAAGGAGGAGAAGT</td>
<td>AGTCACATCTGTGTCAGATGAAG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>TBX2A2R</td>
<td>NM_001066</td>
<td>GGCTCAGCTCCTCCTG</td>
<td>GCACTGTCTGGGCGATGAAG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CLDN12</td>
<td>NM_001185072</td>
<td>AAACTTTGCTCTTCTTTCTG</td>
<td>AAACACTGGAATGTCATTTCAATG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CDH6</td>
<td>NM_004832</td>
<td>ATGACATCAATGACAATGAACCA</td>
<td>TCXCGAGTGACCTTGACAC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>ANPEP</td>
<td>NM_001150</td>
<td>GACCCCGATTTCAAAGTACC</td>
<td>GTTCGTCGTCCTTACATCTAC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>SELE</td>
<td>NM_000450</td>
<td>GGGCTGAGAGGGAGTTGAAAT</td>
<td>AAGGCTGAAACATTTTACACTTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CD69</td>
<td>NM_001781</td>
<td>GCCCAATGCTTCTCCTCTG</td>
<td>TGTCTCTCCTACCTCCGAT</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>EPHA9</td>
<td>NM_004499</td>
<td>ATTTATTGGATTCACGCACTGTC</td>
<td>TCTCCACTTCACCAATCTCTTC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>STAB1</td>
<td>NM_015316</td>
<td>GCCACATTCAAGCCAGAG</td>
<td>TGAAGACGAGGCTCCAGAGA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CD24</td>
<td>NM_001291737</td>
<td>GTCCTCTACCCACGCAGATTTAT</td>
<td>CTTTGTCGTCGAGATTGAG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>STMN2</td>
<td>NM_001199214</td>
<td>AAGCCCAAGAAGACCTTAT</td>
<td>CTTTGACTCCTCCTTCTCTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>RTN1</td>
<td>NM_021136</td>
<td>AATATCCTCTGACTTTAGAAGGCTC</td>
<td>TTGTCACACTCCAGACATTCTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>PRPH</td>
<td>NM_000626</td>
<td>TGCTCCTCTAATAATAGAACGAC</td>
<td>CCGGCTCCTAATGTTCTTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CXCL12</td>
<td>NM_000609</td>
<td>GTTGGTCTGTCGTCTCTCTTCT</td>
<td>TCGACAGTTGCTGCTTAG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>IGF1</td>
<td>NM_000618</td>
<td>GGTGAGATCTTCTCTCAGTTC</td>
<td>CGTCGAGACCTCATTCC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>MAP1B</td>
<td>NM_005900</td>
<td>TGGGAGGAGAGAGAACAGAGGA</td>
<td>GGGGCACACAGATGACT</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>ISL1</td>
<td>NM_002202</td>
<td>GCAGATGAGCATAGATACAGG</td>
<td>GTGTTAGGGATATGGAATTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CLDN1</td>
<td>NM_021101</td>
<td>TCCCCAAAGAGCAAGAAGAAGAG</td>
<td>TCCCAAGAGCCAGAAGAAG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>THBS2</td>
<td>NM_003247</td>
<td>CACCAACGCCACCTAACC</td>
<td>CGTCATGCTCTCAGTCA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>WISP3</td>
<td>NM_003880</td>
<td>TGGCAGACCAATATATAGAAGAC</td>
<td>GAGTACTGGAAGCCTCAAAGA</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CHST11</td>
<td>NM_001173982</td>
<td>CTTTTATCTCCTGCTCTTCTCTTCTTCC</td>
<td>GCTTGATTGGTGTGAGTGTCCTGC</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>SERPINB6</td>
<td>NM_001085</td>
<td>TCACAGGGGAGCCAGAAC</td>
<td>GCACAGGAGGGGTGATTTG</td>
<td>N/A</td>
<td>900</td>
</tr>
<tr>
<td>CHAD</td>
<td>NM_001267</td>
<td>AGCCCAATGCCCCACCTC</td>
<td>TTTATGAAATCTCAGGAGAATTC</td>
<td>N/A</td>
<td>900</td>
</tr>
</tbody>
</table>
Table 5.7 SYBR green qRT-PCR assay mastermix. Volumes are per reaction.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x mastermix</td>
<td>5 µL</td>
</tr>
<tr>
<td>Primer</td>
<td>0.5 µL</td>
</tr>
<tr>
<td>RNase/DNase-free water</td>
<td>2.5 µL</td>
</tr>
</tbody>
</table>

Table 5.8 Thermal cycler programming for qRT-PCR.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature (°C)</th>
<th>Length of time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95.0</td>
<td>20 seconds</td>
</tr>
<tr>
<td>2</td>
<td>95.0</td>
<td>1 second</td>
</tr>
<tr>
<td>3</td>
<td>60.0</td>
<td>20 seconds</td>
</tr>
<tr>
<td>4</td>
<td>Repeat steps (2-3) 40 times</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Data was analysed using the previously described $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

The expression of the genes GRB14, ADORA3, CDH6, ANPEP, CD69, CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 in the adult NP normalised to the housekeeping gene MRPL19 was represented as box and whisker graphs to infer about the expression of these genes in adult NP tissue. In the box and whisker graphs, the top and the bottom boxes represent the first and third quartiles and the line in their centre represents the median. The whiskers reflect the minimum and the maximum of all the data.

The expression of the same genes was then correlated with the degree of degeneration of each disc sample and analysed using Spearman correlation coefficient. This test analyses the relationship between paired data and calculates an r value (range: -1 to 1). Values between -1 and 0 represent a negative correlation and those between 0 and 1 represent a positive correlation. A correlation is perfect if r is ±1. Values between ±0.01 and ±0.19 represent a very
weak correlation, those between $\pm 0.20$ and $\pm 0.39$ represent a weak correlation, those between $\pm 0.40$ and $\pm 0.59$ represent a moderate correlation, those between $\pm 0.60$ and $\pm 0.79$ represent a strong correlation and those between $\pm 0.80$ and $\pm 0.99$ represent a very strong correlation.

5.3.6 Immunohistochemistry

Expression of ADORA3, MAP1B, PRPH and RTN1 in embryonic and foetal spine and adult NP tissues was assessed using immunohistochemistry. Dissected embryonic and foetal spine and adult tissues were fixed, decalcified, processed, mounted and sectioned as described in 2.3.3.

ADOR3, MAP1B, PRPH and RTN1 antibodies were kindly optimised by Mrs Sonal Patel (the rationale for choosing these proteins is given in the results section). Immunohistochemistry was performed as described in 2.3.7. All primary antibodies were diluted in 1% BSA in TBS and incubation was performed overnight at 4°C. Table 5.9 details the antibodies, concentrations and antigen retrieval methods used. For pressure cooker antigen retrieval, 2L of citrate buffer pH 6.0 were heated in the pressure cooker for 15 minutes, after which the slides were placed in the pressure cooker chamber and the lid closed and left for 3 minutes once the pressure was reached. Then, the chamber was placed under running tap water until depressurised, the lid removed and the slides placed under running tap water for 10 minutes. Images were analysed as described in 2.3.9.
Table 5.9 Details of the antibodies and antigen retrieval methods used.

<table>
<thead>
<tr>
<th>Marker</th>
<th>Antibodies (optimised concentration, clonality, manufacturer, catalogue number)</th>
<th>Antigen retrieval method</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADORA3</td>
<td>0.05µg/mL rabbit polyclonal anti-ADORA3 IgG (Sigma Aldrich®, HPA028509)</td>
<td>Heat (Pressure Cooker) Citrate</td>
</tr>
<tr>
<td>MAP1B</td>
<td>0.4µg/mL mouse monoclonal anti-MAP1B IgG1 (Abcam®, ab3095)</td>
<td>No antigen retrieval</td>
</tr>
<tr>
<td>PRPH</td>
<td>0.05µg/mL mouse monoclonal anti-PRPH IgG1 (Vector®, VP-P968)</td>
<td>Heat TrisEDTA</td>
</tr>
<tr>
<td>RTN1</td>
<td>10µg/mL mouse monoclonal anti-RTN1 IgG1 (Abcam®, ab8957)</td>
<td>Heat Citrate</td>
</tr>
</tbody>
</table>

### 5.3.7 Statistical analysis

Statistical analysis was performed using GraphPad InStat software (GraphPad Software Inc.®) using the Mann-Whitney U-test; p values < 0.05 were defined as being representative of a significant difference.
5.4 RESULTS

5.4.1 Validation of the markers identified in the KRT18 microarray analysis

5.4.1.1 Gene expression of KRT18$^{+ve}$ and KRT18$^{-ve}$ markers

To validate the genes identified in the KRT18 microarray analysis, three of the 4 top KRT18$^{+ve}$ markers (GRB14, SLC19A1 and FGF10), the top 3 KRT18$^{-ve}$ (or bottom KRT18$^{+ve}$) markers (NOL12, C5AR1, SERPINB6) and 9 of the top 10 cell surface markers (ADORA3, TBXA2R, CLDN12, CDH6, ANPEP, SELE, CD69, EPHA5 and STAB1) were chosen for qRT-PCR validation. HLA-DQB1 (one of the top KRT18$^{+ve}$ markers) and HLA-DQA1 (one of the top KRT18$^{+ve}$ cell surface markers) were not chosen for validation due to their widespread distribution in several tissues, therefore rendering these molecules inadequate as notochordal markers.

Due to depletion of RNA from KRT18$^{+ve}$ and KRT18$^{-ve}$ cells, gene expression of KRT18 markers was analysed in two different groups of samples: 1) using RNA from all CD24$^{+ve}$ and CD24$^{-ve}$ cells isolated using FACS (Table 5.1) and 2) using RNA from notochordal and sclerotomal tissue microdissected from a larger (18 WPC) foetal sample (Table 5.1).
5.4.1.1 Expression of KRT18^{+ve} and KRT18^{−ve} markers in CD24^{+ve} and CD24^{−ve} cells isolated from developing foetal spines

The expression of the KRT18^{+ve} markers GRB14, SLC19A1 and FGF10 was significantly higher in CD24^{+ve} than in CD24^{−ve} cells confirming these genes as notochordal markers (Figure 5.2 A). The expression of the KRT18^{−ve} marker NOL12 was significantly lower in CD24^{+ve} than in CD24^{−ve} cells, confirming this marker as a sclerotomal marker. However, C5AR1 had significantly higher gene expression in CD24^{+ve} than in CD24^{−ve} cells and the expression of SERPINB6 was not significantly differentially expressed between CD24^{+ve} and CD24^{−ve} cells, and therefore qRT-PCR did not validate the microarray results for these two genes (Figure 5.2 B).

**Figure 5.2** Relative gene expression of KRT18^{+ve} and KRT18^{−ve} markers in CD24^{+ve} and CD24^{−ve} cells isolated from human embryonic and foetal developing spine cells. (A) KRT18^{+ve} markers; (B) KRT18^{−ve} markers. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in CD24^{−ve} cells and presented on a log scale. Error bars represent the standard error of the mean of the samples analysed. ** represents p<0.01; *** represents p<0.001.
The KRT18\textsuperscript{+ve} cell surface markers ADORA3, TBXA2R, CLDN12, CDH6, ANPEP, CD69 and STAB1 all had a significantly higher expression in CD24\textsuperscript{+ve} than in CD24\textsuperscript{-ve} embryonic and foetal developing spine cells. Although there was a trend towards a higher expression of SELE and a lower expression of EPHA5 in CD24\textsuperscript{+ve} than in CD24\textsuperscript{-ve} cells, these differences did not reach statistical significance (Figure 5.3).

![Figure 5.3](image_url)

**Figure 5.3** Relative gene expression of KRT18\textsuperscript{+ve} cell surface markers in CD24\textsuperscript{+ve} and CD24\textsuperscript{-ve} cells isolated from human embryonic and foetal developing spine cells. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in CD24\textsuperscript{-ve} cells and presented on a log scale. Error bars represent the standard error of the mean of the samples analysed. * represents $p<0.05$; ** represents $p<0.01$; *** represents $p<0.001$.

### 5.4.1.1.2 Expression of KRT18\textsuperscript{+ve} and KRT18\textsuperscript{-ve} markers in notochordal and sclerotomal tissues microdissected from an 18 WPC sample

To confirm that at 18WPC, the developing NP was still fully formed by notochordal cells, H&E was used to stain axial IVD sections. It was found that, at this stage, the central IVD region was fully formed by large notochordal cells,
which were encircled by lamellar sclerotome-derived developing AF cells (Figure 5.4). The central notochordal NP region corresponded to the anatomical region from which notochordal tissue was dissected (as shown in Figure 5.1), which confirmed the suitability of the microdissection method to separate notochordal from sclerotomal cells at this developmental stage.

![Figure 5.4](image)

**Figure 5.4** H&E staining of an axial section through the IVD of specimen M784 (18 WPC). The central NP region is populated by large vacuolated notochordal cells, encircled by lamellar sclerotome-derived cells (AF anlagens).

However, since only one older sample was available for analysis no statistical analysis was possible. For this reason, it is not known if differences in gene expression were inherent to this specific sample or if the small differences seen here would be statistically significant had more samples been analysed.

The expression of the KRT18\(^{\text{+ve}}\) markers GRB14, SLC19A1 and FGF10 was higher in notochordal than in sclerotomal tissue microdissected from the 18 WPC foetal spine specimen (Figure 5.5 A). The expression of the KRT18\(^{-\text{ve}}\) marker NOL12 was higher in notochordal than in sclerotomal tissue while the expression of the KRT18\(^{-\text{ve}}\) marker SERPINB6 was higher in sclerotomal tissue. The
expression of C5AR1 did not differ between notochordal and sclerotomal tissue microdissected from the 18 WPC foetal spine (Figure 5.5 B).

**Figure 5.5** Relative gene expression of the KRT18^+ve and KRT18^−ve markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC). (A) KRT18^+ve markers. (B) KRT18^−ve markers. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in sclerotomal tissue and presented on a log scale. As RNA was extracted from only one specimen, no statistical analysis was performed. Error bars represent standard error of the mean of the technical repeats.

The analysis of the expression of the KRT18^+ve cell surface markers in notochordal and sclerotomal tissue microdissected from an 18 WPC sample showed that the markers CDH6, ANPEP and EPHA5 had higher expression in notochordal than in sclerotomal tissues and that CLDN12 had higher expression in sclerotomal than in notochordal tissue. TBX2R and CD69 had a slightly higher and STAB1 had a slightly lower expression in notochordal tissue compared to sclerotomal tissues. There was no observable difference in the expression of ADORA3 and SELE in notochordal and sclerotomal tissue microdissected from this sample (Figure 5.6).
5.4.1.2 Protein expression of the KRT18\textsuperscript{+ve} marker ADORA3 in embryonic and foetal IVD tissues

ADORA3 was the top cell surface marker identified in the KRT18 microarray analysis and qRT-PCR analysis confirmed that this marker was differentially expressed between CD24\textsuperscript{+ve} and CD24\textsuperscript{-ve} developing spine cells. For this reason, this marker was chosen for protein validation in a cohort of embryonic and foetal spine tissues ranging between 6-19 WPC.

ADORA3 was specifically expressed by notochordal cells at all stages analysed; this expression was localised to the cell membrane, reflecting the characteristic cell surface localisation of this marker (Figure 5.7 A-H panels 2); it was weaker in the earlier stages (7-10 WPC) (Figure 5.7 A-B, panels 2) and stronger between 10-18 WPC (Figure 5.7 C-H, panels 2). Sclerotomal cells in the developing VB

Figure 5.6 Relative gene expression of the KRT18\textsuperscript{+ve} cell surface markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC). Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in sclerotomal tissue and presented on a log scale. As RNA was extracted from only one specimen, no statistical analysis was performed. Error bars represent standard error of the mean of the technical repeats.
(Figure 5.7 A-B panels 1 and 2) and in the developing AF (Figure 5.7 A-B panels 2 and C-H panels 3) did not express this marker at any of the stages analysed.
Isolation and phenotypic characterisation of human notochordal cells
Ricardo Rodrigues Pinto
Figure 5.7 ADORA3 immunostaining of a cohort of developing spines showing notochord-specific expression (arrows) of this marker. No ADORA3 expression was seen in the surrounding sclerotomal cells in the developing AF and VB. A: M736 (8 WPC); B: M636 (9.5 WPC); C: M685 (10 WPC); D: M404 (11.5); E: M776 (13 WPC); F: M739 (14 WPC); G: M777 (17 WPC); H: M784 (18 WPC); I: Pancreas positive control; J: M742 (12 WPC) isotype control. For each age a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. ADORA3 was optimised using ADORA3 antibody on pancreas sections (+ve control) and with foetal spines stained with isotype control antibody (-ve control).
5.4.2 Validation of the markers identified in the CD24 microarray analysis

5.4.2.1 Gene expression of CD24⁺ve and CD24⁻ve markers

The top 10 CD24⁺ve and top 4 CD24⁻ve (bottom CD24⁺ve) differentially expressed genes identified in the CD24 microarray analysis were selected for validation with qRT-PCR. Gene expression was analysed in three different groups of samples: 1) using RNA from the same CD24⁺ve and CD24⁻ve samples used in the CD24 microarray analysis; 2) using RNA from CD24⁺ve and CD24⁻ve cells isolated from two additional foetal samples; 3) using RNA from notochordal and sclerotomal tissue microdissected from a larger (18 WPC) foetal sample.

5.4.2.1.1 Expression of CD24⁺ve and CD24⁻ve markers in sorted CD24⁺ve and CD24⁻ve cells isolated from the developing foetal spines used in the CD24 microarray analysis

Using the same samples used in the CD24 microarray analysis, qRT-PCR showed that all CD24⁺ve markers CD24, STMN2, RTN1, PRPH, CXCL12, IGF1, MAP1B, ISL1, CLDN1 and THBS2 had a significantly higher expression in CD24⁺ve than in CD24⁻ve cells (Figure 5.8 A). Similarly, the CD24⁻ve markers WISP3, CHST11, SERPINA3 and CHAD had a lower expression in CD24⁺ve than in CD24⁻ve cells; this difference was statistically significant for all markers except for CHST11, which did not reach statistical significance (Figure 5.8 B).
Figure 5.8 Relative gene expression of CD24\(^{+}\)ve and CD24\(^{-}\)ve markers in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from the same human embryonic and foetal specimens used in the CD24 microarray analysis. (A) CD24\(^{+}\)ve markers. (B) CD24\(^{-}\)ve markers. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in CD24\(^{-}\)ve cells and presented on a log scale. Error bars represent the standard error of the mean of the samples analysed. * represents p<0.05; ** represents p<0.01; *** represents p<0.001.

5.4.2.1.2 Expression of CD24\(^{+}\)ve and CD24\(^{-}\)ve markers in sorted CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from developing foetal spines from two additional specimens

To exclude that the concordant results between the CD24 microarray analysis and qRT-PCR could be inherent to the biology of the samples used in the microarrays, the expression of the same genes was also analysed in RNA from CD24\(^{+}\)ve and
CD24<sup>ve</sup> cells isolated using the same methodology from 2 additional specimens (M792 and M801) with similar ages (9 and 11 WPC, respectively) (Table 5.1).

Again, qRT-PCR showed that the top 10 CD24<sup>ve</sup> markers had significantly higher expression in the CD24<sup>ve</sup> cells compared to CD24<sup>ve</sup> cells (Figure 5.9 A) and that the top 4 CD24<sup>ve</sup> markers had significantly higher expression in the CD24<sup>ve</sup> cells than in the CD24<sup>ve</sup> cells (Figure 5.9 B); these differences were statistically significant for all genes.

**Figure 5.9** Relative gene expression of the CD24<sup>ve</sup> and CD24<sup>ve</sup> markers in CD24<sup>ve</sup> and CD24<sup>ve</sup> cells isolated from specimens M792 and M801. (A) CD24<sup>ve</sup> markers. (B) CD24<sup>ve</sup> markers. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in CD24<sup>ve</sup> cells and presented on a log scale. Error bars represent the standard error of the mean of the samples analysed. * represents p<0.05; ** represents p<0.01; *** represents p<0.001
5.4.2.1.3 Expression of CD24^+ve and CD24^-ve markers in notochordal and sclerotomal tissues microdissected from an 18 WPC sample

qRT-PCR analysis of the CD24^+ve markers in notochordal and sclerotomal tissues microdissected from an 18 WPC foetal sample showed that CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 had higher expression in notochordal than in sclerotomal tissue. STMN2, CXCL12 and IGF1 had higher expression in sclerotomal than in notochordal tissue (Figure 5.10 A). The CD24^-ve markers WISP3, CHST11 and SERPINA3 had higher expression in sclerotomal than in notochordal tissue, while CHAD had a lower expression in sclerotomal than in notochordal tissue (Figure 5.10B).
Figure 5.10 Relative gene expression of the CD24\(^{+}\)ve and CD24\(^{-}\)ve markers in notochordal and sclerotomal tissue microdissected from sample M784 (18 WPC). (A) CD24\(^{+}\)ve markers (B) CD24\(^{-}\)ve markers. Gene expression was normalised to the housekeeping gene GAPDH and then to the expression in sclerotomal tissue and presented on a log scale. As RNA was extracted from only one specimen, no statistical analysis was performed. Error bars represent standard error of the mean of the technical repeats.

5.4.2.2 Protein expression of CD24\(^{+}\)ve (notochordal) markers in embryonic and foetal IVD tissues

CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 were validated as notochordal markers using qRT-PCR in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from the same samples used in the CD24 microarray analysis, in CD24\(^{+}\)ve and CD24\(^{-}\)ve cells isolated from 2 additional foetal samples and in notochordal and sclerotomal tissues microdissected from an 18WPC sample. From these, RTN1, PRPH and
MAP1B were empirically chosen for protein validation in a cohort of embryonic and foetal spine tissues ranging between 6-19 WPC.

In the earlier developmental stages (6-12 WPC) RTN1 immunopositivity was seen in the cytoplasm of all notochordal cells (Figure 5.11 A-E panels 2); at those stages, the intensity of notochordal staining was weak and sclerotomal cells (AF or VB anlagens) did not express this protein (Figure 5.11 A-E panels 2 and D-E panels 3). Between 13-19 WPC, and with the exception of the 17 WPC sample, in which no staining was identified in any developing spine cell (Figure 5.11 G panels 2 and 3), RTN1 became co-expressed by notochordal and sclerotomal cells surrounding the developing NP (Figure 5.11 F-I panels 2). The staining was cytoplasmic in all cases. Between 13-14 WPC notochordal and sclerotomal staining was more intense than in previous stages (Figure 5.11 F-G panels 2), but the staining of both cell types was weaker again in the 19 WPC (Figure 5.11 H-I panels 2). Between 13-19 WPC sclerotomal cells in the AF region did not express RTN1 (Figure 5.11 F-I panels 3).
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 5.11. RTN1 immunostaining of a cohort of developing spines. RTN1 was notochord-specific (arrow) until 10 WPC, after which it became co-expressed by sclerotomal cells surrounding the central developing NP (*), with the exception of the 17 WPC sample in which no staining was identified in any developing spine cell: A: M741 (7 WPC); B: M736 (8 WPC); C: M636 (9.5 WPC); D: M439 (12 WPC); E: M776 (13 WPC); F: M739 (14 WPC); G: M777 (17 WPC); H: M809 (19 WPC); I: Pancreas positive control; J: Pancreas isotype control. For each age a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. RTN1 was optimised using pancreas sections stained with RTN1 antibody (positive control) and isotype control antibody (negative control).
PRPH was not expressed by any embryonic or foetal spine cell, notochordal (Figure 5.12 A-I panels 2) or sclerotomal (Figure 5.12 A-I panels 2 and 3). There was, however, consistently expression of PRPH in the cytoplasm of the spinal nerve roots neighbouring the developing spine (Figure 5.12 A-F panels 1).
Isolation and phenotypic characterisation of human notochordal cells
RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 5.12. PRPH immunostaining of a cohort of developing spines. PRPH was not confirmed as a notochord-specific marker at the protein level in the developing spines. No spine cells (notochordal (straight arrows) or sclerotomal) expressed this marker. However, in the periphery of the developing spine, nerve roots were positive for this marker. 

- A: M741 (7 WPC); B: M736 (8 WPC); C: M636 (9.5 WPC); D: M415 (10.5 WPC); E: M742 (12 WPC); F: M394 (12.5 WPC); G: M739 (14 WPC); H: M777 (17 WPC); I: M809 (19 WPC); J: Pancreas positive control; K: Pancreas isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. PRPH was optimised using PRPH antibody on colon sections (positive control) and isotype control antibody on colon sections (negative control).
MAP1B expression was localised to the cytoplasm of all notochordal cells at all stages analysed (Figure 5.13 A-I panels 2). While this expression was notochord-specific between 6-8 WPC (Figure 5.13 A-B panels 2) and 11-13 WPC (Figure 5.13 E-F panels 2), it was co-expressed by sclerotomal cells in the developing VB in the stages between 9-11 WPC (Figure 5.13 C-E panels 3) and 14-19 WPC (Figure 5.13 G-I panels 3). MAP1B staining was not localised to the sclerotomal cells in the developing AF in any of the stages analysed (Figure 5.13 A-C panels 2 and D-I panels 3).
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto
Figure 5.13. MAP1B immunostaining of a cohort of developing spines showing notochord-specific expression (arrows) of this marker between 6-8 and 11-13 WPC. Between 9-11 and 14-19 WPC it was co-expressed by the sclerotomal cells in the developing VB. A: M741 (7 WPC); B: M736 (8 WPC); C: M483 (10 WPC); D: M685 (10 WPC); E: M742 (12 WPC); F: M776 (13 WPC); G: M739 (14 WPC); H: M777 (17 WPC); I: M809 (19 WPC); J: Pancreas positive control; K: M404 (11.5 WPC) isotype control. For each stage a lower magnification depicting at least one spinal segment, with the developing notochordal NP and the surrounding sclerotome (panels 1) is shown. For each panel 1, one or two higher magnifications are highlighted: panels 2 are centred to the developing notochordal NP and panels 3 are centred to the developing sclerotomal AF. MAP1B was optimised using MAP1B antibody on pancreas sections (+ve control) and isotype control antibody in foetal spine sections (-ve control).
5.4.7 Expression of notochordal KRT18<sup>+</sup> markers in adult NP samples

5.4.7.1 Gene expression of notochordal KRT18<sup>+</sup> markers in adult NP samples

Adult NP samples from 17 patients undergoing spinal surgery for disc degeneration-related conditions were analysed for the expression of the top KRT18<sup>+</sup> marker GRB14 and for the KRT18<sup>+</sup> cell surface markers ADORA3, CDH6, ANPEP and CD69. All genes were expressed at various levels in the NP of the cohort analysed. ANPEP was the gene with the highest expression (relative to the housekeeping gene MRPL19) and CD69 the one with the lowest expression. CDH6 and CD69 had the highest variability between samples (Figure 5.14).

![Figure 5.14](image-url) Box and whisker graph depicting the gene expression of GRB14, ADORA3, CDH6, ANPEP and CD69 in the adult NP. Gene expression values were normalised to the housekeeping gene MRPL19 and then plotted on a log scale.
The expression of each gene was then correlated with the histological degeneration grade of each NP sample. The expression of GRB14 and ADORA3 had a moderate negative correlation, the expression of CDH6 and ANPEP had a weak negative correlation and the expression of CD69 had a very weak positive correlation with degenerative grade (Figure 5.15 and Table 5.10). These correlations, however, failed to reach statistical significance (Table 5.10).
Figure 5.15 Scatter plots depicting the correlation between the gene expression of GRB14, ADORA3, CDH6, ANPEP and CD69 in the adult NP and the degeneration grade. Gene expression values were normalised to the housekeeping gene MRPL19 and plotted on a log scale.

Table 5.10 Spearman r values for the correlation between the adult NP expression of GRB14, ADORA3, CDH6, ANPEP and CD69 and the degeneration grade.

<table>
<thead>
<tr>
<th>Gene</th>
<th>GRB14</th>
<th>ADORA3</th>
<th>CDH6</th>
<th>ANPEP</th>
<th>CD69</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman r</td>
<td>-0.4227</td>
<td>-0.4141</td>
<td>-0.2659</td>
<td>-0.3187</td>
<td>0.1195</td>
</tr>
<tr>
<td>p value</td>
<td>0.0909</td>
<td>0.0984</td>
<td>0.3023</td>
<td>0.2125</td>
<td>0.6479</td>
</tr>
</tbody>
</table>
5.4.7.2 Protein expression of notochordal KRT18 markers in adult NP samples

The expression of the top differentially expressed cell surface marker ADORA3 was analysed by immunohistochemistry in 11 adult NP samples with degeneration grades ranging from 2-12 (Table 5.11). The NP of all samples was composed of small round cells. NP cells were found as single and as organised cell clusters. All samples but two contained ADORA3\textsuperscript{+ve} cells in the tissue fragment analysed. There was no consistent pattern of ADORA3 staining with single positive and negative cells and clustered positive and negative cells being found in most samples, independently of degeneration grade (Table 5.11 and Figure 5.16).

<table>
<thead>
<tr>
<th>ID</th>
<th>Grade</th>
<th>Positive cells</th>
<th>Morphology of positive cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH0520</td>
<td>2</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0512</td>
<td>3</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0268</td>
<td>4</td>
<td>Absent</td>
<td>--</td>
</tr>
<tr>
<td>HH0269</td>
<td>5</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0422</td>
<td>6</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0418</td>
<td>7</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0422</td>
<td>8</td>
<td>Absent</td>
<td>--</td>
</tr>
<tr>
<td>HH0440</td>
<td>9</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0408</td>
<td>10</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0593</td>
<td>11</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0386</td>
<td>12</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
</tbody>
</table>
Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto

Grade 3

Grade 7
Figure 5.16. ADORA3 immunostaining of adult NP tissues. (A): sample HH0512 grade 3; (B): sample HH0418 grade 7; (C): sample HH593 grade 11; (D): sample HH0386 grade 12. For each grade, a lower magnification (panels 1) and two higher magnifications corresponding to the regions depicted in panels 1 (panels 2 and 3) are presented. Adult NP cells were round, non-vacuolated and scarce, with some cells being found as single cells (A2, B2, B3, C3, D2 and D3) and others being organised in clusters (A3 and C2). ADORA3⁺ve (panels 2) and ADORA3⁻ve (panels 3) cells were found in all samples depicted here, independently of their degeneration state and in single and clustered cells.
5.4.8 Expression of notochordal CD24 markers in adult NP samples

5.4.8.1 Gene expression of notochordal CD24 markers in adult NP samples

Adult NP samples from 17 patients undergoing spinal surgery for disc degeneration-related conditions were analysed for the expression of the CD24+ve markers CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1. All genes were expressed at various levels in the NP of the cohort analysed. CD24, MAP1B and CLDN1 were the genes with the highest expression and ISL1 was the gene with the lowest expression. MAP1B was the gene with the highest variability between samples (Figure 5.17).

Figure 5.17 Box and whisker graph depicting the relative gene expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 in the adult NP. Gene expression values were normalised to the housekeeping gene MRPL19 and then plotted on a log scale.
The expression of each gene was then correlated with the histological degeneration grade of each NP sample. All genes analysed showed a negative correlation with degeneration grade. That correlation was highest for MAP1B, which showed a strong correlation. CLDN1 showed a moderate, CD24 and ISL1 a weak and PRPH and RTN1 a very weak negative correlation (Figure 5.18 and Table 5.12). This negative correlation was statistically significant for MAP1B and ISL1.
Figure 5.18 Scatter plots depicting the correlation between the expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 in the adult NP and the degeneration grade. Gene expression values were normalised to the housekeeping gene MRPL19 and plotted on a log scale.

Table 5.12 Spearman r values for the correlation between the adult NP expression of CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 and the degeneration grade.

<table>
<thead>
<tr>
<th>Gene</th>
<th>CD24</th>
<th>RTN1</th>
<th>PRPH</th>
<th>MAP1B</th>
<th>ISL1</th>
<th>CLDN1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman r</td>
<td>-0.3598</td>
<td>-0.05813</td>
<td>-0.1051</td>
<td>-0.7977</td>
<td>-0.3008</td>
<td>-0.4408</td>
</tr>
<tr>
<td>p value</td>
<td>0.156</td>
<td>0.8246</td>
<td>0.688</td>
<td>0.0023</td>
<td>0.0465</td>
<td>0.0766</td>
</tr>
</tbody>
</table>
5.4.8.2 Protein expression of notochordal CD24 markers in adult NP samples

The expression of MAP1B was analysed using immunohistochemistry in adult NP samples from 11 patients with different degrees of IVD degeneration (Table 5.13). NP cells were small and round and presented either as single cells or organised in clusters – this organisation was found in all samples, independently of degeneration grade. MAP1B positivity was found in all but one sample analysed. There was no consistent pattern of MAP1B staining with single positive and negative cells and clustered positive and negative cells being found in most samples, independently of degeneration grade (Table 5.13 and Figure 5.19).

<table>
<thead>
<tr>
<th>ID</th>
<th>Grade</th>
<th>Positive cells</th>
<th>Morphology of positive cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH0520</td>
<td>2</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0512</td>
<td>3</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0268</td>
<td>4</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0269</td>
<td>5</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0422</td>
<td>6</td>
<td>Absent</td>
<td>--</td>
</tr>
<tr>
<td>HH0418</td>
<td>7</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0422</td>
<td>8</td>
<td>Present</td>
<td>Single</td>
</tr>
<tr>
<td>HH0440</td>
<td>9</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0408</td>
<td>10</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0593</td>
<td>11</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
<tr>
<td>HH0386</td>
<td>12</td>
<td>Present</td>
<td>Single and clusters</td>
</tr>
</tbody>
</table>

Table 5.13 MAP1B staining of adult NP samples. MAP1B was expressed in all but one of the samples analysed. Its expression was identified in single and in clustered cells.
Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO
Figure 5.19. MAP1B immunostaining of adult NP tissues. (A): sample HH0268 grade 4; (B): sample HH418 grade 7; (C): sample HH408 grade 10; (D): sample HH0386 grade 12. For each grade, a lower magnification (panels 1) and two higher magnifications corresponding to the regions depicted in panels 1 (panels 2 and 3) are presented. Adult NP cells were round, non-vacuolated and scarce, with some cells being found as single cells (A2, A3, B2, B3, C3, D2 and D3) and others being organised in clusters (C2). MAP1B⁺ve (panels 2) and MAP1B⁻ve (panels 3) cells were found in all samples depicted here, independently of their degeneration state and in single and clustered cells.
Tables 5.14 and 5.15 summarise the results presented in this chapter.
Table 5.14 Summary of the validation results for the markers identified in the KRT18 microarray analysis.

<table>
<thead>
<tr>
<th></th>
<th>Embryonic and foetal</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gene CD24 sorted cells</td>
<td>Gene microdissected tissue</td>
</tr>
<tr>
<td>Notochordal markers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRB14</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>SLC19A1</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>FGF10</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>Sclerotomal markers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOL12</td>
<td>✔✔</td>
<td>✖</td>
</tr>
<tr>
<td>C5AR1</td>
<td>✖✖</td>
<td>✖</td>
</tr>
<tr>
<td>SERP1NB6</td>
<td>✖</td>
<td>✔</td>
</tr>
<tr>
<td>Cell surface markers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADORA3</td>
<td>✔✔</td>
<td>✖</td>
</tr>
<tr>
<td>TBXA2R</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>CLDN12</td>
<td>✔✔</td>
<td>✖</td>
</tr>
<tr>
<td>CDH6</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>ANPEP</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>SELE</td>
<td>✔</td>
<td>✖</td>
</tr>
<tr>
<td>CD69</td>
<td>✔✔</td>
<td>✔</td>
</tr>
<tr>
<td>EPHA5</td>
<td>✖</td>
<td>✔</td>
</tr>
<tr>
<td>STAB1</td>
<td>✔✔</td>
<td>✖</td>
</tr>
</tbody>
</table>

✔ validated as notochordal marker; ✔✔ validated as notochordal marker (difference statistically significant); ✖ not validated as notochordal marker; ✖✖ not validated as notochordal marker (difference statistically significant); ++ expressed NT: not tested; ✷ positive correlation with degeneration; ❌ negative correlation with degeneration.
### Table 5.15 Summary of the validation results for the markers identified in the CD24 microarray analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Embryonic and foetal</th>
<th>Adult</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CD24 sorted cells (microarray)</td>
<td>CD24 sorted cells (additional samples)</td>
</tr>
<tr>
<td>Notochordal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD24</td>
<td>✔✔</td>
<td>✔✔</td>
</tr>
<tr>
<td>STMN2</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>RTN1</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>PRPH</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>CXCL12</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>IGF1</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>MAP1B</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>ISL1</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>CLDN1</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>THBS2</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Sclerotomal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WISP3</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>CHST11</td>
<td>❌</td>
<td>✔</td>
</tr>
<tr>
<td>SERPINA3</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>CHAD</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

✔ validated as notochordal/ sclerotomal marker; ✔✔ validated as notochordal/ sclerotomal marker (difference statistically significant); ✖ not validated as notochordal/ sclerotomal marker; ✖✖ not validated as notochordal/ sclerotomal marker (difference statistically significant); ++ expressed NT: not tested; ▼ negative correlation with degeneration.
5.5 DISCUSSION

In the previous chapters embryonic and foetal notochordal markers were identified and KRT18 and CD24 were independently used to isolate developing human notochordal cells. The phenotype of those cells was then compared with that of sclerotomal cells using microarrays and used to identify notochordal cell markers and to investigate the mechanisms and pathways regulating notochordal cell development and function.

Microarrays are useful as a screening tool that may aid in the identification of candidate genes for future studies. They may, however, and as aforementioned, detect false positives, which may be due to random chance, sampling bias, cross-hybridisation or even experimental artefacts (Simon, 2008). For this reason, microarray results are often validated at the gene level using qRT-PCR (Minogue et al., 2010b, Minogue et al., 2010a, Morey et al., 2006, Power et al., 2011). This validation becomes even more necessary in the case of this investigation, since the methodologies used (enzymatic tissue digestion, cell labelling and sorting, with or without prior fixation and permeabilisation) may themselves impact on the gene expression of the analysed cells (Autengruber et al., 2012, Diez et al., 1999, Esser et al., 1995, Li et al., 2013).

Finally, validation of the identified genes at the protein level using techniques such as immunohistochemistry (Ip et al., 2015), ELISA (Italiani et al., 2014) and western blotting (Sudo et al., 2013) provides additional information on the function of the cell being analysed.
Therefore, in this chapter, the expression of the genes identified using the previous two cell isolation techniques (KRT18 and CD24 labelling and sorting) was analysed using qRT-PCR in notochordal and sclerotomal cells isolated using the same methodology, but importantly, also in notochordal and sclerotomal tissue that had not been submitted to enzymatic digestion, cell labelling and sorting.

To understand which of the validated genes are translated to proteins, their expression was analysed in a cohort of developing human spines using immunohistochemistry.

From the genes identified in the KRT18 microarray analysis, GRB14, SLC19A1 and FGF10 were validated as notochordal gene markers and TBXA2R, CDH6, ANPEP and CD69 were validated as cell surface notochordal gene markers. ADORA3, despite not being differentially expressed at the gene level between 18 WPC notochordal and sclerotomal tissue, was validated as a notochord-specific marker at the protein level in developing spines between 7-19 WPC.

From the genes identified in the CD24 microarray analysis, CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 were validated as notochordal gene markers and WISP3, CHST11 and SERPINA3 as sclerotomal gene markers. MAP1B was further validated as notochord-specific marker at the protein level in developing spines between 6-8 and 11-13 WPC, while RTN1 was notochord-specific between 6-12 WPC.

These markers constitute the human notochordal and sclerotomal phenotype and provide, for the first time, valuable information regarding the transcriptome and the proteome of human developing notochordal cells. The fact that the genes were
validated in cells isolated using different techniques adds strength to the data and eliminates possible bias that each isolation technique could individually introduce. However, the differential expression of some genes identified in the KRT18 microarray analysis was not seen in the RNA from sorted CD24\(^{+ve}\) and CD24\(^{-ve}\) cells. C5AR1 and SERPINB6, which were identified as sclerotomal markers had higher expression in CD24\(^{+ve}\) cells and EPHA5, which was identified as a notochordal cell surface marker showed a trend towards a higher expression in CD24\(^{+ve}\) cells. These differences may be explained by several reasons: 1) those genes may be false positives, not confirmed by qRT-PCR; 2) those genes may, in fact, have higher differential expression in sclerotomal than in notochordal cells at 9 WPC (age of the sample used in the KRT18 microarray analysis) but their expression in other ages may be different (the age of the CD24 samples used for validation varied between 7.5-14 WPC); 3) the differential expression may be inherent to the specific sample used in the microarray analysis; 4) differences in gene expression may be explained by differences in the methodology used (KRT18 sorted samples were fixed and permeabilised prior to labelling and sorting, whereas CD24 samples did not suffer this treatment).

When analysing the expression of the same panel of genes in notochordal and sclerotomal tissue from an 18WPC sample, the KRT18 sclerotomal marker NOL12 had a higher differential expression in notochordal tissue and C5AR1 was not differentially expressed between tissues. The KRT18 cell surface markers ADORA3, CLDN12, SELE and STAB1 also did not show the same pattern of differential expression. Differences between the expression of genes in CD24 sorted cells and in the 18WPC microdissected sample may be attributable to: 1) the different age of the samples, therefore implying that the genes that are
validated as markers in CD24$$^{\text{ve}}$$ samples and in the 18 WPC sample are those that, more likely, are not affected by the different microenvironment in the developing spine at different ages and may constitute a notochordal cell phenotype independently of specimen age; 2) the different techniques used to isolate cells: as discussed above, enzymatic tissue digestion, cell labelling (with or without prior fixation and permeabilisation) and cell sorting may affect gene expression; 3) these differences may be inherent to the specific samples: it is one limitation to this study the fact that only one older specimen was available for validation, therefore preventing any statistical analysis on these samples.

The genes identified in the CD24 microarray analysis all showed a similar differential expression in the same cells used in the CD24 microarray analysis and in the two additional CD24 sorted samples. This confirms that the thresholds used for gene detection allowed identification of those genes that were indeed differentially expressed in those samples. STMN2, CXCL12, IGF1 and THBS2 which were identified as notochordal markers and CHAD, which was identified as a sclerotomal marker, however, did not show the same pattern of differential expression between the notochordal and sclerotomal tissue from the 18 WPC sample. As aforementioned, these differences may be attributable to: 1) the different age of the samples, indicating that the expression of such genes changes between 14 and 18 WPC; 2) the different techniques used for cell/ tissue isolation.

To understand if notochord-derived cells persist in the adult NP, the expression of notochordal markers was analysed in the NP from adult patients undergoing surgery for disc degeneration. All the genes investigated were present to different degrees in the adult NP with ANPEP (KRT18 notochordal marker), CD24,
MAP1B and CLDN1 (CD24 notochordal markers) having the highest expression levels.

Immunohistochemical analysis of notochordal markers in adult NP samples, identified adult NP cells expressing notochordal cell markers (ADOR3/MAP1B) and adult NP cells lacking the expression of notochordal cell markers. These findings, although derived from the analysis of a limited number of samples confirm the presence of notochord-derived cells in the adult. Other studies have previously hypothesised that cells with a notochordal phenotype were present in the human adult NP. Such studies, however, used the markers KRT8, KRT18, KRT19, GAL3 and T (Minogue et al., 2010b, Tang et al., 2012, Weiler et al., 2010), as notochord-specific markers. While KRT8, KRT18, KRT19 were confirmed as being human notochord-specific (chapter 2), their expression in the adult NP was not investigated in this study, since they were not top differentially expressed genes in the microarray analyses. GAL3 and T, however, which are often referred to as notochord-specific markers were shown to be co-expressed by sclerotomal cells in the human developing spine and, therefore, are not useful as notochordal markers.

The question as to whether the adult NP is composed of sub-populations of cells remains, however, unanswered. While the fact that some cells express notochordal markers and others don’t may suggest that two sub-populations reside in the adult NP (one that is notochord-derived and other that is sclerotome-derived), two other hypotheses would also be possible: one, in which a single population of notochord-derived cells resides in the adult NP that, with degeneration, loses the expression of some notochordal markers and other in which, with degeneration, notochord-derived cells disappear to be replaced by sclerotome-derived cells.
These hypotheses are suggested by the negative correlation encountered between the expression of notochordal markers, particularly that of MAP1B, CLDN1, ISL1, GRB14 and ADORA3 and degeneration grade. These findings were, however found in a limited number of adult samples and, therefore, require validation in a larger cohort. Furthermore, it remains to be elucidated if the cells expressing ADORA3 are the same as those expressing MAP1B – this could be confirmed by performing dual staining for those proteins in adult NP sections.

To fully understand if different sub-populations of cells co-exist in the adult NP it would be important to, firstly, expand the number of adult NP tissues in which the analysis was performed, analyse the expression of other identified notochordal markers and perform dual or even multiple staining for notochordal markers. Secondly, it would be important to analyse the gene expression of the two putative sub-populations of cells. Being a cell surface marker, ADORA3 could be used to identify, isolate and extract RNA from ADORA3\(^{+ve}\) and ADORA3\(^{-ve}\) adult NP cells. Their gene expression could then be compared between each other and also with that of CD24\(^{+ve}\) (notochordal) and CD24\(^{-ve}\) (sclerotomal) developing spine cells. Tools such as PCA and hierarchical clustering could help to investigate a possible common genetic signature between them and an extensive analysis of the expressed genes would further allow a better understanding on the number of sub-populations of cells in the adult NP. These questions, however, were not the main purpose of this study and, therefore, remain to be clarified.
CHAPTER 6

Conclusions and future work
6.1 GENERAL DISCUSSION AND CONCLUSIONS

This is, to the author’s knowledge, the first study to isolate and phenotype human notochordal cells. Since in humans cells with distinguishable notochordal cell morphology are only present during embryonic and foetal development and in the immature stages, cells were isolated from human embryos and foetuses.

The first aim of this thesis was to characterise the tissue morphology of the human embryonic and foetal spine and determine its suitability as a source of human notochordal cells. Large vacuolated notochordal cells were found in all development stages analysed (3.5-18 WPC). While they were initially localised along the embryo longitudinal axis, they later became restricted to the central developing IVD, encircled by lamellar sclerotomal AF anlagen cells and having sclerotomal VB anlagens cells above and below. It was also found that, at least in the younger specimens analysed, and due to the small size of the developing notochord and its close proximity to the neighbouring sclerotomal cells, conventional microdissection techniques would not be able to accurately separate notochordal from sclerotomal cells without cross-contamination.

As such, the second aim of this thesis was to identify a notochord-specific marker that could be used to label and isolate (using FACS) notochordal cells from their adjacent sclerotomal cells. This would allow a separation of pure populations of cells that could then be used for phenotyping studies. KRT8, KRT18 and KRT19 were found to be notochord-specific within the developing spine at all stages analysed and CD24 was notochord-specific between 5.5-18 WPC. Other markers,
which have previously been associated with notochordal cell phenotype, such as GAL3, CD55, CTGF, BASP1 and T were found to also be co-expressed by AF and/or VB sclerotomal cells. CD90, E-Cad and Tie2 were not expressed by any developing spine cell at the stages analysed.

The third aim of this thesis was to characterise the phenotype of human notochordal cells using microarrays. Two methodologies were used: one in which KRT18 was used as a notochord-specific marker and the other in which CD24 was used – these overlapped in time.

To phenotype notochordal cells isolated using their KRT18 specificity a novel methodology was developed to allow for the extraction of high quality RNA from limited numbers of cells that had previously been fixed, permeabilised, labelled with an intracellular marker (anti-KRT18 antibody) and sorted. The development of this methodology involved the identification of the ideal cell and tissue handling techniques, fixation and permeabilisation agent, incubation time periods and centrifugation speeds. It was found that a mixture of 95% ethanol and 5% acetic acid was the fixation and permeabilisation agent that had the least detriment effect on RNA quality, integrity and concentration. The methodology and its development, which are detailed in chapter 3, constitute a novel technique in molecular biology, whose applications extend beyond the IVD research field. Due to the technical difficulties inherent to the development and application of this methodology, one of its main objectives was to identify a notochordal cell surface marker that could then be used to isolate notochordal cells without the need for prior cell fixation and permeabilisation.

Concurrently, and based on the immunohistochemistry findings, CD24 was also used to isolate and separate notochordal and sclerotomal cells using FACS. Being
a cell surface marker, and therefore, posing less technical challenges, this methodology allowed for the isolation of larger RNA quantities (which were later available to validate the genes identified in the microarray analysis of notochordal and sclerotomal cells isolated using both techniques) and also to isolate RNA from 7 different specimens.

Microarrays were utilised to characterise the phenotype of notochordal cells using the two aforementioned techniques and to identify notochord-specific markers, cell surface notochordal markers and pathways, networks, upstream regulators and downstream functions of notochordal cells. Microarrays were validated using qRT-PCR on RNA from notochordal and sclerotomal cells isolated using two different methodologies: sorted CD24^{+ve} and CD24^{-ve} cells from 7.5-14 WPC human developing spines and notochordal and sclerotomal tissues obtained from an older and larger specimen (18 WPC) and isolated using microdissection techniques. The utilisation of these two different methodologies for cell isolation adds robustness to the validation, since it eliminates possible biases inherent to one the techniques individually. To understand which genes were translated to protein, a panel of genes were further analysed using immunohistochemistry in developing human spines. Microarray validation was the fourth aim of this thesis.

The analysis of notochordal and sclerotomal cells isolated using KRT18 specificity identified and validated GRB14, SLC19A1 and FGF10 as notochord-specific markers and TBXA2R, CDH6, ANPEP and CD69 as notochord-specific cell surface markers. ADORA3 was the top cell surface marker identified in this microarray analysis and its differential expression was validated using qRT-PCR in notochordal and sclerotomal cells isolated based on their CD24 expression; while this gene was not differentially expressed in notochordal and sclerotomal
tissues microdissected from the 18 WPC specimen, the expression of the protein it encodes was found to be notochord-specific in a cohort of 7-19 WPC foetal spines and therefore we put forward ADORA3 as a notochord-specific cell surface marker.

The analysis of notochordal and sclerotomal cells isolated using CD24 specificity identified and validated CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 as notochord-specific markers and WISP3, CHST11 and SERPINA3 as sclerotome-specific markers. The expression of the proteins translated by RTN1 and MAP1B was found to be notochord-specific during certain developmental stages.

These markers constitute the human notochordal cell phenotype. Notochordal cells, or factors they produce, have been proposed to have an anabolic effect in the IVD and to protect NP cells from degeneration and cell death and, for that reason, they have been proposed as the ideal cells to repair/ regenerate the degenerate disc (Aguiar et al., 1999, Cappello et al., 2006, Cornejo et al., 2015, Erwin and Inman, 2006, Erwin et al., 2011, Erwin et al., 2009, Miyazaki et al., 2009, Purmessur et al., 2015, Purmessur et al., 2011). As such, several studies have attempted to differentiate and to assess differentiation of stem cells to notochordal cells (Acosta et al., 2011, Chen et al., 2013, Korecki et al., 2010, Liu et al., 2015, Liu et al., 2014, Potier et al., 2014, Purmessur et al., 2011). However, due to the lack of notochordal cell markers, markers from studies in animals that retain notochordal cells were used. There are, however, important interspecies phenotype variations; this had previously been shown for the adult NP (Minogue et al., 2010b, Minogue et al., 2010a, Rodrigues-Pinto et al., 2013) and is shown in this thesis in which several genes that had previously been proposed to be notochordal cell markers in animal models were found to, in humans, also be expressed by sclerotomal cells.
This first characterisation of the human notochordal cell phenotype provides a fundamental tool for tissue engineering and regenerative medicine – markers such as GRB14, SLC19A1, FGF10, ADORA3, TBXA2R, CDH6, ANPEP, CD69, CD24, RTN1, PRPH, MAP1B, ISL1 and CLDN1 should be used in future studies to assess the notochordal cell phenotype.

Confirming that, such as with the NP phenotype (Minogue et al., 2010b, Minogue et al., 2010a, Rodrigues-Pinto et al., 2013), the notochordal cell phenotype varies between different species, from all the human notochordal cell markers identified in this study only CD24, which has been identified as rat NP marker in a stage in which notochordal cells are still present (Fujita et al., 2005), had previously been associated with the notochordal cell phenotype.

Relevant to the study of notochordal cells, it is also important to understand the mechanisms through which they exert their anabolic and anti-catabolic effects and the pathways leading to notochordal cell differentiation. Hence, part of the third aim of this thesis was to investigate the pathways, networks, upstream regulators and downstream functions of notochordal cells. IPA was used to analyse the data from both microarray analyses and, interestingly, and despite the list of differentially expressed genes differed between both microarrays, two important physiological mechanisms that may have relevant roles in notochordal cell biology and function were highlighted by both methodologies: inhibition of inflammation and inhibition of vascularisation.

The anti-inflammatory role of notochordal cells was highlighted in the KRT18 microarray analysis by the identification of ADORA3 and CD69, which have been shown to be negative regulators of inflammation (Odashima et al., 2005, Radulovic and Niess, 2015), as KRT18+ve cell surface markers. In the CD24
microarray analysis, the pro-inflammatory cytokines TNF, IL-1β and IFNγ were found to be negative regulators of notochordal cells and IL-2, TNF, TNFSF13B, IL-5 and IL-20 were negative master regulators of notochordal cells. IL-1RA, the antagonist for the receptor of IL-1β, which has previously been proposed as a therapeutic agent for IVD degeneration (Hoyland et al., 2008, Le Maitre et al., 2007c) was found to be a positive upstream regulator and master regulator of notochordal cells.

The anti-angiogenic role of notochordal cells was highlighted mostly in the CD24 microarray analysis by the identification of THBS2, which has been shown to be important in the maintenance of the avascular state of the IVD (Gruber et al., 2006), as a notochordal cell marker, and by the fact that among the top functions predicted to be inhibited by notochordal cells were development of blood vessels, vasculogenesis and angiogenesis. Adding to this, CCN5, Noggin and EDN2, which are all involved in angiogenesis inhibition (Cornejo et al., 2015, Lake et al., 2003, Rattner et al., 2013), were found to be positive master regulators and VEGF was found to be a negative regulator of notochordal cells. In the KRT18 microarray analysis HGF was found to be a positive regulator of notochordal cells and to drive the upregulation of NRP-1, the receptor for Sem3A, which has been proposed to act as a barrier to neuronal ingrowth (Masuda et al., 2004, Tolofari et al., 2010) – an event that is usually accompanied by vascular ingrowth into the healthy IVD (Freemont et al., 1997, Freemont et al., 2002).

One of the most debated topics in notochordal cell research concerns the presence of notochord-derived cells in the adult NP. Because the adult NP is mostly composed of small non-vacuolated cells it is not known if notochordal cells undergo a morphological change to become small NP cells and if, by doing so,
they change their phenotype, if they die to be replaced by cells migrating from neighbouring tissues or even if both events occur. To investigate that, the fifth aim of this thesis was to investigate the presence of notochord-specific markers in NP tissue harvested from patients undergoing surgery for IVD degeneration-related conditions and to correlate notochordal gene expression with degeneration grade. While this investigation was performed in a limited number of tissues and using a limited number of notochordal markers, it was found that within the adult NP, some cells express the notochordal markers ADORA3 or MAP1B and some don’t. It was also found that the expression of notochordal markers was negatively correlated with the degeneration grade. These findings suggest that at least some notochord-derived cells are retained in the adult NP and that, with degeneration, the phenotype of adult NP cells may change by loosing the expression of such markers.

The results presented here have important implications for the regeneration and repair of the degenerate IVD. Firstly, the human notochordal cell phenotype and notochord-specific markers have been identified. This knowledge can now be used to assess progenitor cell differentiation to a notochordal cell phenotype. Secondly, since it provides a methodology (CD24 sorting) that can be used to isolate viable human notochordal cells, the function of human notochordal cells can now be investigated in vitro. This will allow an understanding as to how human notochordal cells interact with NP cells and whether they confer resistance to conditions, such as IL-1 or Fas-Ligand, that have been shown to induce degeneration or cell death in NP cells alone (Erwin et al., 2011, Hoyland et al., 2008). Additionally, culturing notochordal cells, will also allow to understand the mechanisms through which notochordal cells exert their function – it is possible
that certain factors or molecules they synthesise are responsible for such a
protective role and, in that case, the isolation of such molecules could
theoretically permit their usage for tissue engineering strategies, such as growth
factor delivery or injection. Finally, this study identified inhibition of
inflammation and of vascularity among the most relevant notochordal functions,
which may be involved in the notochordal cell role in the IVD. Strategies aimed at
inhibiting these two mechanisms should be investigated as their regulation can
potentially retard or even prevent IVD degeneration.
6.2 FUTURE WORK

The investigation in this thesis has identified and validated human notochord-specific markers, has highlighted the possible roles of notochordal cells in IVD biology and elucidated on the presence of notochord-derived cells in the adult NP. However, a number of questions were also raised by this investigation that warrant future research.

Firstly, while a panel of notochord and sclerotomal markers were selected for validation, it would be important to expand the number of samples (mainly of larger non-sorted microdissected specimens) to use for validation. Additionally, only a few notochordal markers were chosen for protein expression analysis. Future studies should investigate the expression of other notochordal markers in the developing spine. The inclusion of a larger number of samples, would also allow for the understanding of whether and how the expression of those markers changes within the developmental stages analysed, as this would elucidate as to which mechanisms are involved in notochordal cell biology at each developmental stage.

Secondly, the isolation of notochordal and sclerotomal cells using CD24 provides the opportunity to culture and study the biological function of human notochordal cells. Future studies should investigate the appropriate way to culture and expand human notochordal cells and to understand (using notochordal and small NP cell co-culture systems with or without catabolic factors, for example) if and how notochordal cells exert their protective role on NP cells. A challenging but
valuable research project would be to immortalise human notochordal cells – this would allow an endless resource for future studies.

While the findings in this thesis suggest that notochord-derived cells are retained in the adult NP, it is not known if the cells expressing the notochordal marker ADORA3 are the same expressing the notochordal marker MAP1B. Dual labelling of both markers would elucidate on this aspect. Furthermore, while the findings in this study suggest that the expression of notochordal cell markers is lost with degeneration, this investigation should be expanded to include a larger number of samples. To better understand the ontogeny of NP cells future studies should compare the phenotype of ADORA3⁺ve and ADORA3⁻ve adult NP cells between each other and with that of CD24⁺ve (notochordal) and CD24⁻ve (sclerotomal) developing spine cells as this would allow understanding if there is a common genetic signature between them. Assuming that ADORA3⁺ve adult NP cells have a phenotype that is closer to that of notochordal cells it would also be important to investigate whether ADORA3⁺ve adult NP cells (which could be isolated using FACS) behave differently in culture, show different resistance to catabolic factors such as IL-1, proliferate differently, or resist differently to induced apoptosis in comparison with ADORA3⁻ve cells.

Finally, this study has highlighted inhibition of inflammation and of vascularisation as two relevant biological roles of notochordal cells that may have implications for the understanding of the protective role notochordal cells have in the IVD and, therefore for the development of cell-based therapies for IVD degeneration. Future studies should focus on the role of HGF, IL-1RA and Noggin in protecting NP cells from degeneration and in their putative clinical applications.
CHAPTER 7

References


Isolation and phenotypic characterisation of human notochordal cells

Ricardo Rodrigues Pinto

410


PRATSINIS, H. & KLETSAS, D. 2007. PDGF, bFGF and IGF-I stimulate the proliferation of intervertebral disc cells in vitro via the activation of the ERK and Akt signaling pathways. Eur Spine J, 16, 1858-66.


Isolation and phenotypic characterisation of human notochordal cells

RICARDO RODRIGUES PINTO


RODRIGUES-PINTO, R., RICHARDSON, S. M. & HOYLAND, J. A. 2014. An understanding of intervertebral disc development, maturation and cell phenotype provides clues to direct cell-based tissue regeneration therapies for disc degeneration. Eur Spine J.


SEKI, S., KAWAGUCHI, Y., CHIBA, K., MIKAMI, Y., KIZAWA, H., OYA, T., MIO, F., MORI, M., MIYAMOTO, Y., MASUDA, I., TSUNODA, T., KAMATA, M., KUBO, T., TOYAMA, Y., KIMURA, T., NAKAMURA,


SHU, C., SMITH, S. S., LITTLE, C. B. & MELROSE, J. 2013. Comparative immunolocalisation of perlecan, heparan sulphate, fibroblast growth factor-18, and fibroblast growth factor receptor-3 and their prospective roles in chondrogenic and osteogenic development of the human foetal spine. Eur Spine J, 22, 1774-84.


---

*Isolation and phenotypic characterisation of human notochordal cells*

RICARDO RODRIGUES PINTO

428


