A Membranome-Centered Approach Defines Novel Biomarkers for Cellular Subtypes in the Intervertebral Disc

DOI:
10.1177/1947603518764260

Document Version
Accepted author manuscript

Cartilage

A membranome-centered approach defines novel biomarkers for cellular subtypes in the intervertebral disc

Journal:	*Cartilage*
Manuscript ID	CART-17-0076.R3
Manuscript Type:	Original Article
Date Submitted by the Author:	05-Feb-2018
Complete List of Authors:	Akker, Guus; Maastricht University Medical Centre, Department of Molecular Genetics; Maastricht University Medical Centre, Department of Orthopedic Surgery; Radboud University Medical Center, Department of Experimental Rheumatology Eijssen, Lars; Maastricht University Medical Centre, Department of Bioinformatics Richardson, Stephen; The University of Manchester, Centre For Regenerative Medicine, Institute of Inflammation and Repair Rhijn, Lodewijk; Maastricht University Medical Centre, Department of Orthopedic Surgery Hoyland, Judith; The University of Manchester, Centre For Regenerative Medicine, Institute of Inflammation and Repair Welting, Tim; Maastricht University Medical Centre, Department of Orthopedic Surgery Voncken, Willem; Maastricht University Medical Centre, Department of Molecular Genetics
Keywords:	Intervertebral disc < Diagnosis, Cell lines, Nucleus Pulposus, Annulus Fibrosus, Biomarkers < Diagnostics
Abstract:	Objective: Lack of specific marker-sets prohibits definition and functional distinction of cellular subtypes in the intervertebral disc (IVD), such as those from the annulus fibrosus (AF) and the nucleus pulposus (NP). Design: We recently generated immortalized cell lines from human NP and AF tissues; these comprise a set of functionally distinct clonal subtypes. Whole transcriptome analyses were performed of twelve phenotypically distinct clonal cell lines (4x NP-Responder, 4x NP-nonResponder, 2x AF-Sheet forming and 2x AF-nonSheet forming). Data sets were filtered for membrane-associated marker genes and compared to literature. Results: Comparison of our immortal cell lines to published primary NP, AF and Articular Chondrocytes (AC) transcriptome datasets revealed preservation of AF and NP phenotypes. NP-specific membrane-associated genes were defined by comparison to AF cells in both the primary dataset (46 genes) and immortal cell-lines (161 genes). Definition of AF specific membrane-associated genes yielded 125 primary AF cell and 92 immortal cell-line markers. Overlap between primary and immortal NP cells yielded high-confidence NP-specific marker genes for NP-R (CLDN11, TMEFF2, CA12, ANXA2, CD44) and NP-nR (EFNA1, NETO2, SLC2A1). Overlap

	between AF and immortal AF subtypes yielded specific markers for AF-S (COLEC12, LPAR1) and AF-nS: (CHIC1). Conclusions: The current study provides a reference platform for preclinical evaluation of novel membrane-associated cell type-specific markers in the IVD. Future research will focus on their biological relevance for IVD function in development, homeostasis and degenerate conditions.

Reviewer 1 raised no further issues.

Answers to Reviewer 2

1: we have added the full description of AC in the abstract: articular chondrocytes (AC).

2: in line with the recommendation of Reviewer 2 to add reference to differences between species with respect to the presence of NC cells, we have added the following to the indicated section:

"Rodents, in contrast to humans, retain notochordal cells (NC) throughout their life time (ref); for this reason published NP markers in e.g. rat studies include NC markers."

A membranome-centered approach defines novel biomarkers for cellular subtypes in the intervertebral disc

Guus GH van den Akker[1,2], Lars MT Eijssen[3], Stephen M Richardson[4], Lodewijk W van Rhijn[1], Judith A Hoyland[4], Tim JM Welting*[1], Jan Willem Voncken*[#2]

[1]Department of Orthopedic Surgery, [2]Department of Molecular Genetics, [3]Department of Bioinformatics, Maastricht University Medical Centre, Maastricht, the Netherlands; [4]Centre For Regenerative Medicine, Institute of Inflammation and Repair, University of Manchester, Manchester, United Kingdom.

[$] Current address: Department of Experimental Reumathology, Radboud University Medical Centre Nijmegen, the Netherlands

* Equal contribution research groups

Corresponding author: JWV: Universiteitssingel 50; 6229ER Maastricht, the Netherlands; tel: +31 43 3882919; fax: +31 43 3884574

van den Akker GGH, PhD	guus.vandenakker@radboudumc.nl
Eissen LMT, PhD	l.eijssen@maastrichtuniversity.nl
Richardson SM, PhD	s.richardson@manchester.ac.uk
van Rhijn LW, MD, PhD	lw.vanrhijn@maastrichtuniversity.nl
Hoyland JA, PhD	judith.a.hoyland@manchester.ac.uk
Welting TJM, PhD	t.welting@maastrichtuniversity.nl
Voncken JW, PhD, Eng	w.voncken@maastrichtuniversity.nl

Keywords: intervertebral disc, cell lines, nucleus pulposus, annulus fibrosus, biomarkers, membranome

Abstract

Objective: Lack of specific marker-sets prohibits definition and functional distinction of cellular subtypes in the intervertebral disc (IVD), such as those from the annulus fibrosus (AF) and the nucleus pulposus (NP).

Design: We recently generated immortalized cell lines from human NP and AF tissues; these comprise a set of functionally distinct clonal subtypes. Whole transcriptome analyses were performed of twelve phenotypically distinct clonal cell lines (4x NP-Responder, 4x NP-nonResponder, 2x AF-Sheet forming and 2x AF-nonSheet forming). Data sets were filtered for membrane-associated marker genes and compared to literature.

Results: Comparison of our immortal cell lines to published primary NP, AF and Articular Chondrocytes (AC) transcriptome datasets revealed preservation of AF and NP phenotypes. NP-specific membrane-associated genes were defined by comparison to AF cells in both the primary dataset (46 genes) and immortal cell-lines (161 genes). Definition of AF specific membrane-associated genes yielded 125 primary AF cell and 92 immortal cell-line markers. Overlap between primary and immortal NP cells yielded high-confidence NP-specific marker genes for NP-R (*CLDN11*, *TMEFF2*, *CA12*, *ANXA2*, *CD44*) and NP-nR (EFNA1, NETO2, SLC2A1). Overlap between AF and immortal AF subtypes yielded specific markers for AF-S (*COLEC12*, *LPAR1*) and AF-nS: (*CHIC1)*.

Conclusions: The current study provides a reference platform for preclinical evaluation of novel membrane-associated cell type-specific markers in the IVD. Future research will focus on their biological relevance for IVD function in development, homeostasis and degenerate conditions.

Introduction

Low back pain has an extensive socio-economic impact (1, 2). Degeneration of the intervertebral disc is strongly associated with low back pain and disc herniation (3). Severe disc degeneration is estimated to occur in 10% of the general population at 50 years of age, and up to 60% in 70 years old individuals (4). The intervertebral disc (IVD) degenerates more rapidly than other and surrounding musculoskeletal tissues (5). In particular lumbar IVDs are susceptible to damage and injuries, underscoring important unique anatomical/mechanical features that contribute to the etiology of disc degeneration. IVD physiology is unique in that it is subject to a number of specific stressors. It is the most hypoxic tissue naturally occurring in the human body. Its nutrient supply to the nucleus pulposus (NP) and the surrounding annulus fibrosus (AF) occurs exclusively through the cartilaginous endplates (CEP; Figure 1A, B) and vessels surrounding the AF (5, 6). In addition, the IVD is subject to constant heavy loading, predominantly by muscle pressure (7). The osmolarity of the NP is higher compared to articular cartilage (AC) thereby retaining tissue-associated water to maintain disc height (8). The AF confines the NP to its central position within the IVD; this mediates vertical pressure absorbance within the spine. In addition, the AF endows the spinal column with a degree of flexibility.

Degenerative disc disease (DDD) is characterized by reduced proteoglycan content, loss of structural integrity and height of the disc and CEP erosion (Figure 1C, D). The CEPs are thought to be the weak link in the context of mechanical compression of the IVD. Loss of CEP integrity allows the NP to bulge into the vertebral body; this loss of structural integrity within the IVD in turn results in overloading of the AF. Consequential wear and tear may lead to AF rupture and NP

protrusion (disc herniation). The exact nature of the processes that lead to irreversible disc degeneration are poorly understood (6). A number of cellular changes have been observed in the degenerate IVD: a decrease in cell numbers, the appearance of senescent cells and cell cluster formation (9, 10). Concomitant production of inflammatory mediators and catabolic enzymes lead to a decrease in proteoglycan content in the NP and loss of structural integrity of the AF (11, 12). Finally, post-translational modifications of ECM molecules (13-15) and alterations in nutrient supply have been reported to correlate with DDD (16, 17). Incomplete understanding of the underlying biology of functionally distinct cell populations in the IVD, however, hampers the identification of signaling pathways and factors important for IVD development, homeostasis and disease.

We recently generated immortal clonal NP and AF cell lines: we identified multiple recurrent, functionally distinct cellular NP and AF phenotypes from independent young (13-15 years old) donors (18-20). While these findings provided vital proof-of-concept that multiple relevant cellular functions can be distinguished in the NP and AF, a clear understanding of the contribution of these cells to IVD biology *in vivo* awaits further elucidation. Heterogeneity of cells in the NP was first described in the early 80s (21, 22). Importantly, differences in cell phenotypes and cellular composition in current human IVD literature are likely to contribute to large variation in reported NP markers in various species (Supplemental Table S1)(23-32). In addition, variations in age, gender or clinical stage of DDD inevitably introduce substantial variation in human datasets which limits understanding of these markers in healthy human disc biology (33-35). The availability of single cell-derived clonal NP and AF subpopulations provides an obvious advantage in the context of defining cellular heterogeneity in the IVD. Morphologically and functionally distinctive cell

types provide an excellent starting point for gene expression profile-based, comparative phenotypic analysis. Biological specificity of cell types within diverse tissues is determined by specialized proteins, such as membrane-associated proteins, which have structural roles, determine morphology and movement and mediate signaling from the microenvironment (e.g. cytokines, growth factors) and/or interaction with other cells and ECM components. A comprehensive analysis of molecules expressed at the cell surface of the different IVD cell types and an detailed understanding of underlying molecular signaling networks is expected to help identify specific cellular subtypes and to define cellular responses to their environment in development and disease. Definition of cancer cell *membranomes* have yielded novel druggable targets and biomarkers for tumor-staging (36-38), and provided important proof-of-principle that crucial advances can be made in translational research by using selective approaches to identify unique features of cells. To obtain an unbiased molecular description of our immortal NP and AF cellular phenotypes, we adapted this approach in the current study by combining unbiased comparative whole-transcriptomic analysis with a selective filtering strategy specifically tailored toward identification of novel membrane-associated markers for IVD cell subtypes. The herein outlined *membranomics* approach using our unique AF and NP models aims to generate a reference platform for comprehensive preclinical evaluation of novel cell type-specific markers in the IVD.

Methods

Human IVD tissue.

Human IVD tissue from young idiopathic scoliosis patients (age 8-15 years) was obtained as surplus material during scoliosis correction surgery (18, 20). Degenerate

and non-degenerate lumbar IVD tissue was obtained *post-mortem* from a 63 years old donor (19). Acquisition and use of all human materials was approved by the MUMC-Medical Ethical Review Committee (Approval ID 08-4-021, July 11, 2012) (18-20, 26).

Cell isolation, culturing; RNA extraction

The isolation of primary NP and AF cells, the immortalization procedures and the establishment and characterization of immortal NP and AF cell lines have been described in detail elsewhere (18, 20). Briefly, 12 immortal cell clones derived from one male donor (D5, 13 years old, male *spina bifida* patient) were used for RNA isolation and comparative transcriptome analysis. Culturing procedures are described in detail in the Supplementary Methods section. RNA isolation, quality assessment and real time PCR analysis were essentially carried out as published (18-20). Procedures are described in detail in the Supplementary Methods section.

Whole transcriptome expression datasets.

Raw datasets from published whole transcriptome expression measurements (Affymetrix U133 Plus 2.0 array) of primary human AF (3), NP (3) and AC (3) cell donors (age range: 46-60 years; n=9) were derived from a previous study (26). Processing of RNA samples, labeling, hybridization of twelve Illimuna humanHT-12v4 arrays, data extraction and analyses are described in detail in the Supplementary Methods section. Briefly, an arbitrary cut-off of fold change >2.0 up or down and an average expression above 100 or higher (mean fluorescence intensity) in the cell type of interest was used to define and select expressed genes. Modified t-tests were used to evaluate statistically significant differences between groups (Benjamini Hochberg FDR adjusted p value <0.05 was considered significant).

Results and Discussion

High-throughput Expression Phenotyping of Novel Immortal NP and AF Clones

We previously defined functionally distinct immortal NP and AF cell lines based on their (in)ability to respond to established chondrocyte-differentiation conditions (*cf.* (non)Responder NP clones: NP-R and NP-nR; (18, 20), and based on their (in)ability to process procollagens *in vitro* (non) Sheet-forming AF clones: AF-S and AF-nS (20). Four representative clones for both the NP-R and NP-nR cellular phenotypes were chosen for large scale expression analyses, in addition to two AF-S and two AF-nS clones (Supplemental Figure S1A). We first evaluated expression of previously reported marker-sets for NP-R cells (*FOXF1, CA12; (18)*, NP-nR (*CD24*) and AF cells (*SFRP2, COL12A1; (20)*) prior to whole array expression analyses. This initial marker analysis confirmed that *FOXF1* and *CA12* discriminated between NP-R, NP-nR and AF cell clones (Supplemental Figure S1B). Expression levels of *COL12A1* and *SFRP2* were higher in AF and NP-nR cell clones compared to NP-R (Supplemental Figure S1B). NP-nR and AF cell clones showed similar levels of *SFRP2* and *COL12A1* mRNA. *CD24* levels, however, were significantly higher in NP-nR compared to NP-R or AF cell clones. This initial combined distinctive marker expression profiling of clonal IVD cell subtypes provided a solid basis for comparative genome-wide transcriptomic analysis.

To determine the degree of transcriptome preservation of our immortalized AF and NP cell clones, we compared our immortal array-dataset with published array datasets of primary isolates, comparing three donor-matched AC, AF and NP primary cell isolates (26). Principal Component Analysis (PCA) on the normalized datasets produced a high-resolution clustering overview. The first component strongly

separated the two datasets, which were generated on different array platforms (see: Supplemental Methods section). The second and third components clearly separated AC cells from all IVD cells, in line with their divergent tissues of origin (Figure 2A). Notably, primary AF and immortal clonal AF cell lines, all from unrelated donors, were presented as a distinctive cluster based on gene expression profiling. Although primary NP isolates showed considerable variation (three independent donors), the immortal NP-R and NP-nR clones all clustered close to the primary NP isolates. Hierarchical clustering of immortal clones resulted in the anticipated separation of the PCA clusters (Figure 2B; Supplemental Figure S2). In line with their earlier observed overlap in marker expression of *SFRP2* and *COL12A1*, NP-nR clones clustered closer to the AF clones than the NP-R group. Thus, PCA analyses indicated significantly distinctive gene expression profiles among immortal clonal IVD cellular subtypes and supported considerable preservation of primary AF and NP transcriptomic phenotype.

To obtain insight into genetic network activity underlying the preservation of immortal phenotypes in the NP, we first generated a list of putative NP, AC and AF markers in rat, canine and human IVDs. Rodents, in contrast to humans, retain notochordal cells (NC) throughout their life time (39); for this reason published NP markers in e.g. rat studies include NC markers. These putative markers were based on whole transcriptome expression studies performed on freshly isolated cells (23-26). We identified a total of 37 unique markers that showed differential expression between NP and AF, NP and AC or NP and notochordal cells in one or more of these studies (Supplemental Table S1). Twenty-five positive NP markers (showing no overlap with AF or AC specific genes) were selected to build a *literature-based NP gene-network*

(Figure 3A; Supplemental Table S2); this network was then used to visualize clone-specific differences in gene expression between NP-R and NP-nR clones. NP-R cells showed relatively higher expression of *CA12, CLEC2B, DSC2, KRT19, FOXF1* and *PTN,* while NP-nR cells expressed *A2M, SNAP25, SOSTDC1* and *VCAN* (Figure 3A). The differential expression of *FOXF1, CA12, KRT19* and *PTN* is consistent with our previous qPCR-based analyses (18). *A2M, CLEC2B*, *DSC2*, *SNAP25* and *SOSTDC1* potentially provide additional functional definition of NP cell subtypes. The NP-nR specific Alpha 2 Macroglobullin (A2M) is a general protease inhibitor which functionally interacts with important cytokines like bFGF, PDGF, NGF, IL1β and IL6 that are thought to be involved in DDD (40-42). The C-type lectin domain family 2, member B (CLEC2B) prevents blood vessel formation by endothelial cells and may fulfill a similar function in the NP (43). Desmocollin 2 (DSC2) is an ECM glycoprotein and component of the rat NP matrix (23). Synaptosomal-associated protein 25 (SNAP25) is involved in intracellular membrane trafficking and has been extensively studied in the context of brain function (44). Genes with known functions in the brain, such as axon guidance, have been reported before as NP markers (45, 46). Finally, Sclerostin domain-containing protein 1 (SOSTDC1) is a BMP antagonist. SOSTDC1 deficiency was recently shown to promote fracture healing through expansion of periosteal mesenchymal stem cells (47).

To visualize relative gene activity in this set of 25 markers in primary NP cells, the published transcriptome dataset (26) was plotted within the same gene network. This approach revealed that NP markers *KRT18, FoxF1* and *VCAN* were relatively highly expressed in primary NP cells (Figure S3A). Comparative gene network analysis thus reveals a number of striking differences in activity profiles of previously proposed general NP markers between immortal NP cellular subtypes and primary isolates,

and emphasizes the importance of functional definition of cellular heterogeneity in the NP.

Progenitor Cells in the Mature NP

Progenitor or stem cells in the mature IVD were first isolated and studied in 1995 and have since been studied in both NP and AF of various species (48-54). Progenitor cells were first isolated from the IVD by fluorescence-activated cell sorting (FACS) using a discriminatory candidate cell surface marker-based approach (55). NP progenitor cells were isolated by FACS from mouse, human and recently cow intervertebral discs using a combination of membrane-associated Cluster of Differentiation 24 (CD24), Ganglioside G2 (GD2) and/or receptor tyrosine kinase, epithelial specific (Tie2) cell surface markers (56, 57). A summary of these reported findings shows little overlap in the markers deployed to identify progenitor cells in the mature IVD (Table 1).

In our immortal clonal NP cell lines, distinctive gene signatures were identified that correlated with morphological characteristics: wave-like NP-Responder (NP-R) clones expressed high levels of *FoxF1* and *CA12,* whereas cobble stone-shaped NP-non-Responder (NP-nR) clones expressed high levels of membrane-associated CD24 (18). NP-R clones exhibited spontaneous spheroid-forming capacity on specific culture substrates (Aggrecan (ACAN), Matrigel). This stemness-associated feature has previously been utilized to identify progenitor cells in multiple systems (58, 59). Hence, we hypothesized that the immortal NP-R clones harbor a progenitor-like phenotype; in contrast our analyses suggested NP-nR clones maintain a more differentiated phenotype (18-20).

To obtain additional support for this hypothesis, we focused on the "*cell heterogeneity and progenitor*" genes listed in Table 1 and used these to generate a *progenitor gene-network* representing the NP progenitor cell phenotype. This network was then used to visualize relative expression differences between NP-R and NP-nR clones (Figure 3B). Significantly higher expression of *ANGPT1, CD44, CD63, FLT1 (VEGFR1)*, *KIT, NT5E (CD73) and TEK* was observed in NP-R clones. FLT1 (VEGFR1) and ANGPT1 are known as important signaling molecules mediating survival of NP cells (56). In the same study TEK (Tie-2) was identified as an NP progenitor marker. The expression of these markers in NP-R clones strongly supports the progenitor-like nature of this cell type.

NP-nR clones were characterized by increased *CD24, JAG1*, *PROM1 (CD133)* and *THY1 (CD90)* expression. Primary NP cell isolates (mixed cell population) had high expression of *CD24, ENO2, ITGB1 (CD29), POU5F1 (OCT3/4)* and *VEGFA*, whereas *TEK1* and *ANGPT1* were expressed at relatively low levels (Supplemental Figure S3B). Of note, the expression level of progenitor cell marker genes in the primary NP cell dataset was expected to be low, as progenitor cells are not likely to represent a predominant phenotype in the mature human IVD.

Although we evaluated a significant proportion of these cell markers, it is currently not clear how the two distinctive cellular NP phenotypes, NP-R and NP-nR, can be positioned on an imaginary differentiation scheme that was proposed based on membrane expression of CD24, CD44, Tie2 and GD2 (Table 2) (56). Our findings fit this model to certain extent: our progenitor cells (NP-R) are indeed CD24 negative (CD24-) whereas the more differentiated NP-nR cells are CD24+. None of our immortal clones were Tie2+ however, which was proposed as a hallmark of NP progenitor cells. Although NP-R clones were initially negative for GD2, GD2-

expression could be induced using chondrogenic differentiation stimuli (18). In summary, NP-R clones displayed a gene signature compatible with NP progenitor characteristics.

Novel Cell Surface Markers in the NP

A previous study identified membrane-associated genes specific for the NP compared with AC, by using an *in silico* screening of differentially expressed genes for the presence of transmembrane domains (26). To test whether these results could be reproduced in an independent study, we compared 90 previously identified membrane-associated genes (27) with primary NP/AC dataset (26) (Figure 4A). Thirty-four overlapping NP membrane-associated marker genes (NP/AC) were identified between these two studies; the expression of these markers was at least 2 fold higher in primary NP cells in comparison to AC (Table 3). *CA12* remained the top hit (3 probes in the top 10; relative expression level: >20 fold). In addition expression of *ABCG1, SLC27A2, TMEM71, FLRT2, CLDN11 and DNER* was more than 10 fold higher in primary NP isolates. As these genes may define functional differences between NP cells and ACs, their potential involvement in IVD biology is briefly discussed below. ATP binding cassette subfamily G member 1 (ABCG1*)* is expressed in articular cartilage and down-regulated in osteoarthritis and has a regulatory function in cholesterol and phospholipid transport. It (60). Solute carrier family 27 member2 (SLC27A2) is a fatty acid transport protein. TMEM71 function is unknown; it is regulated by the candidate tumor suppressor gene inhibitor of growth 2 (ING2) (61). Fibronectin-like domain-containing leucine-rich transmembrane protein (FLRT2) interacts with fibronectin and has distinct expression patterns in mouse development that are linked to FGF signaling (62, 63). *Claudin 11 (CLDN11)* is a tight

junction protein originally identified in oligodendrocytes and is known to be expressed in a small cellular subpopulation in the human IVD (64, 65). Delta- and notch-like epidermal growth factor-related receptor (*DNER*) plays an important role in brain development (66). Taken together, 34 membrane-associated marker genes that show a relatively higher transcriptional activity in NP compared to AC were identified in two independent datasets, further emphasizing the unique phenotype of NP cells. With the exception of CA12 these markers have not been studied in the IVD.

Comparative Membranome Analysis of primary NP, AF and AC isolates

To identify additional NP- and AF-specific cell surface markers, we first analyzed the primary NP, AF and AC cell datasets (26) and filtered for differentially expressed genes (FC>2.0; adjusted p value <0.05; see: Methods section for additional expression criteria). This yielded 898 genes of which expression was significantly higher in NP vs AF and 226 genes that were more highly expressed in NP vs AC. From these two gene lists 143 NP-specific markers (significantly higher in NP compared to AF and AC) were extracted. Using published meta-analyses (36), 6134 membrane-associated proteins were defined. Of the 143 NP-specific markers, 46 were found to be cell membrane-associated (Figure 4B, C). This marker set definition in NP and AF cell-lines showed overlap with our previous analyses in published primary NP cells and AC datasets (*Cf.* Table 3 with Table 4): *CA12, CLDN11,* Heparan Sulfate 6-0-sulfotransferase 3 (HS6ST3), Neuropilin- and tolloid-like 2 (*NETO2*), Calcium Release-Activated Calcium Modulator 3 (*ORAI3, TMEM142C*) and Choline transporter-like protein 1 (*SLC44A1*) (Table 4; bold print). The latter 4 genes have thus far not been studied in the musculoskeletal system and appear highly specific for the NP in two independent human IVD cell studies.

NP Clone-specific Membrane-associated Markers

In parallel, we identified unique membrane-associated markers for the NP clonal subtypes. The immortal cell clone dataset was filtered using a published *membranome* dataset (36). Of 161 differentially expressed membrane-associated genes, 79 were more highly expressed in the NP-R and 82 in the NP-nR (Figure 4B; Supplemental Tables S3, S4). These genes and their possible relevance for the NP cell phenotype are discussed below.

NP Responder-specific Membrane-associated Markers

Among NP-R-specific membrane markers, three genes showed a larger than 10-fold change: 1) plasticity related gene 1 (LPPR4), 2) transient receptor potential cation channel, subfamily V, member 2 (TRPV2) and 3) protocadherin 10 (PCDH10). LPPR4 is known for its role in brain plasticity (67) and has not been studied in the context of the IVD or articular cartilage. TRPV2 is known to be expressed in chondrogenic high-density cultures of limb bud cells in chickens and mice and was recently associated with genetic skeletal disorders (68, 69). PCDH10 is a negative regulator of the WNT pathway (70). Platelet-Derived Growth Factor Receptor Alpha (PDGFRA) has a fold change of 8.2 (NP-R/NP-nR) and is required for the formation of chondrogenic mesoderm and interacts with CAV1, which inhibits its signaling (71, 72). Interestingly, the NP-marker PAX1 is a transcriptional regulator of PDGFRA (73). The PDGFRA ligand PDGF partakes in a complex autocrine feedback with its receptor and modulates the effect of TGFβ-stimulation (74). In addition, the PDGFR ligand PDGF-BB was recently shown to delay disc degeneration in a rabbit model (75). In contrast, PDGF was shown to potentiate the chondrocyte's response to IL1β

in rheumatoid arthritis (76). Thus, differences in a cell's ability to activate the *PAX1-PDGFRA-PDGF* signaling axis may be relevant to local inflammatory responses and contribute to disc degeneration. *CLDN11* (fold change 6.9 NP-R/nR) expression was recently inversely correlated with *SDC2* expression (77). Of note: NP-nR clones showed a higher expression of *SDC2* than CLDN11 expressing NP-R clones. RGM (repulsive guidance molecule) domain family member B (*RGMB*) was restricted to NP-R clones (fold change 5.5). RGMa and RGMb are both highly expressed in the developing chick notochord (78). RGMb functions as a co-receptor for BMP and acts as an antagonist of noggin, which in turn antagonizes BMP-signaling, and has an important role in the embryonic development of the NP (79). In the NP, *RGM* expression is thought to prevent nerve in-growth in the IVD (45, 80). As qPCR analysis independently confirmed the micro-array-based differences (Supplemental Figure 4A), these markers are likely to have potential with regards to progenitor isolation from the NP.

NP nonResponder-specific Membrane-associated Markers

Among the 82 NP-nR specific membrane genes, protein tyrosine phosphatase, receptor type, F (PTPRF), tumor necrosis factor (ligand) superfamily, member 7 (CD70), chromosome 13 open reading frame 13 (SHISA2), G protein-coupled receptor, family C, group 5, member C (GPRC5C), chemokine orphan receptor 1 (CXCR7) and prostaglandin F2 receptor negative regulator (PTGFRN) all more than 10 fold higher expression in NP-nR clones than in NP-R clones. Strikingly, multiple genes could be traced to the WNT signaling pathway. PTPRF shares an extracellular domain with NCAM and controls beta-catenin signaling (81). NCAM was previously suggested as a NP marker (23, 30). SHISA2 modulates both WNT and FGF signaling

in the developing chicken embryo (82); in *Xenopus* it promotes somatic maturation of precursor cells (83). Low density lipoprotein Receptor-related Protein, 5 (LRP5; fold change 5) is part of the canonical WNT signaling. Induced expression of LRP5 mediates cartilage destruction in ostheoarthritis (84, 85). PTGFRN is involved in cell migration by selective interaction with Collagen type I and Fibronectin. This effect is modulated by the expression of the tetraspanins CD9 and CD81 (86). Differential expression of these genes in NP-nR (*versus* NP-R and AF cells) was confirmed by qPCR (Supplemental Figure 4B).

Markers for NP Cellular Heterogeneity in vivo

To begin to verify the potential usefulness of the clone-specific marker sets to distinguish between NP cell subtypes, we lastly compared the set of 46 primary NP cell-specific membrane-associated genes to NP-R and NP-nR-specific membrane-associated markers (Figure 4C). A subset of 8 genes was identified both as primary NP-specific cell surface markers and as being differentially expressed between immortalized NP clonal cell types. NP-R cells were marked by *CLDN11, TMEFF2, CA12, ANXA2, CD44* expression; NP-nR by *EFNA1, NETO2* and *SLC2A1* (Table 5). TransMembrane protein with EGF-like and two Follistatin-like domains 2 (TMEFF2) regulates non-canonical BMP4/activin signaling and is involved in PI3K and RAS/ERK1/2 signalling (87). Annexin A2 (ANXA2) regulates membrane organization and trafficking including endo- and exocytosis (88). CD44 is the receptor for hyaluronan, known to be expressed by notochordal and mature NP cells (56, 89). Finally, Solute carrier family 2 member 1 (SLC2A1; also known as GLUT1) is a known NP marker (90) and was expressed at substantially higher levels in NP-nR

clones. Importantly, expression of *SLC2A1* is reported to be elevated in the context of human disc degeneration (91).

Cellular Heterogeneity in the AF

Cellular heterogeneity in the inner (iAF) and outer AF (oAF) has been suggested based on morphological differences: outer AF cells have an elongated appearance while iAF cells are more rounded and showed morphological similarities with NP cells. A recent study demonstrated differences in ECM expression after prolonged culturing of iAF and oAF cells on a polycarbonate urethane substrate: oAF cells expressed higher amounts of *COL1A1*, while iAF cells expressed more *COL2A1, ACAN* and *VCAN* (91). *COMP* and Glypican 3 (*GPC3)* have been reported, although not consistently, as AF cell markers *in vivo*, (Supplemental Table S1). GPC3 is a Heparan sulfate proteoglycan involved in WNT, BMP and FGF signaling. Relevantly, GPC3 is an important inhibitor of SHH signaling (92); as SHH is expressed in the embryonic notochord, GPC3 might be an important regulator of NP/AF boundary demarcation in IVD development. Secreted Frizzled Related Protein 2 (*SFRP2*) is an inhibitor of WNT signaling and was identified as an AF marker in primary bovine IVD cells (25), which we validated as an mRNA marker of human AF cells *in vitro (20)*. Finally, differential expression of mitochondrial genes was suggested to differentiate between senescent and non-senescent cells in the AF (93, 94). Overall, this limited gene list stresses the need for identification of additional AF-specific markers.

Novel AF Subclone-specific Membrane-associated Markers

To identify AF-specific membrane markers, we first aimed to define AF-specific genes by comparing primary cell isolates from AF and NP (26); this analysis yielded

1161 genes that showed a significantly higher expression in AF than in NP. Application of the *membranome*-filter to this gene collection produced a subset of 125 AF-specific genes encoding membrane-associated proteins (Figure 4D, Table 7). We have limited discussing the potential relevance of these markers in regards to AF cell type-specific biology to a selection of marker genes. Of interest, we identified two receptors from the canonical Wnt pathway (LRP4 and FZD2). LRP4 was recently shown to facilitate chondrogenic differentiation and articular cartilage homeostasis (95, 96). In addition, expression of the ligand *WNT16* was significantly higher in primary AF. WNT16 was shown to repress osteoclastogenesis and to protect articular cartilage in a murine model for osteoarthritis (97, 98). Combined with our earlier observation that the *SFRP2* is an AF marker *in vitro*, these results suggest that regulation of Wnt signaling separates AF cells from NP cells. Of note, none of these factors have thus far been investigated in the IVD. WNT signaling is known to be involved in IVD development, decreases in the adult mouse IVD and is re-activated in old IVDs.

To expose AF clone-specific markers, we next compared the two AF clone subtypes that were distinguished on their (in)ability to process procollagens *in vitro*: Sheet-forming (AF-S) and nonSheet-forming (AF-nS) AF clones (20). A total of 120 differentially expressed genes 92 remained after *membranome*-filtering (Figure 4D, Supplemental Table S5, S6). *CLDN11*, identified as an NP-R marker gene (*cf.* Supplemental Figure S2), has also been implicated as a polarity gene in human AF cells *in situ* (64). *CLDN11* was preferentially expressed in AF-S clones. Of interest, AF-nS clones showed conspicuous Notch activity, as *NOTCH3, JAG1* and downstream target gene *HES1* were more highly expressed in AF-nS cells compared

to AF-S cells (Supplemental Figure S5C). Notch/Jagged/Hes1 activity was shown to be differentially regulated in NP an AF in response to hypoxia (99). In addition, polymorphisms in the *NOTCH3* gene have been associated with lumbar disc herniation and Notch signaling has been implicated in degenerative disc disease (100, 101). The exact biological implication of increased Notch-signaling in AF cells is currently unknown. It has been suggested that Notch is important for maintaining proliferative capacity of disc cells (99). Our data add to these observations and for the first time identifies a cellular AF sub-phenotype in which Notch signaling has a distinctive function.

Lastly, to determine the validity of these clone-specific markers, we compared primary AF cell specific genes to AF subtype-specific markers (Figure 4E). This identified 3 genes that were expressed both in primary as well as immortal monoclonal AF cell extracts: *Cysteine-rich hydrophobic domain 1* (CHIC1), *Collectin sub-family member 12* (COLEC12) and *Lysophosphatidic acid receptor 1* (LPAR1; Table 7). CHIC1 (also known as AKAP13) is an anchoring protein involved in Protein Kinase A signaling. In addition to its involvement in cardiac hypertrophy, it is known to regulate osmotic responses and integration of inflammatory signals in T lymphocytes (102, 103). Little is known about COLEC12 in the context of IVD biology. LPAR1 was shown to play a role in synovial fibroblast activation during arthritis and plays a role in osteoclastogenesis (104, 105). Taken together, our analyses identify a substantial set of novel AF membrane-associated markers. Subsets among these markers, such as WNT receptors LRP4, LRP5 (NP-nR and AF marker), FZD2 and additional genes described above, discriminate between our

previously established immortal IVD cell types and may prove useful to functionally define cellular heterogeneity in the AF.

Summary and Perspectives

In this study we present a meta-analysis of molecular markers in relation to cellular NP and AF phenotypes in the IVD. We here provide independent evidence of whole-transcriptome-based clustering of our novel immortal NP and AF cell lines with their respective IVD-derived primary isolates; in addition, PCA analysis clearly separates the NP clonal subtypes from each other and from AF cellular subpopulations. Importantly, the data accrued confirms cellular heterogeneity in NP and AF cell populations and supports the notion that cellular phenotypes are epigenetically preserved during the immortalization process.

Our *membranome*-based transcriptomics analysis aimed to identify novel cell surface markers for NP and AF cellular (sub)types. PCA analysis and validation of membrane and non-membrane-associated marker expression using qPCR analysis, consistently reveal a closer similarity of NP-nR with AF clones than with NP-R clones. Besides being lineage specific, a cell's transcriptomic phenotype is also determined through interaction with its microenvironment (106). Hence, it is conceivable that despite their dissimilar ontogeny, original NP-nR and AF cells were derived from areas with a relatively similar microenvironment in their respective tissues of origin.

The novel membrane markers identified in this study will be crucial for the isolation, systematic analysis and functional characterization of NP cellular subtypes and their implication in the context of disc homeostasis and disease. Investigation of specific cell subtypes in the IVD requires histological studies focused on expression analysis of markers identified herein in IVD tissue sections to define their relative locations

and to understand their ontogenic relationships. The availability of a functionally defined set of NP and AF cells enables systematic analysis of altered activity within gene networks, for instance as a function of age, and is expected to significantly contribute to understanding gene-microenvironment interaction in the IVD. It is very likely that microenvironmental alterations (e.g. inflammation, oxygen tension, osmolarity) and accompanying physiological changes related to ageing or degeneration override physiological responses required to maintain cell and tissue homeostasis. Such knowledge will also be pivotal for defining appropriate conditions for isolation, culturing and maintaining cells from the NP and AF and for application of IVD-directed differentiation of stem cells in regenerative procedures.

Authors' contributions

GvdA, LMTE, LWvR, JAH, TJMW, JWV contributed to the conception and design of the study; GvdA, LMTE, SMR contributed to acquisition of data, analysis and interpretation of data; GvdA, LMTE, SMR, JAH, TJMW, JWV drafting and final approval of the submitted version.

Acknowledgements

We thank Dr. Paul Willems for providing adult human IVD tissue for histology. Approval for all experimental sections of the current study, and informed consent for publication of patient details and accompanying images in this manuscript was obtained as an integral part of the MUMC-MERC approval 08-4-021; the approval is held by the authors and is available for review. Funding: Project P2.01 IDiDAS (LWvR, TJMW); LLP14 Grant Dutch Arthritis Foundation (JWV, LWvR).

Declaration of conflicting interests

None of the authors had any conflict of interest. This research forms part of the Project P2.01 IDiDAS of the research program of the BioMedical Materials institute, co-funded by the Dutch Ministry of Economic Affairs, Agriculture and Innovation. None of the funding sources had any role in study design, data collection, analysis and interpretation, writing of the manuscript or the decision to submit the manuscript for publication.

Figure legends

Figure 1: Intervertebral disc morphology and degenerative disc disease. A) Schematic representation of an intervertebral disc. The IVD comprises three distinctive tissues: a central gelatinous nucleus pulposus (NP), surrounded by the lamellar annulus fibrosus (AF). The AF consists of multiple, concentrically arranged lamellae in which cells and ECM components interact to provide function and structure. In the AF two distinct zones are distinguished: the inner and outer AF, primarily based on their collagen/proteoglycan content. The NP and AF are flanked at each side by cartilaginous endplates (CEP) which mediate the contact between the IVD body and vertebra. B) Sections of a healthy AF (15 years old male donor) stained with Safranin O and Fast green. The outer AF predominantly contains collagens while the inner AF contains proteoglycans. The characteristic lamellar structure is clearly visible; bar represents 200 μm. C) A non-degenerate L1/L2 intervertebral disc of 63 years old male donor. The lamellar annulus fibrosus is clearly recognizable. The central nucleus pulposus displays no aberrant morphology; the cartilaginous endplate is relatively thin. D) The L4/L5 intervertebral disc of the same donor. Degenerative changes cause apparent loss of lamellar structure in the AF. The NP is no longer distinguishable from the AF and additional ossification is found near the CEP.

Figure 2: Principle component analysis reveals overlap between primary and immortalized NP and AF clones. Transcriptome comparison between immortal cell clone and primary cell datasets generated on Affymetrix and Illumina platforms (1st component), respectively. PCA analysis was preceded by ranking genes (descending

expression value; see: Supplemental Methods for detailed description). A) PCA plot of the 2nd and 3rd component of primary AC, NP and AF samples from three independent healthy donors (circles) and immortalized NP-Responder (n=4), NP non-Responder (n=3) and AF (n=3) cell clones from a single donor (squares); these clones have been described in previous reports (18, 20). The 1st component separated the dataset on platform (primary cells: *Affymetrix* platform, immortalized cell lines: *Illumina* platform). B) Dendrogram showing the hierarchical clustering of immortalized NP-Responder (n=4), NP non-Responder (n=4) and AF (n=4) cell clones; clustering was based on Pearson correlation. The NP-R, AF, NP-nR clusters clearly differ. NP-nR clone 102 fell outside the NP-nR cluster (Supplemental Figure S2). As this effect persisted in subsequent analyses, this clone was omitted from further analysis. Analyses were performed and figures generated using the open source scripting language R (version 2.13.0) and R packages of Bioconductor 2.8.

Figure 3: Differential expression of published NP markers among NP cellular subtypes. A) Selected marker genes (NP cell markers; Supplementary Table S1, S2) were used as input for *GeneMANIA* network building using *Cytoscape* visualization software. Gene expression differences between NP-R (red) and NP-nR (blue) clones are depicted as log2-based fold change (FC) in colored circles. Colored lines in the network represent established biological connections between markers; parentheses: percentage representation. The network is strongly inter-connected based on co-expression and genetic interactions studies. Of note: established NP markers are not equally expressed by isolated cellular sub-populations derived from the NP. B) Progenitor-associated gene expression in NP-Responder clones. Selected genes (NP progenitor markers; Table 1, Supplementary Table S2) were used as

input for network building using *GeneMANIA* and *Cytoscape* visualization software. Gene expression differences between NP-R (red) and NP-nR (blue) clones are depicted as log2-based fold change (FC) in colored circles. Colored lines in the network represent established biological connections between markers; parentheses: percentage representation.

Figure 4: Identification of membrane-associated NP marker genes. Comparison strategy of whole transcriptome datasets of published primary IVD cellular isolates and clonal cell lines. A) Marker genes significantly more highly expressed in primary NP compared to AC from two independent published datasets (26, 27), were compared. B) Definition of primary NP-specific membrane marker genes compared to AF and AC (26) and definition of NP cell clone (R/nR) specific marker genes (18). C) Subsequently both primary and clonal NP membrane marker gene sets were compared to each other. D) Definition of primary AF-specific membrane marker genes compared to NP and definition of AF cell clone (S/nS) specific marker genes. E) Subsequent comparison of primary and clonal AF membrane marker genes.

References

1. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA.* 2008;299:656-64
2. Maniadakis N, Gray A. The economic burden of back pain in the UK. *PAIN.* 2000;84:95-103
3. Luoma K, Riihimäki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A. Low Back Pain in Relation to Lumbar Disc Degeneration. *Spine.* 2000;25:487-92
4. Miller J, Schmatz C, Schultz A. Lumbar disc degeneration: correlation with age, sex, and spine level in 600 autopsy specimens. *Spine.* 1988;13:173-8
5. Urban J, Roberts S. Degeneration of the intervertebral disc. *Arthritis Res Ther.* 2003;5:120-30
6. Adams MA, Roughley PJ. What is intervertebral disc degeneration, and what causes it? *Spine.* 2006;31:2151-61
7. Adams MA, Dolan P, McNally DS. The internal mechanical functioning of intervertebral discs and articular cartilage, and its relevance to matrix biology. *Matrix Biology.* 2009;28:384-89
8. Mwale F, Roughley P, Antoniou1 J. Distinction between the extracellular matrix of the nucleus pulposus and hyaline cartilage: a requisite for tissue engineering of intervertebral disc. *Eurepean Cells & Materials.* 2004;8:58-64
9. Le Maitre C, Freemont A, Hoyland J. Accelerated cellular senescence in degenerate intervertebral discs: a possible role in the pathogenesis of intervertebral disc degeneration. *Arthritis Research & Therapy.* 2007;9:R45
10. Wang F, Cai F, Shi R, Wang XH, Wu XT. Aging and age related stresses: a senescence mechanism of intervertebral disc degeneration. *Osteoarthritis and Cartilage.* 2016;24:398-408
11. Risbud MV, Shapiro IM. Role of Cytokines in Intervertebral Disc Degeneration: Pain and Disc-content. *Nature reviews. Rheumatology.* 2014;10:44-56
12. Phillips KLE, Cullen K, Chiverton N, Michael ALR, Cole AA, Breakwell LM, et al. Potential roles of cytokines and chemokines in human intervertebral disc degeneration: interleukin-1 is a master regulator of catabolic processes. *Osteoarthritis and Cartilage.* 2015;23:1165-77
13. Liu H-F, Zhang H, Qiao G-X, Ning B, Hu Y-L, Wang D-C, et al. A novel rabbit disc degeneration model induced by fibronectin fragment. *Joint Bone Spine.* 2013;80:301-06
14. Sivan S, Hayes A, Wachtel E, Caterson B, Merkher Y, Maroudas A, et al. Biochemical composition and turnover of the extracellular matrix of the normal and degenerate intervertebral disc. *European Spine Journal.* 2014;23:344-53
15. Sivan SS, Wachtel E, Roughley P. Structure, function, aging and turnover of aggrecan in the intervertebral disc. *Biochimica et Biophysica Acta (BBA) - General Subjects.* 2014;1840:3181-89
16. Huang Y-C, Urban JPG, Luk KDK. Intervertebral disc regeneration: do nutrients lead the way? *Nature reviews. Rheumatology.* 2014;10:561-66
17. Wang D, Nasto LA, Roughley P, Leme AS, Houghton AM, Usas A, et al. Spine degeneration in a murine model of chronic human tobacco smokers. *Osteoarthritis and Cartilage.* 2012;20:896-905
18. van den Akker GGH, Surtel DAM, Cremers A, Rodrigues-Pinto R, Richardson SM, Hoyland JA, et al. Novel immortal human cell lines reveal subpopulations in the nucleus pulposus. *Arthritis Research & Therapy.* 2014;16:R135-R35
19. van den Akker GG, Surtel DA, Cremers A, Hoes MF, Caron MM, Richardson SM, et al. EGR1 controls divergent cellular responses of distinctive nucleus pulposus cell types. *BMC Musculoskelet Disord.* 2016;17:124
20. van den Akker GG, Surtel DA, Cremers A, Richardson SM, Hoyland JA, van Rhijn LW, et al. Novel Immortal Cell Lines Support Cellular Heterogeneity in the Human Annulus Fibrosus. *PLoS ONE.* 2016;11:e0144497
21. Trout JJ, Buckwalter JA, Moore KC. Ultrastructure of the human intervertebral disc: II. Cells of the nucleus pulposus. *The Anatomical Record.* 1982;204:307-14
22. Trout JJ, Buckwalter JA, Moore KC, Landas SK. Ultrastructure of the human intervertebral disc. I. Changes in notochordal cells with age. *Tissue and Cell.* 1982;14:359-69
23. Lee CR, Sakai D, Nakai T, Toyama K, Mochida J, Alini M, et al. A phenotypic comparison of intervertebral disc and articular cartilage cells in the rat. *European Spine Journal.* 2007;16:2174-85
24. Sakai D, Nakai T, Mochida J, Alini M, Grad S. Differential phenotype of intervertebral disc cells: microarray and immunohistochemical analysis of canine nucleus pulposus and anulus fibrosus. *Spine.* 2009;34:1448-56

25. Minogue BM, Richardson SM, Zeef LAH, Freemont AJ, Hoyland JA. Transcriptional profiling of bovine intervertebral disc cells: implications for identification of normal and degenerate human intervertebral disc cell phenotypes. *Arthritis Research & Therapy.* 2010b;12:R22-R22
26. Minogue BM, Richardson SM, Zeef LAH, Freemont AJ, Hoyland JA. Characterization of the human nucleus pulposus cell phenotype and evaluation of novel marker gene expression to define adult stem cell differentiation. *Arthritis & Rheumatism.* 2010a;62:3695-705
27. Power KA, Grad S, Rutges JPHJ, Creemers LB, van Rijen MHP, O'Gaora P, et al. Identification of cell surface–specific markers to target human nucleus pulposus cells: Expression of carbonic anhydrase XII varies with age and degeneration. *Arthritis & Rheumatism.* 2011;63:3876-86
28. Fujita N, Miyamoto T, Imai J-i, Hosogane N, Suzuki T, Yagi M, et al. CD24 is expressed specifically in the nucleus pulposus of intervertebral discs. *Biochemical and Biophysical Research Communications.* 2005;338:1890-96
29. Gilson A, Dreger M, Urban J. Differential expression level of cytokeratin 8 in cells of the bovine nucleus pulposus complicates the search for specific intervertebral disc cell markers. *Arthritis Research & Therapy.* 2010;12:R24
30. Rutges J, Creemers LB, Dhert W, Milz S, Sakai D, Mochida J, et al. Variations in gene and protein expression in human nucleus pulposus in comparison with annulus fibrosus and cartilage cells: potential associations with aging and degeneration. *Osteoarthritis and Cartilage.* 2009;18:416-23
31. Risbud MV, Schoepflin ZR, Mwale F, Kandel RA, Grad S, Iatridis JC, et al. Defining the Phenotype of Young Healthy Nucleus Pulposus Cells: Recommendations of the Spine Research Interest Group at the 2014 Annual ORS Meeting. *Journal of Orthopaedic Research.* 2015;33:283-93
32. Rodrigues-Pinto R, Berry A, Piper-Hanley K, Hanley N, Richardson SM, Hoyland JA. Spatiotemporal analysis of putative notochordal cell markers reveals CD24 and keratins 8, 18, and 19 as notochord-specific markers during early human intervertebral disc development. *Journal of Orthopaedic Research.* 2016;1327-40
33. Tang X, Jing L, Richardson WJ, Isaacs RE, Fitch RD, Brown CR, et al. Identifying molecular phenotype of nucleus pulposus cells in human intervertebral disc with aging and degeneration. *Journal of Orthopaedic Research.* 2016;34:1316-26
34. Thorpe AA, Binch AL, Creemers LB, Sammon C, Le Maitre CL. Nucleus pulposus phenotypic markers to determine stem cell differentiation: fact or fiction? *Oncotarget.* 2016;7:2189-200
35. Wáng YXJ, Wáng J-Q, Káplár Z. Increased low back pain prevalence in females than in males after menopause age: evidences based on synthetic literature review. *Quantitative Imaging in Medicine and Surgery.* 2016;6:199-206
36. Uva P, Lahm A, Sbardellati A, Grigoriadis A, Tutt A, de Rinaldis E. Comparative *Membranome* Expression Analysis in Primary Tumors and Derived Cell Lines. *PLoS ONE.* 2010;5:e11742
37. Rust S, Guillard S, Sachsenmeier K, Hay C, Davidson M, Karlsson A, et al. Combining phenotypic and proteomic approaches to identify membrane targets in a 'triple negative' breast cancer cell type. *Molecular Cancer.* 2013;12:11
38. Ghosh D, Beavis RC, Wilkins JA. The Identification and Characterization of Membranome Components. *Journal of Proteome Research.* 2008;7:1572-83
39. Alini M, Eisenstein SM, Ito K, Little C, Kettler AA, Masuda K, et al. Are animal models useful for studying human disc disorders/degeneration? *European Spine Journal.* 2008;17:2-19
40. Rehman AA, Ahsan H, Khan FH. Alpha-2-macroglobulin: A physiological guardian. *Journal of Cellular Physiology.* 2013;228:1665-75
41. LA Binch A, Cole A, Breakwell L, Michael A, Chiverton N, Cross A, et al. Expression and regulation of neurotrophic and angiogenic factors during human intervertebral disc degeneration. *Arthritis Research & Therapy.* 2014;16:416
42. Barcelona PF, Saragovi HU. A Pro-Nerve Growth Factor (proNGF) and NGF Binding Protein, α2-Macroglobulin, Differentially Regulates p75 and TrkA Receptors and Is Relevant to Neurodegeneration Ex Vivo and In Vivo. *Molecular and Cellular Biology.* 2015;35:3396-408
43. Osada M, Inoue O, Ding G, Shirai T, Ichise H, Hirayama K, et al. Platelet Activation Receptor CLEC-2 Regulates Blood/Lymphatic Vessel Separation by Inhibiting Proliferation, Migration, and Tube Formation of Lymphatic Endothelial Cells. *Journal of Biological Chemistry.* 2012;287:22241-52
44. McNew JA, Parlati F, Fukuda R, Johnston RJ, Paz K, Paumet F, et al. Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature.* 2000;407:153-59

45. Tolofari S, Richardson S, Freemont A, Hoyland J. Expression of semaphorin 3A and its receptors in the human intervertebral disc: potential role in regulating neural ingrowth in the degenerate intervertebral disc. *Arthritis Research & Therapy.* 2010;12:R1
46. Bu G, Hou S, Ren D, Wu Y, Shang W, Huang W. Increased Expression of Netrin-1 and Its Deleted in Colorectal Cancer Receptor in Human Diseased Lumbar Intervertebral Disc Compared With Autopsy Control. *Spine.* 2012;37:2074-81
47. Collette NM, Yee CS, Hum NR, Murugesh DK, Christiansen BA, Xie L, et al. Sostdc1 deficiency accelerates fracture healing by promoting the expansion of periosteal mesenchymal stem cells. *Bone.* 2016;88:20-30
48. Chelberg MK, Banks GM, Geiger DF, Oegema TR. Identification of heterogeneous cell populations in normal human intervertebral disc. *Journal of Anatomy.* 1995;186:43-53
49. Sharp CA, Roberts S, Evans H, Brown SJ. Disc cell clusters in pathological human intervertebral discs are associated with increased stress protein immunostaining. *European Spine Journal.* 2009;18:1587-94
50. Henriksson H, Thornemo M, Karlsson C, Hagg O, Junevik K, Lindahl A, et al. Identification of cell proliferation zones, progenitor cells and a potential stem cell niche in the intervertebral disc region: a study in four species. *Spine.* 2009;34:2278-87
51. Feng G, Yang X, Shang H, Marks IW, Shen FH, Katz A. Multipotential differentiation of human anulus fibrosus cells: an in vitro study. *Journal of Bone & Joint Surgery.* 2010;92:675-85
52. Liu C, Guo Q, Li J, Wang S, Wang Y, Li B, et al. Identification of Rabbit Annulus Fibrosus-Derived Stem Cells. *PLoS ONE.* 2014;9:e108239
53. Liu LT, Huang B, Li CQ, Zhuang Y, Wang J, Zhou Y. Characteristics of stem cells derived from the degenerated human intervertebral disc cartilage endplate. *PLoS ONE.* 2011;6:e26285
54. Erwin WM, Islam D, Eftekarpour E, Inman RD, Karim MZ, Fehlings MG. Intervertebral disc-derived stem cells: implications for regenerative medicine and neural repair. *Spine.* 2013;38:211-6
55. Risbud M, Guttapalli A, Tsai T, Lee J, Danielson K, Vaccaro A, et al. Evidence for Skeletal Progenitor Cells in the Degenerate Human Intervertebral Disc. *Spine.* 2007;32:2537-44
56. Sakai D, Nakamura Y, Nakai T, Mishima T, Kato S, Grad S, et al. Exhaustion of nucleus pulposus progenitor cells with ageing and degeneration of the intervertebral disc. *Nature Communications.* 2012;3:1264
57. Tekari A, Chan SCW, Sakai D, Grad S, Gantenbein B. Angiopoietin-1 receptor Tie2 distinguishes multipotent differentiation capability in bovine coccygeal nucleus pulposus cells. *Stem Cell Research & Therapy.* 2016;7:1-12
58. Cheng N-C, Wang S, Young T-H. The influence of spheroid formation of human adipose-derived stem cells on chitosan films on stemness and differentiation capabilities. *Biomaterials.* 2012;33:1748-58
59. Tesei A, Zoli W, Arienti C, Storci G, Granato AM, Pasquinelli G, et al. Isolation of stem/progenitor cells from normal lung tissue of adult humans. *Cell Proliferation.* 2009;42:298-308
60. Collins-Racie LA, Yang Z, Arai M, Li N, Majumdar MK, Nagpal S, et al. Global analysis of nuclear receptor expression and dysregulation in human osteoarthritic articular cartilage. *Osteoarthritis and Cartilage.* 2009;17:832-42
61. Ythier D, Larrieu D, Binet R, Binda O, Brambilla C, Gazzeri S, et al. Sumoylation of ING2 regulates the transcription mediated by Sin3A. *Oncogene.* 2010;29:5946-56
62. Haines BP, Wheldon LM, Summerbell D, Heath JK, Rigby PWJ. Regulated expression of FLRT genes implies a functional role in the regulation of FGF signalling during mouse development. *Developmental Biology.* 2006;297:14-25
63. Flintoff KA, Arudchelvan Y, Gong S-G. FLRT2 Interacts With Fibronectin in the ATDC5 Chondroprogenitor Cells. *Journal of Cellular Physiology.* 2014;229:1538-47
64. Gruber HE, Ingram J, Hoelscher GL, Norton HJ, Hanley ENJ. Cell Polarity in the Anulus of the Human Intervertebral Disc: Morphologic, Immunocytochemical, and Molecular Evidence. *Spine.* 2007;32:1287-94
65. Bronstein JM, Kozak CA, Chen X-N, Wu S, Danciger M, Korenberg JR, et al. Chromosomal Localization of Murine and Human Oligodendrocyte-Specific Protein Genes. *Genomics.* 1996;34:255-57
66. Eiraku M, Tohgo A, Ono K, Kaneko M, Fujishima K, Hirano T, et al. DNER acts as a neuron-specific Notch ligand during Bergmann glial development. *Nature Neuroscience.* 2005;8:873-80

67. Strauss U, Bräuer AU. Current views on regulation and function of plasticity-related genes (PRGs/LPPRs) in the brain. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids.* 2013;1831:133-38
68. Somogyi C, Matta C, Foldvari Z, Juhász T, Katona É, Takács Á, et al. Polymodal Transient Receptor Potential Vanilloid (TRPV) Ion Channels in Chondrogenic Cells. *International Journal of Molecular Sciences.* 2015;16:18412
69. Bonafe L, Cormier-Daire V, Hall C, Lachman R, Mortier G, Mundlos S, et al. Nosology and classification of genetic skeletal disorders: 2015 revision. *American Journal of Medical Genetics Part A.* 2015;167:2869-92
70. Xu Y, Yang Z, Yuan H, Li Z, Li Y, Liu Q, et al. PCDH10 inhibits cell proliferation of multiple myeloma via the negative regulation of the Wnt/β-catenin/BCL-9 signaling pathway. *Oncology Reports.* 2015;34:
71. Craft AM, Ahmed N, Rockel JS, Baht GS, Alman BA, Kandel RA, et al. Specification of chondrocytes and cartilage tissues from embryonic stem cells. *Development.* 2013;140:2597-610
72. Yamamoto M, Toya Y, Jensen RA, Ishikawa Y. Caveolin Is an Inhibitor of Platelet-Derived Growth Factor Receptor Signaling. *Experimental Cell Research.* 1999;247:380-88
73. Joosten PHLJ, Hol FA, van Beersum SEC, Peters H, Hamel BCJ, Afink GB, et al. Altered regulation of platelet-derived growth factor receptor-α gene-transcription in vitro by spina bifida-associated mutant Pax1 proteins. *Proceedings of the National Academy of Sciences of the United States of America.* 1998;95:14459-63
74. Battegay EJ, Raines EW, Seifert RA, Bowen-Pope DF, Ross R. TGF-beta induces bimodal proliferation of connective tissue cells via complex control of an autocrine PDGF loop. *Cell.* 1990;63:515-24
75. Paglia DN, Singh HMD, Karukonda T, Drissi H, Moss IL. PDGF-BB Delays Degeneration of the Intervertebral Discs in a Rabbit Preclinical Model. *Spine.* 2016;41:449-58
76. Smith RJ, Justen JM, Sam LM, Rohloff NA, Ruppel PL, Brunden MN, et al. Platelet-derived growth factor potentiates cellular responses of articular chondrocytes to interleukin-1. *Arthritis & Rheumatism.* 1991;34:697-706
77. Ullah M, Sittinger M, Ringe J. Extracellular matrix of adipogenically differentiated mesenchymal stem cells reveals a network of collagen filaments, mostly interwoven by hexagonal structural units. *Matrix Biology.* 2013;32:452-65
78. Jorge EC, Ahmed MU, Bothe I, Coutinho LL, Dietrich S. RGMa and RGMb expression pattern during chicken development suggest unexpected roles for these repulsive guidance molecules in notochord formation, somitogenesis, and myogenesis. *Developmental Dynamics.* 2012;241:1886-900
79. Dipaola CP, Farmer JC, Manova K, Niswander LA. Molecular signaling in intervertebral disc development. *Journal of Orthopaedic Research.* 2005;23:1112-19
80. Sugiura A, Ohtori S, Yamashita M, Inoue G, Yamauchi K, Koshi T, et al. Existence of Nerve Growth Factor Receptors, Tyrosine Kinase A and p75 Neurotrophin Receptors in Intervertebral Discs and on Dorsal Root Ganglion Neurons Innervating Intervertebral Discs in Rats. *Spine.* 2008;33:2047-51
81. Stoker AW. Isoforms of a novel cell adhesion molecule-like protein tyrosine phosphatase are implicated in neural development. *Mechanisms of development.* 1994;46:201-17
82. Hedge T, Mason I. Expression of Shisa2, a modulator of both Wnt and Fgf signaling, in the chick embryo. *The International journal of developmental biology.* 2008;52:81-5
83. Nagano T, Takehara S, Takahashi M, Aizawa S, Yamamoto A. Shisa2 promotes the maturation of somitic precursors and transition to the segmental fate in Xenopus embryos. *Development.* 2006;133:4643-54
84. Lambert C, Dubuc J-E, Montell E, Vergés J, Munaut C, Noël A, et al. Gene Expression Pattern of Cells From Inflamed and Normal Areas of Osteoarthritis Synovial Membrane. *Arthritis & Rheumatology.* 2014;66:960-68
85. Shin Y, Huh Y, Kim K, Kim S, Park K, Koh J-T, et al. Low-density lipoprotein receptor-related protein 5 governs Wnt-mediated osteoarthritic cartilage destruction. *Arthritis Research & Therapy.* 2014;16:R37
86. Chambrion C, Le Naour F. The Tetraspanins CD9 and CD81 Regulate CD9P1-Induced Effects on Cell Migration. *PLoS ONE.* 2010;5:e11219
87. Labeur M, Wölfel B, Stalla J, Stalla GK. TMEFF2 is an endogenous inhibitor of the CRH signal transduction pathway. *Journal of Molecular Endocrinology.* 2015;54:51-63

88. Grewal T, Wason Sundeep J, Enrich C, Rentero C. 2016. Annexins – insights from knockout mice. In *Biological Chemistry*, pp. 1031
89. Stevens JW, Kurriger GL, Carter AS, Maynard JA. CD44 expression in the developing and growing rat intervertebral disc. *Developmental Dynamics.* 2000;219:381-90
90. Richardson SM, Knowles R, Tyler J, Mobasheri A, Hoyland JA. Expression of glucose transporters GLUT-1, GLUT-3, GLUT-9 and HIF-1α in normal and degenerate human intervertebral disc. *Histochemistry and Cell Biology.* 2008;129:503-11
91. Iu J, Santerre JP, Kandel RA. Inner and Outer Annulus Fibrosus Cells Exhibit Differentiated Phenotypes and Yield Changes in Extracellular Matrix Protein Composition In Vitro on a Polycarbonate Urethane Scaffold. *Tissue Engineering Part A.* 2014;20:3261-69
92. Capurro MI, Xu P, Shi W, Li F, Jia A, Filmus J. Glypican-3 inhibits Hedgehog signaling during development by competing with patched for Hedgehog binding. *Developmental Cell.* 2008;14:700-11
93. Gruber H, Hoelscher G, Ingram J, Zinchenko N, Hanley E. Senescent vs. non-senescent cells in the human annulus in vivo: Cell harvest with laser capture microdissection and gene expression studies with microarray analysis. *BMC Biotechnology.* 2010;10:5
94. Gruber HE, Watts JA, Hoelscher GL, Bethea SF, Ingram JA, Zinchenko NS, et al. Mitochondrial gene expression in the human annulus: in vivo data from annulus cells and selectively harvested senescent annulus cells. *The Spine Journal.* 2011;11:782-91
95. Asai N, Ohkawara B, Ito M, Masuda A, Ishiguro N, Ohno K. LRP4 induces extracellular matrix productions and facilitates chondrocyte differentiation. *Biochemical and Biophysical Research Communications.* 2014;451:302-07
96. Eldridge S, Nalesso G, Ismail H, Vicente-Greco K, Kabouridis P, Ramachandran M, et al. Agrin mediates chondrocyte homeostasis and requires both LRP4 and α-dystroglycan to enhance cartilage formation in vitro and in vivo. *Annals of the Rheumatic Diseases.* 2016;75:1228-35
97. Moverare-Skrtic S, Henning P, Liu X, Nagano K, Saito H, Borjesson AE, et al. Osteoblast-derived WNT16 represses osteoclastogenesis and prevents cortical bone fragility fractures. *Nature Medicine.* 2014;20:1279-88
98. Nalesso G, Thomas BL, Sherwood JC, Yu J, Addimanda O, Eldridge SE, et al. WNT16 antagonises excessive canonical WNT activation and protects cartilage in osteoarthritis. *Annals of the Rheumatic Diseases.* 2016;76:218-26
99. Hiyama A, Skubutyte R, Markova D, Anderson DG, Yadla S, Sakai D, et al. Hypoxia activates the notch signaling pathway in cells of the intervertebral disc: Implications in degenerative disc disease. *Arthritis & Rheumatism.* 2011;63:1355-64
100. Yamada H, Yasuda T, Kotorii S, Takahashi K, Tabira T, Sunada Y. Report of a patient with CADASIL having a novel missense mutation of the Notch 3 gene-association with alopecia and lumbar herniated disc. *Clinical neurology.* 2001;41:144-6
101. Wang H, Tian Y, Wang J, Phillips KLE, Binch ALA, Dunn S, et al. Inflammatory Cytokines Induce NOTCH Signaling in Nucleus Pulposus Cells: IMPLICATIONS IN INTERVERTEBRAL DISC DEGENERATION. *Journal of Biological Chemistry.* 2013;288:16761-74
102. Johnson KR, Nicodemus-Johnson J, Spindler MJ, Carnegie GK. Genome-Wide Gene Expression Analysis Shows AKAP13-Mediated PKD1 Signaling Regulates the Transcriptional Response to Cardiac Hypertrophy. *PLoS ONE.* 2015;10:e0132474
103. Kino T, Takatori H, Manoli I, Wang Y, Tiulpakov A, Blackman MR, et al. Brx Mediates the Response of Lymphocytes to Osmotic Stress Through the Activation of NFAT5. *Science signaling.* 2009;2:ra5-ra5
104. Miyabe Y, Miyabe C, Iwai Y, Yokoyama W, Sekine C, Sugimoto K, et al. Activation of fibroblast-like synoviocytes derived from rheumatoid arthritis via lysophosphatidic acid–lysophosphatidic acid receptor 1 cascade. *Arthritis Research & Therapy.* 2014;16:461
105. David M, Machuca-Gayet I, Kikuta J, Ottewell P, Mima F, Leblanc R, et al. Lysophosphatidic acid receptor type 1 (LPA1) plays a functional role in osteoclast differentiation and bone resorption activity. *Journal of Biological Chemistry.* 2014;289:6551-64
106. Bhat R, Bissell MJ. Of plasticity and specificity: dialectics of the microenvironment and macroenvironment and the organ phenotype. *Wiley Interdisciplinary Reviews: Developmental Biology.* 2014;3:147-63

Supplemental information

Extended Experimental Methods

Human IVD tissue

Research involving human material was performed in accordance with the Declaration of Helsinki. Human IVD tissue from idiopathic scoliosis patients was obtained as surplus material from scoliosis correction surgery (approval ID MEC 08-4-021). Immortalization and cloning procedures of NP and AF cells have been extensively described before (1, 2). Adult human IVD tissue was obtained from a 63 years old, non-heart beating donor (3): two IVDs L1/L2 (without overt signs of degeneration; healthy) and L4/L5 (clearly degenerated; *cf.* Figure 1C, D) were dissected by an orthopedic surgeon. Approval for all experimental sections of the current study and informed consent for publication of patient details and accompanying images in this manuscript was obtained as an integral part of the MUMC-Medical Ethical Review Committee (METC approval 08-4-021; July 11, 2012); the approval is held by the authors (LWvR) and is available for review by the Editor-in-Chief. Macroscopic pictures of tissues from L1/L2 and L4/L5 IVDs were captured with an Olympus Zuiko mirror reflex camera. Tissue was dissected; fixed in formalin, paraffin embedded and 5 µm sections were stained with Safranin O/Hematox according to standard protocols (4). Microscopic images were captured using a NIKON TE200 Eclipse microscope and photographed using a NIKON DXM1200 digital camera.

Cell culture

The original isolation of primary NP and AF material from young donors (age 8-15 years), the immortalization of NP and AF cells and the establishment and

characterization of immortal NP and AF cell lines have been described in detail elsewhere (1). NP and AF cell clones were standardly cultured in DMEM-F12/Glutamax (Gibco), 10% fetal calf serum (FCS; Biowhittaker, cat no DE14-801F), 1% penicillin/streptomycin (Gibco), 1% non-essential amino acids (NEAA; Gibco). All cell clones used for expression array analysis were derived from one donor (D5; 13 years old, male *spina bifida* patient; IVD levels T6-L1) (1, 2). For array expression analyses 4 independent cell clones representing different clonal phenotypes (*cf.* Figure 2B) were cultured on 158 cm^2 culture dishes (Greiner) and harvested during the exponential growth phase at 48 hours post-plating, at approximately 30-40% confluency, under standard culturing conditions (Supplementary Figure 1A). Samples were harvested and processed for RNA isolation and subsequent whole transcriptome analysis or cDNA synthesis for quantitative PCR.

RNA isolation and quantitative PCR

RNA was isolated according to manufacturer's instructions with the RNeasy kit (QIAGEN). Total RNA (500ng) for each sample was converted into first strand cDNA with the iScript cDNA synthesis kit (Bio-Rad, USA) according to manufacturer's instructions. Gene expression was determined using Real-time quantitative PCR (RT-qPCR) with SYBR® Green (Eurogentec). Primer sets were validated and are depicted in Table 7. For amplification, an Applied Biosystems ABI PRISM 7700 Sequence Detection System was used: initial denaturation 95°C for 10 min was followed by 40 cycles of DNA amplification. Data was analyzed using the standard curve method and normalized to *Cyclophillin A*. Statistical significance ($p<0.05$) was determined by two-tailed Student's *t*-test.

Whole transcriptome expression datasets

Raw data from published whole transcriptome expression measurements of primary human AF, NP and AC cells from 3 different donors per tissue (n=9 in total), was provided by Prof. J.H. Hoyland and Dr. S.M. Richardson (Manchester, UK). Briefly, high-quality RNA with an RNA integrity number of at least 7 taken from 3 histologically "normal" IVD samples (2 male and 1 female donor, ages 46–57 years [mean age 51 years]; histologic grade 1 or 2) and articular cartilage samples (1 male and 2 female donors, ages 55–60 years [mean age 58 years]; Mankin grade 1 or 2) were used for the microarray analyses (5). The calculated mean age of all six AF and NP was 55 years.

Immortal cell clone RNA samples were processed for quality control, labeling, hybridization and data extraction at Service XS B.V. (Leiden, the Netherlands). The RNA integrity and concentration was determined using lab-on-chip analysis on the Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA). Biotinylated cRNA was prepared using the illumine TotalPrep RNA Amplification Kit (Ambion, Inc., Austin, TX, USA) according to manufacturer's specifications with an input of 200 ng total RNA. Representative, high-quality RNA preparations from triplicate samples for each of the 12 distinct immortal cell clones, were characterized using 12 arrays (*cf.* Figure 2B). Per clone, 750 ng of the obtained biotinylated cRNA samples was hybridized onto the Illimuna humanHT-12v4 (Illumina, Inc., San Diego, CA, USA). Hybridization and washing were performed according to the Illumina manual "Direct hybridization assay guide". Scanning was performed on the Illumina iScan (Illumina, Inc., San Diego, CA, USA). Image analysis and extraction of raw expression data was performed with Illumina GenomeStudio v2011.1 gene expression software with default settings (no background correction and no normalization).

The immortal cell clone NP-R, NP-nR and AF array dataset and primary NP, AF, AC dataset were analyzed using the open source scripting language R (version 2.13.0) and R packages of Bioconductor 2.8. Established quality control, visualization, normalization and statistical methods were combined in the workflow using the pipeline from ArrayAnalysis.org (6).

The limma package (7) was used to apply linear regression modelling with modified t-tests to evaluate statistically significant differences between groups: primary NP to AC, NP to AF and immortal NP-R to NP-nR, NP-R to AF, NP-nR to AF (Benjamini Hochberg FDR adjusted p value <0.05 was considered significant). For the comparison of AF-S and AF-nS cell clones the regular p value <0.05 was used. Additional criteria were used to identify relevant expression differences in a more stringent procedure: fold change larger than 2.0 and an average expression value of 100 or higher (mean fluorescence intensity) in the cell type of interest.

To enable principle component analysis (PCA) using datasets derived from distinct microarray platforms, comparison of immortal cell clone and primary cell datasets generated on Affymetrix and Illumina platforms, respectively, was done by ranking genes by descending expression value and assigning an ordinal number reflecting their relative position within each dataset.

Network analyses were performed using Cytoscape and the GeneMANIA plugin (8, 9). Membrane expressed genes, as determined by meta-analyses (10), and were coupled to our whole transcriptome datasets in R based on Entrez Gene ID resulting in a list of 6,134 matched membrane expressed genes.

Supplemental References

1. van den Akker GGH, Surtel DAM, Cremers A, Rodrigues-Pinto R, Richardson SM, et al. Novel immortal human cell lines reveal subpopulations in the nucleus pulposus. *Arthritis Research & Therapy.* 2014;16:R135-R35
2. van den Akker GG, Surtel DA, Cremers A, Richardson SM, Hoyland JA, et al. Novel Immortal Cell Lines Support Cellular Heterogeneity in the Human Annulus Fibrosus. *PLoS ONE.* 2016;11:e0144497
3. van den Akker GG, Surtel DA, Cremers A, Hoes MF, Caron MM, et al. EGR1 controls divergent cellular responses of distinctive nucleus pulposus cell types. *BMC Musculoskelet Disord.* 2016;17:124
4. Welting T, Caron M, Emans P, Janssen M, Sanen K, et al. Inhibition of cyclooxygenase-2 impacts chondrocyte hypertrophic differentiation during endochondral ossification. *eCM.* 2011;22:420-37
5. Minogue BM, Richardson SM, Zeef LAH, Freemont AJ, Hoyland JA. Characterization of the human nucleus pulposus cell phenotype and evaluation of novel marker gene expression to define adult stem cell differentiation. *Arthritis & Rheumatism.* 2010a;62:3695-705
6. Eijssen LM, Jaillard M, Adriaens ME, Gaj S, de Groot PJ, et al. User-friendly solutions for microarray quality control and pre-processing on ArrayAnalysis.org. *Nucleic acids research.* 2013;41:W71-6
7. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research.* 2015;43:e47-e47
8. Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics.* 2010;26:2927-28
9. Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. Cytoscape Web: an interactive web-based network browser. *Bioinformatics.* 2010;26:2347-8
10. Uva P, Lahm A, Sbardellati A, Grigoriadis A, Tutt A, de Rinaldis E. Comparative *Membranome* Expression Analysis in Primary Tumors and Derived Cell Lines. *PLoS ONE.* 2010;5:e11742

Supplemental Figure legends

Figure S1: Morphological characteristics of immortalized cell clones. A) Phase contrast images of NP-R, NP-nR and AF exponentially growing cell cultures prior to RNA isolation; bar represents 20 micron. B) Gene expression analysis of *FOXF1, CA12, COL12A1, SFRP2* and *CD24* in indicated cell clones (n=3 biological repeats per clone). Consistent with earlier observations, *FOXF1* and *CA12* expression was highest in NP-R clones, while *CD24* mRNA levels were highest in NP-nR clones (9). *SFRP2* was strongly expressed in AF and NP-nR clones and was nearly undetectable in NP-R clones. *COL12A1* was significantly higher expressed in both AF and NP-nR clones compared to NP-R clones. Single dots represent a single measurement; horizontal lines represent the mean of all measurements. Correspondence of symbols to clone identity is indicated in the table (bottom right). Gene expression was normalized to *Cyclophillin A*. Statistical significant differences are indicated by an asterisk (two-tailed Student's *t*-test; p value <0.05).

Figure S2: Principle component analysis clusters immortalized NP and AF clones. A) PCA plot of immortalized NP-Responder (n=4; red square symbols), NP non-Responder (n=4; blue square symbols) and AF (n=4; green triangular symbols) cell clones. NP-nR clone 102 was considered an outlier based on these analyses and additional gene expression measurements and was not included in subsequent statistical analyses.

Figure S3: Published NP marker expression in primary NP cells. The gene networks depicted in Figure 3 were used to visualize expression levels of corresponding genes in primary NP cell isolates for either NP marker genes (A) or NP progenitor-associated genes (B). A) Selected NP cell marker gene network. Circle-size indicates an arbitrary log2-based expression value (*i.e.* based on fluorescence intensity on the micro-array); a log2-based expression value below 6.0 indicates low/absent gene expression (*i.e.* not expected to be biologically relevant). Colored lines in the network represent established biological connections between markers; parentheses: percentage representation. The highly expressed genes (*KRT18, FOXF1*) are among the best known NP cell marker genes. B) NP progenitor-associated gene network. Gene network showing log2-based expression values of the indicated genes in primary NP cells (*cf.* Figure 3B). Circle-size indicates an arbitrary log2-based expression value (*i.e.* based on fluorescence intensity on the micro-array. *VEGFA, ENO2, CD24, CD63, ITGB1, ITGAV* and *POU5F1* are strongly expressed by primary NP cells. Of note: expression of genes strongly linked to stem cells and NP progenitor cells is low/not detectable in primary isolates.

Figure S4: Validation of membrane-associated NP and AF cellular heterogeneity markers. A) rtPCR measurements for NP-R marker genes *LPPR4, CLDN11, PDGFRA*, and *RGMB* in four NP-R, NP-nR and AF clones under normal proliferation conditions. B) rtPCR measurements for NP-nR marker genes *PTPRF, SHISA2, PTGFRN* and *LRP5* in four NP-R, NP-nR and AF clones. C) rtPCR measurements for AF-S and AF-nS marker genes *NOTCH3, JAG1* and *HES1*. Gene selection was arbitrarily based on the fold expression difference and previous reports in literature. Correspondence of symbols to clone identity is indicated in the table (bottom right). Gene expression was normalized to

the housekeeper gene *Cyclophillin A*. Statistical significant differences are indicated by asterisks (two-tailed Student's *t*-test; p value <0.05).

Supplemental Tables

Species *(ref#)*	Rat *(23)*	Canine *(24)*	Human *(26)*
Methods	micro arrays	micro arrays	micro arrays
Tissues	NP, AF, AC	NP, AF, AC	NP, AF
Markers validated (rtPCR)			
NP positive (vs AC)	***KRT19*, *PTN*, *GPC3*, *ANX3*, *VIM***	*KRT18*, *A2M*, *DSC2*, *NCAM1*	***PAX1*, *FOXF1*, *CA12*, *HBB*, *OVOS2***
NP positive (vs AF)	***KRT19*, *GPC3*, *PTN*, *ANX3*, *VIM***		
IVD positive (vs AC)			
NC positive (vs NP)			
AC positive (vs NP)	***COMP*, *MGP***	*MGP*, ***COMP*, *PTN***	*GDF10*, *CYTL*, *IBSP*, ***FBLN1***
AF positive (vs NP)	***COMP***	*GPC3*, ***COMP***	
AF & AC (vs NP)	***COMP***		
not a marker (AF vs NP)			
not a marker (NP)	***COMP***		
not a marker (IVD)			
Markers validated (IHC)	-	on IVD tissue A2M, NCAM1, DSC2, KRT18	in differentiated MSCs PAX1, FOXF1, IBSP, FBLN1

Species *(ref)*	Bovine *(25)*	Human *(30)*	Human *(27)*
Methods	micro arrays	candidate approach	micro arrays
Tissues	NP, NC, AF, AC	NP, AF, AC	NP, AF, AC
Markers validated (rtPCR)			
NP positive (vs AC)	***KRT19*, *KRT18*, *VCAN*, *SNAP25*, *CDH2*, *SOSTDC1***	*KRT18*, ***KRT19***, *VIM*, *NCAM1*	***CA12*, *CLEC2B*, *SGCG*, *TYRO3***
NP positive (vs AF)	*SNAP25*, *CDH2*, *SOSTDC1*	***KRT19***, *NCAM1*, *A2M*, *DSC2*	
IVD positive (vs AC)	*VCAN*, *TNMD*, *BASP1*, *TNFAIP6*, ***FOXF1***, *FOXF2*, *AQP1*		
NC positive (vs NP)	***T***		
AC positive (vs NP)		***COMP***, *MGP*, ***PTN***	
AF positive (vs NP)		*GPC3*, ***COMP***	
AF & AC (vs NP)	***PTN***		
not a marker (AF vs NP)	***COL2A1, AGC, CD24***		
not a marker (NP)	*IBSP*, *GPC3*, *ANX3*, *VIM*, ***COMP***, *MGP*, *A2M*, *ANX4*, *DSC2*, *NCAM1*	***CD24***	
not a marker (IVD)	***FBLN1***		
Markers validated (IHC)	-	on IVD tissue KRT19, MGP	on IVD tissue CA12

Table S1: Overview of published NP, AF and AC markers. Depicted are marker genes identified by whole transcriptome expression studies performed on freshly isolated cells, isolated from the indicated tissue and species. The genes indicated in bold showed consistent results while underlined genes did not under normal culturing or chondrogenic conditions. Interspecies variation is relatively high and certain markers (e.a. *GPC3, MGP, PTN*) are not exclusively linked to one cell type in literature. #: Citations correspond to references in main article.

cf. Figure	network[A]	Gene markers
Figure 3A Figure S3A	NP phenotype	*CD49F, CD56 (NCAM1), CD73 (NT5E), CD90, CD105, CD166, KRT19, PTN, GPC3, ANXA3, VIM, KRT18, A2M, DSC2, PAX1, FOXF1, CA12, HBB, VCAN, SNAP25, CDH2, SOSTDC1, CLEC2B, SGCG, TYRO3*
Figure 3B Figure S3B	NP maturity	*ALCAM, ANGPT1, CD24, CD63, ENG, ENO2, FLT1, GNL3, HSPB1, ITGAV, ITGB1, JAG1, KIT, NANOG, NGFR, NOTCH1, NT5E, PLCD4, POU5F1, PROM1, PTPRC, SHH, SNAI1, SOX2, TEK, THY1, VEGFA*

Table S2: Genes used for network building. [A]: network generation in GenMania/Cytoscape.

Gene	ID[A]	FC[B]	*p*[C]	Gene description
LPPR4	**9890**	**12,2**	**0,0001**	**plasticity related gene 1**
TRPV2	51393	11,3	0,0010	transient receptor potential cation channel, subfamily V, member 2
PCDH10	57575	10,2	0,0012	protocadherin 10
MXRA8	54587	9,1	0,0001	matrix-remodelling associated 8
PDGFRA	**5156**	**8,2**	**0,0005**	**platelet-derived growth factor receptor, alpha polypeptide**
CLDN11	**5010**	**6,9**	**0,0113**	**claudin 11 (oligodendrocyte transmembrane protein)**
GPER	2852	6,9	0,0019	G protein-coupled receptor 30
SCARA3	51435	6,8	0,0009	scavenger receptor class A, member 3
ELTD1	64123	6,1	0,0129	EGF, latrophilin and seven transmembrane domain containing 1
NLGN1	22871	6,0	0,0049	neuroligin 1
RGMB	**285704**	**5,5**	**0,0000**	**RGM domain family, member B**
TMEFF2	23671	5,5	0,0053	transmembrane protein with EGF-like and two follistatin-like domains 2
ADAM23	8745	5,0	0,0103	a disintegrin and metalloproteinase domain 23
GYPC	2995	4,7	0,0015	glycophorin C (Gerbich blood group)
PERP	64065	4,5	0,0145	PERP, TP53 apoptosis effector
ADCY3	109	4,3	0,0001	adenylate cyclase 3
TSPAN32	10077	4,3	0,0121	pan-hematopoietic expression
SLC16A5	9121	4,2	0,0011	solute carrier family 16 (monocarboxylic acid transporters), member 5
STEAP1	26872	4,2	0,0006	six transmembrane epithelial antigen of the prostate 1
COLEC12	81035	4,2	0,0302	collectin sub-family member 12
FAM162B	221303	4,1	0,0049	chromosome 6 open reading frame 189
RSPO3	84870	4,1	0,0082	thrombospondin, type I, domain containing 2
LPAR1	1902	3,9	0,0121	Endothel. diff., lysophosphatidic acid G-protein-coupled receptor, 2
KIAA1324L	222223	3,8	0,0006	hypothetical protein FLJ31340
KCNJ8	3764	3,7	0,0079	K^+ inwardly-rectifying channel, subfamily J, member 8
TMEM26	219623	3,6	0,0254	transmembrane protein 26
EPHA3	2042	3,6	0,0027	EPH receptor A3
BDKRB1	623	3,5	0,0019	bradykinin receptor B1
PCDH18	54510	3,4	0,0237	protocadherin 18
ANPEP	290	3,3	0,0111	alanyl (membrane) aminopeptidase (N, M, microsomal, CD13, p150)
TSPAN5	10098	3,3	0,0172	transmembrane 4 superfamily member 9
SLC17A9	63910	3,2	0,0003	chromosome 20 open reading frame 59
S1PR3	1903	3,2	0,0110	endothelial differentiation, sphingolipid G-protein-coupled receptor, 3
TNFRSF19	55504	3,2	0,0309	tumor necrosis factor receptor superfamily, member 19
LPHN2	23266	3,1	0,0205	latrophilin 2
VNN2	8875	3,1	0,0026	vanin 2
TMEM71	137835	3,1	0,0071	hypothetical protein FLJ33069
ITGA2	3673	3,0	0,0276	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)
LEPR	3953	3,0	0,0240	leptin receptor
POPDC3	64208	3,0	0,0046	popeye domain containing 3
LRP11	84918	2,7	0,0025	low density lipoprotein receptor-related protein 11
CA12	771	2,7	0,0254	carbonic anhydrase XII

TSPAN10	83882	2,7	0,0061	oculospanin
IL11RA	3590	2,6	0,0012	interleukin 11 receptor, alpha
CD68	968	2,6	0,0036	CD68 antigen
MGC42105	167359	2,6	0,0060	hypothetical protein MGC42105
NT5E	4907	2,6	0,0289	5'-nucleotidase, ecto (CD73)
DSEL	92126	2,5	0,0098	chromosome 18 open reading frame 4
ITM2C	81618	2,5	0,0025	integral membrane protein 2C
ANXA2	302	2,5	0,0086	annexin A2
ITGA6	3655	2,5	0,0052	integrin, alpha 6
PQLC3	130814	2,5	0,0043	chromosome 2 open reading frame 22
SSFA2	6744	2,4	0,0156	sperm specific antigen 2
SCN2A	6326	2,4	0,0333	Na^+ channel, voltage-gated, type II, alpha 2
ANTXR2	118429	2,4	0,0146	anthrax toxin receptor 2
FAM176B	55194	2,3	0,0225	hypothetical protein FLJ10647
NINJ2	4815	2,3	0,0442	ninjurin 2
FDFT1	2222	2,3	0,0016	farnesyl-diphosphate farnesyltransferase 1
SLC9A9	285195	2,3	0,0187	solute carrier family 9 (sodium/hydrogen exchanger), isoform 9
GPR162	27239	2,3	0,0177	gene rich cluster, A gene
FAM176A	84141	2,3	0,0042	hypothetical protein FLJ13391
MERTK	10461	2,3	0,0366	c-mer proto-oncogene tyrosine kinase
GSG1	83445	2,2	0,0072	germ cell associated 1
STEAP3	55240	2,2	0,0099	dudulin 2
CLEC2B	9976	2,2	0,0412	C-type lectin domain family 2, member B
TMEM38B	55151	2,2	0,0172	transmembrane protein 38B
CD44	960	2,2	0,0422	CD44 antigen (homing function and Indian blood group system)
F2RL1	2150	2,1	0,0309	coagulation factor II (thrombin) receptor-like 1
EVI2B	2124	2,1	0,0128	ecotropic viral integration site 2B
TSPAN4	7106	2,1	0,0265	transmembrane 4 superfamily member 7
SLCO3A1	28232	2,1	0,0218	solute carrier organic anion transporter family, member 3A1
IL1R1	3554	2,1	0,0189	interleukin 1 receptor, type I
VAT1	10493	2,1	0,0196	vesicle amine transport protein 1 homolog (T californica)
DCHS1	8642	2,1	0,0321	dachsous 1 (Drosophila)
LRRC8C	84230	2,1	0,0206	factor for adipocyte differentiation 158
KIAA1715	80856	2,1	0,0018	KIAA1715
C2orf28	51374	2,1	0,0019	chromosome 2 open reading frame 28
LY6K	54742	2,0	0,0079	lymphocyte antigen 6 complex, locus K
LRFN5	145581	2,0	0,0160	leucine rich repeat and fibronectin type III domain containing 5

Table S3: NP-R marker genes (compared to NP-nR). [A]: ENTREZ Gene ID; [B]: Fold Change (NP-R/NP-nR); [C]: adjusted p value.

Gene	ID[A]	FC[B]	*p*[C]	Gene description
PTPRF	**5792**	**-22,5**	**0,0001**	**protein tyrosine phosphatase, receptor type, F**
CD70	970	-20,3	0,0006	tumor necrosis factor (ligand) superfamily, member 7
SHISA2	**387914**	**-18,9**	**0,0000**	**chromosome 13 open reading frame 13**
GPRC5C	55890	-17,4	0,0001	G protein-coupled receptor, family C, group 5, member C
CXCR7	57007	-13,7	0,0004	chemokine orphan receptor 1
PTGFRN	**5738**	**-11,4**	**0,0000**	**prostaglandin F2 receptor negative regulator**
CHST15	51363	-8,7	0,0053	B cell RAG associated protein
VLDLR	7436	-8,1	0,0005	very low density lipoprotein receptor
HEPH	9843	-6,0	0,0013	hephaestin
PDPN	10630	-5,9	0,0057	lung type-I cell membrane-associated glycoprotein
DLL3	10683	-5,8	0,0003	delta-like 3 (Drosophila)
CDH6	1004	-5,6	0,0005	cadherin 6, type 2, K-cadherin (fetal kidney)
SLC6A9	6536	-5,5	0,0001	solute carrier family 6 (neurotransmitter transporter, glycine), member 9
PLD6	201164	-5,5	0,0012	similar to CG12314 gene product
LRP5	**4041**	**-5,1**	**0,0001**	**low density lipoprotein receptor-related protein 5**
EFNB3	1949	-4,8	0,0333	ephrin-B3
SLC2A1	6513	-4,6	0,0224	solute carrier family 2 (facilitated glucose transporter), member 1
PTPRE	5791	-4,5	0,0008	protein tyrosine phosphatase, receptor type, E
CDH13	1012	-4,1	0,0019	cadherin 13, H-cadherin (heart)
BST1	683	-4,0	0,0430	bone marrow stromal cell antigen 1
PLXDC2	84898	-3,9	0,0391	plexin domain containing 2
CLDN1	9076	-3,8	0,0411	claudin 1
NETO2	81831	-3,6	0,0022	neuropilin (NRP) and tolloid (TLL)-like 2
ROR1	4919	-3,6	0,0099	receptor tyrosine kinase-like orphan receptor 1
PMEPA1	56937	-3,5	0,0385	transmembrane, prostate androgen induced RNA
ITGA11	22801	-3,4	0,0471	integrin, alpha 11
SLC38A1	81539	-3,4	0,0010	solute carrier family 38, member 1
APCDD1L	164284	-3,4	0,0470	hypothetical protein FLJ90166
TSPAN13	27075	-3,2	0,0256	transmembrane 4 superfamily member 13
EFNA1	1942	-3,2	0,0034	ephrin-A1
THY1	7070	-3,2	0,0347	Thy-1 cell surface antigen
SUSD3	203328	-3,1	0,0188	sushi domain containing 3
SLC7A5	8140	-2,9	0,0008	solute carrier family 7 (cation.aminoacid transporter, y+ system), member 5
ABCA3	21	-2,8	0,0172	ATP-binding cassette, sub-family A (ABC1), member 3
TNFSF9	8744	-2,8	0,0196	tumor necrosis factor (ligand) superfamily, member 9
GRAMD3	65983	-2,8	0,0005	HCV NS3-transactivated protein 2
CELSR2	1952	-2,7	0,0015	cadherin, EGF LAG seven-pass G-type receptor 2 (flamingo homolog *D.mel*)
PVR	5817	-2,7	0,0012	poliovirus receptor
IL27RA	9466	-2,7	0,0049	interleukin 27 receptor, alpha
PRUNE2	158471	-2,6	0,0080	KIAA0367
NPDC1	56654	-2,6	0,0249	neural proliferation, differentiation and control, 1
MAMDC4	158056	-2,5	0,0008	apical early endosomal glycoprotein precursor

OXTR	5021	-2,5	0,0031	oxytocin receptor
NTM	50863	-2,5	0,0129	Neurotrimin
MUC1	4582	-2,5	0,0277	mucin 1, transmembrane
C4orf49	84709	-2,5	0,0110	ovary-specific acidic protein
PROM1	8842	-2,5	0,0214	prominin 1
SDC2	6383	-2,5	0,0148	syndecan 2 (heparan sulfate proteoglycan 1, cell surface-assoc. fibroglycan)
CTXN1	404217	-2,4	0,0153	cortexin 1
ENPP4	22875	-2,4	0,0342	ectonucleotide pyrophosphatase/phosphodiesterase 4
CMTM8	152189	-2,4	0,0208	chemokine-like factor super family 8
CMTM7	112616	-2,3	0,0216	chemokine-like factor super family 7
PLXNB1	5364	-2,3	0,0042	plexin B1
ACCN2	41	-2,3	0,0484	amiloride-sensitive cation channel 2, neuronal
RHBDF2	79651	-2,3	0,0281	rhomboid, veinlet-like 6 (Drosophila)
SLC2A6	11182	-2,3	0,0148	solute carrier family 2 (facilitated glucose transporter), member 6
GDPD5	81544	-2,3	0,0314	hypothetical protein PP1665
P2RY11	5032	-2,3	0,0430	purinergic receptor P2Y, G-protein coupled, 11
REEP2	51308	-2,3	0,0197	chromosome 5 open reading frame 19
C14orf37	145407	-2,3	0,0120	chromosome 14 open reading frame 37
SIRPA	140885	-2,3	0,0117	protein tyrosine phosphatase, non-receptor type substrate 1
ADAM12	8038	-2,2	0,0052	a disintegrin and metalloproteinase domain 12 (meltrin alpha)
C11orf75	56935	-2,2	0,0154	FN5 protein
F2R	2149	-2,2	0,0030	coagulation factor II (thrombin) receptor
SLC26A6	65010	-2,2	0,0058	solute carrier family 26, member 6
KIAA1324	57535	-2,2	0,0159	maba1
TMEM56	148534	-2,2	0,0037	hypothetical protein FLJ31842
TGFA	7039	-2,1	0,0189	transforming growth factor, alpha
GPC2	221914	-2,1	0,0149	glypican 2 (cerebroglycan)
TMEM39A	55254	-2,1	0,0103	transmembrane protein 39A
FZD9	8326	-2,1	0,0482	frizzled homolog 9 (Drosophila)
HBEGF	1839	-2,1	0,0385	heparin-binding EGF-like growth factor
FAIM3	9214	-2,1	0,0141	regulator of Fas-induced apoptosis
CD200	4345	-2,1	0,0479	CD200 antigen
RRP12	23223	-2,1	0,0183	KIAA0690
FIBCD1	84929	-2,0	0,0319	fibrinogen C domain containing 1
IFRD1	3475	-2,0	0,0053	interferon-related developmental regulator 1
CHRNA5	1138	-2,0	0,0237	cholinergic receptor, nicotinic, alpha polypeptide 5
SLC1A4	6509	-2,0	0,0480	solute carrier family 1 (glutamate/neutral amino acid transporter) member 4
HPN	3249	-2,0	0,0262	hepsin (transmembrane protease, serine 1)
CACNA1C	775	-2,0	0,0421	Ca^{2+} channel, voltage-dependent, L type, alpha 1C subunit
IL1RAP	3556	-2,0	0,0237	interleukin 1 receptor accessory protein

Table S4: NP-nR marker genes (compared to NP-R). [A]: ENTREZ Gene ID; [B]: Fold Change (NP-R/NP-nR); [C]: adjusted p value.

Gene	ID[A]	FC[B]	*p*[C]	Gene description
NOTCH3	**4854**	**4,8**	**0,0010**	**notch 3**
KCNG1	3755	4,4	0,0002	K^+ voltage-gated channel, subfamily G, member 1
CD14	929	3,6	0,0017	CD14 molecule
JAG1	**182**	**3,5**	**0,0000**	**jagged 1**
HAS1	3036	3,2	0,0047	hyaluronan synthase 1
KCNMB1	3779	3,2	0,0035	K^+ large conductance $Ca2^+$-activated channel, subfamily M, beta member 1
GHSR	2693	3,0	0,0000	growth hormone secretagogue receptor
SLC26A7	115111	2,7	0,0028	solute carrier family 26, member 7
P2RY2	5029	2,7	0,0148	purinergic receptor P2Y, G-protein coupled, 2
CADM3	57863	2,6	0,0178	cell adhesion molecule 3
GPM6B	2824	2,5	0,0007	glycoprotein M6B
NPR3	4883	2,5	0,0244	(atrio)natriuretic peptide receptor C/guanylate cyclase C
DYSF	8291	2,5	0,0333	dysferlin, limb girdle muscular dystrophy 2B (autosomal recessive)
PMEPA1	56937	2,5	0,0071	prostate transmembrane protein, androgen induced 1
SUSD2	56241	2,5	0,0002	sushi domain containing 2
LRRC25	126364	2,5	0,0204	leucine rich repeat containing 25
GPR124	25960	2,5	0,0263	G protein-coupled receptor 124
ITGA11	22801	2,4	0,0198	integrin, alpha 11
TTYH2	94015	2,4	0,0394	tweety homolog 2 (Drosophila)
KCNK17	89822	2,3	0,0037	K^+ channel, subfamily K, member 17
LRRC32	2615	2,3	0,0335	leucine rich repeat containing 32
GPR115	221393	2,3	0,0006	G protein-coupled receptor 115
EFNB1	1947	2,3	0,0044	ephrin-B1
CHIC1	53344	2,3	0,0017	cysteine-rich hydrophobic domain 1
THSD7B	80731	2,2	0,0012	thrombospondin, type I, domain containing 7B
ATP2B2	491	2,2	0,0089	ATPase, Ca++ transporting, plasma membrane 2
ITPRIPL2	162073	2,2	0,0006	inositol 1,4,5-trisphosphate receptor interacting protein-like 2
ABCA10	10349	2,2	0,0116	ATP-binding cassette, sub-family A (ABC1), member 10
C3orf35	339883	2,2	0,0024	chromosome 3 open reading frame 35
SLC43A1	8501	2,2	0,0014	solute carrier family 43, member 1
ZP4	57829	2,2	0,0438	zona pellucida glycoprotein 4
CD6	923	2,2	0,0019	CD6 molecule
PKHD1	5314	2,2	0,0233	polycystic kidney and hepatic disease 1 (autosomal recessive)
CDHR5	53841	2,2	0,0014	cadherin-related family member 5
SLC27A6	28965	2,2	0,0176	solute carrier family 27 (fatty acid transporter), member 6
OR52J3	119679	2,1	0,0021	olfactory receptor, family 52, subfamily J, member 3
SLC29A4	222962	2,1	0,0052	solute carrier family 29 (nucleoside transporters), member 4
SLCO2B1	11309	2,1	0,0185	solute carrier organic anion transporter family, member 2B1
CTAGE1	64693	2,1	0,0110	cutaneous T-cell lymphoma-associated antigen 1
ALPL	249	2,1	0,0300	alkaline phosphatase, liver/bone/kidney
FAP	2191	2,1	0,0129	fibroblast activation protein, alpha
SPNS3	201305	2,1	0,0040	spinster homolog 3 (Drosophila)

FOLH1	2346	2,1	0,0075	folate hydrolase (prostate-specific membrane antigen) 1
CDH26	60437	2,1	0,0208	cadherin 26
LST1	7940	2,1	0,0024	leukocyte specific transcript 1
KCNN3	3782	2,1	0,0025	K^+ intermediate/small conductance Ca2+-activated channel, subfam N, mmbr 3
SUCO	51430	2,1	0,0295	SUN domain containing ossification factor
LGR6	59352	2,1	0,0111	leucine-rich repeat containing G protein-coupled receptor 6
SLCO1A2	6579	2,1	0,0062	solute carrier organic anion transporter family, member 1A2
SEMA5B	54437	2,1	0,0049	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5B
PTCHD4	442213	2,1	0,0165	patched domain containing 4
CMTM1	113540	2,1	0,0322	CKLF-like MARVEL transmembrane domain containing 1
CNGA4	1262	2,0	0,0168	cyclic nucleotide gated channel alpha 4
LINC00340	401237	2,0	0,0097	long intergenic non-protein coding RNA 340
PCDHGA1	56114	2,0	0,0102	protocadherin gamma subfamily A, 1
OPN1MW	2652	2,0	0,0409	opsin 1 (cone pigments), medium-wave-sensitive
KIR2DL1	3802	2,0	0,0451	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 1
GJA5	2702	2,0	0,0445	gap junction protein, alpha 5, 40kDa

Table S5: AF-nS marker genes (compared to AF-S). [A]: ENTREZ Gene ID; [B]: Fold Change (AF-nS/AF-S); [C]: adjusted p value.

Gene	ID[A]	FC[B]	*p*[C]	Gene description
CD24	100133941	-7,9	0,0085	CD24 molecule
HAS3	3038	-4,8	0,0177	hyaluronan synthase 3
CLDN11	**5010**	**-3,5**	**0,0014**	**claudin 11**
GPRC5A	9052	-3,3	0,0003	G protein-coupled receptor, family C, group 5, member A
C1QTNF3	114899	-3,3	0,0073	C1q and tumor necrosis factor related protein 3
TMEM158	25907	-3,1	0,0052	transmembrane protein 158 (gene/pseudogene)
PERP	64065	-3,0	0,0009	PERP, TP53 apoptosis effector
GPER	2852	-3,0	0,0019	G protein-coupled estrogen receptor 1
SLCO5A1	81796	-3,0	0,0029	solute carrier organic anion transporter family, member 5A1
GPR26	2849	-2,7	0,0324	G protein-coupled receptor 26
CD36	948	-2,6	0,0022	CD36 molecule (thrombospondin receptor)
CLDN1	9076	-2,5	0,0234	claudin 1
GABRA1	2554	-2,5	0,0017	gamma-aminobutyric acid (GABA) A receptor, alpha 1
LOC440292	440292	-2,5	0,0002	uncharacterized LOC440292
OR51A2	401667	-2,5	0,0232	olfactory receptor, family 51, subfamily A, member 2
MSLN	10232	-2,5	0,0095	mesothelin
LRIT2	340745	-2,5	0,0183	leucine-rich repeat, immunoglobulin-like and transmembrane domains 2
APCDD1L	164284	-2,4	0,0131	adenomatosis polyposis coli down-regulated 1-like
CNIH3	149111	-2,4	0,0103	cornichon homolog 3 (Drosophila)
DLK2	65989	-2,3	0,0245	delta-like 2 homolog (Drosophila)
SUN5	140732	-2,3	0,0030	Sad1 and UNC84 domain containing 5
MAL	4118	-2,3	0,0302	mal, T-cell differentiation protein
COLEC12	81035	-2,3	0,0432	collectin sub-family member 12
SYNPR	132204	-2,2	0,0315	synaptoporin
TMEFF2	23671	-2,2	0,0224	transmembrane protein with EGF-like and two follistatin-like domains 2
SCN9A	6335	-2,2	0,0243	Na^+ channel, voltage-gated, type IX, alpha subunit
GJA8	2703	-2,2	0,0069	gap junction protein, alpha 8, 50kDa
KCNMA1	3778	-2,2	0,0049	K^+ large conductance Ca^{2+} -activated channel, subfamily M, alpha member 1
C10orf111	221060	-2,1	0,0145	chromosome 10 open reading frame 111
OR5M11	219487	-2,1	0,0110	olfactory receptor, family 5, subfamily M, member 11
KLRG1	10219	-2,1	0,0032	killer cell lectin-like receptor subfamily G, member 1
KCNS3	3790	-2,1	0,0291	K^+ voltage-gated channel, delayed-rectifier, subfamily S, member 3
LPAR1	1902	-2,0	0,0232	lysophosphatidic acid receptor 1
SYT13	57586	-2,0	0,0075	synaptotagmin XIII

Table S6: AF-S marker genes (compared to AF-nS). [A]: ENTREZ Gene ID; [B]: Fold Change (AF-nS/AF-S); [C]: adjusted p value.

Gene	Gene name	Forward primer	Reverse primer
CA12	Carbonic anhydrase XII	AAAGGCCAGGAAGCATTCGT	CCCCGGTAGCGGTAATATTCA
CD24	CD24 Antigen	CCACGCAGATTTATTCCAGTGA	GCCAACCCAGAGTTGGAAGTAC
CLDN11	Claudin-11	TGGCTGGTGTTTTGCTCATTC	CGCACACAGGGAACCAGAT
COL12A1	Collagen type XII, collagen α1	TGACAACCCTTTCCGACACA	CTCCTCACGGTTCTAAAATTTGC
FOXF1	Forkead box F1	CCCACACAGGAATTCTGCTGA	TTCCCCCACTTCTGCCATT
HES1	hairy and enhancer of split-1	AGGCGGACATTCTGGAAATG	CGGTACTTCCCCAGCACACTT
JAG1	Jagged 1	CCGGCGTGCTGGGTAGAG	CGGCCGCAGGTAACACAAT
LPPR4	Lipid phosphate phosphatase-related 4	TGGTAGCAGCAGTGATGGAATT	GAGCTAGCATCTCTGTGGTTTCG
LRP5	Low density lipoprotein receptor-related protein 5	AACATCAAGCGAGCCAAGGA	CGGCTGTAGATGTCGATGCT
NOTCH3	Notch homolog 3 (Drosophila)	TCTCAGACTGGTCCGAATCCAC	ACACTTGCCTCTTGGGGGTAAC
PDGFRA	Platelet-derived growth factor receptor, alpha	CTTATTGTCCTGGTTGTCATTTGG	CAATGACCCTCCAGCGAATT
PTGFRN	Prostaglandin F2 receptor negative regulator	GTTCAACCCCCAGCACAGAT	CTTTAACCTGCACTGTGTCCTCAT
PTPRF	Receptor-type tyrosine-protein phosphatase F	CCGACCTGGCGGACAAC	GTTTGAATTCTCCCACGTGAACT
RGMB	RGM domain family, member B	CCAACAGCCAGCCCAATG	TGAGAAGTCAGGGACACGAAGTC
SFRP2	Secreted frizzled-related protein 2	TGTGCCACGGCATCGA	TCGTGGCCCAGCAGGTT
SHISA2	Shisa Homolog 2 (Xenopus Laevis)	CGTGCCGTTCCTCATTGTT	CCAGGGACCCCAAGATGATA
PPIA	Cyclophillin A	CCTGCTTCCACCGGATCAT	CGTTGTGGCGCGTAAAGTC

Table S7: qPCR primer sequences

Figure 1

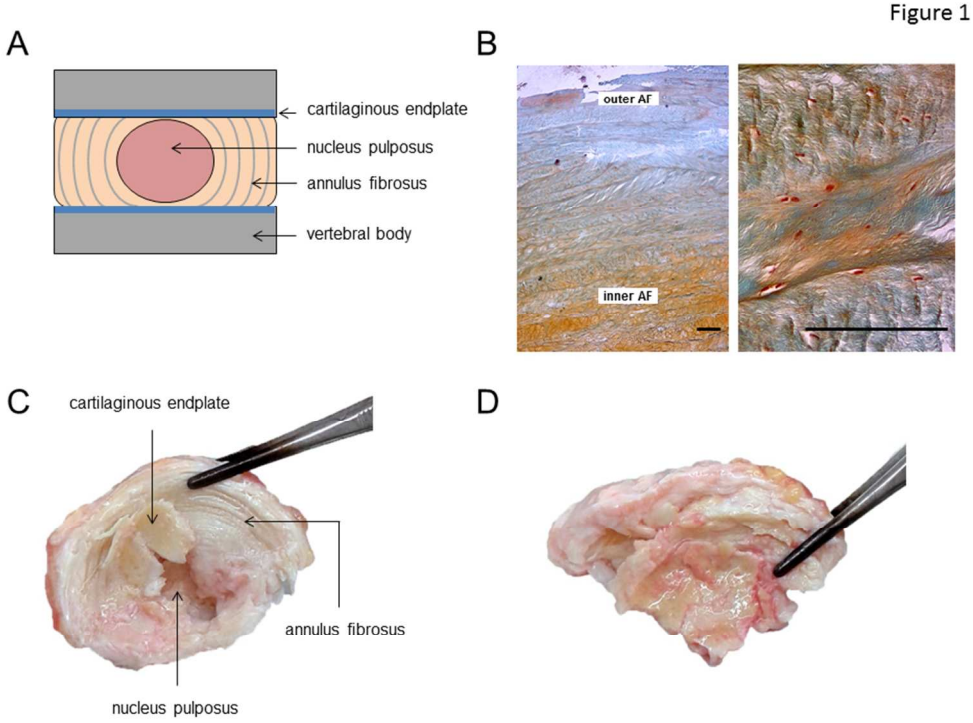

Figure 1: Intervertebral disc morphology and degenerative disc disease. A) Schematic representation of an intervertebral disc. The IVD comprises three distinctive tissues: a central gelatinous nucleus pulposus (NP), surrounded by the lamellar annulus fibrosus (AF). The AF consists of multiple, concentrically arranged lamellae in which cells and ECM components interact to provide function and structure. In the AF two distinct zones are distinguished: the inner and outer AF, primarily based on their collagen/proteoglycan content. The NP and AF are flanked at each side by cartilaginous endplates (CEP) which mediate the contact between the IVD body and vertebra. B) Sections of a healthy AF (15 years old male donor) stained with Safranin O and Fast green. The outer AF predominantly contains collagens while the inner AF contains proteoglycans. The characteristic lamellar structure is clearly visible; bar represents 200 µm. C) A non-degenerate L1/L2 intervertebral disc and of 63 years old male donor. The lamellar annulus fibrosus is clearly recognizable. The central nucleus pulposus displays no aberrant morphology; the cartilaginous endplate is relatively thin. D) The L4/L5 intervertebral disc of the same donor. Degenerative changes cause apparent loss of lamellar structure in the AF. The NP is no longer distinguishable from the AF and additional ossification is found near the CEP.

254x190mm (96 x 96 DPI)

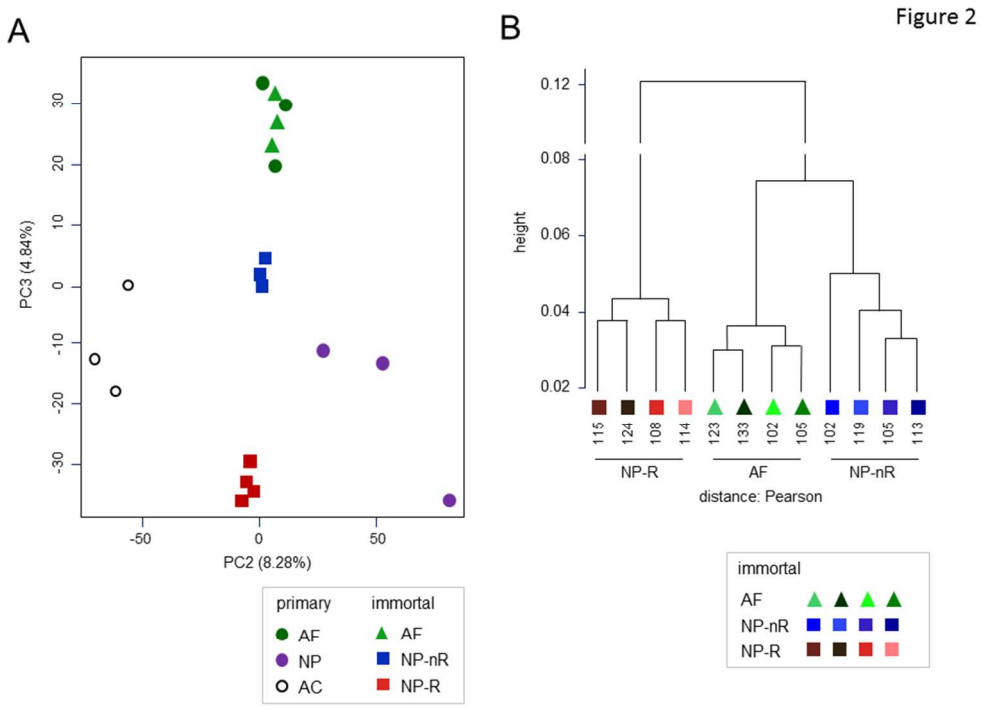

Figure 2: Principle component analysis reveals overlap between primary and immortalized NP and AF clones. Transcriptome comparison between immortal cell clone and primary cell datasets generated on Affymetrix and Illumina platforms (1st component), respectively. PCA analysis was preceded by ranking genes (descending expression value; see: Supplemental Methods for detailed description). A) PCA plot of the 2nd and 3rd component of primary AC, NP and AF samples from three independent healthy donors (circles) and immortalized NP-Responder (n=4), NP non-Responder (n=3) and AF (n=3) cell clones from a single donor (squares); these clones have been described in previous reports (18, 20). The 1st component separated the dataset on platform (primary cells: Affymetrix platform, immortalized cell lines: Illumina platform). B) Dendrogram showing the hierarchical clustering of immortalized NP-Responder (n=4), NP non-Responder (n=4) and AF (n=4) cell clones; clustering was based on Pearson correlation. The NP-R, AF, NP-nR clusters clearly differ. NP-nR clone 102 fell outside the NP-nR cluster (Supplemental Figure S2). As this effect persisted in subsequent analyses, this clone was omitted from further analysis. Analyses were performed and figures generated using the open source scripting language R (version 2.13.0) and R packages of Bioconductor 2.8.

254x190mm (96 x 96 DPI)

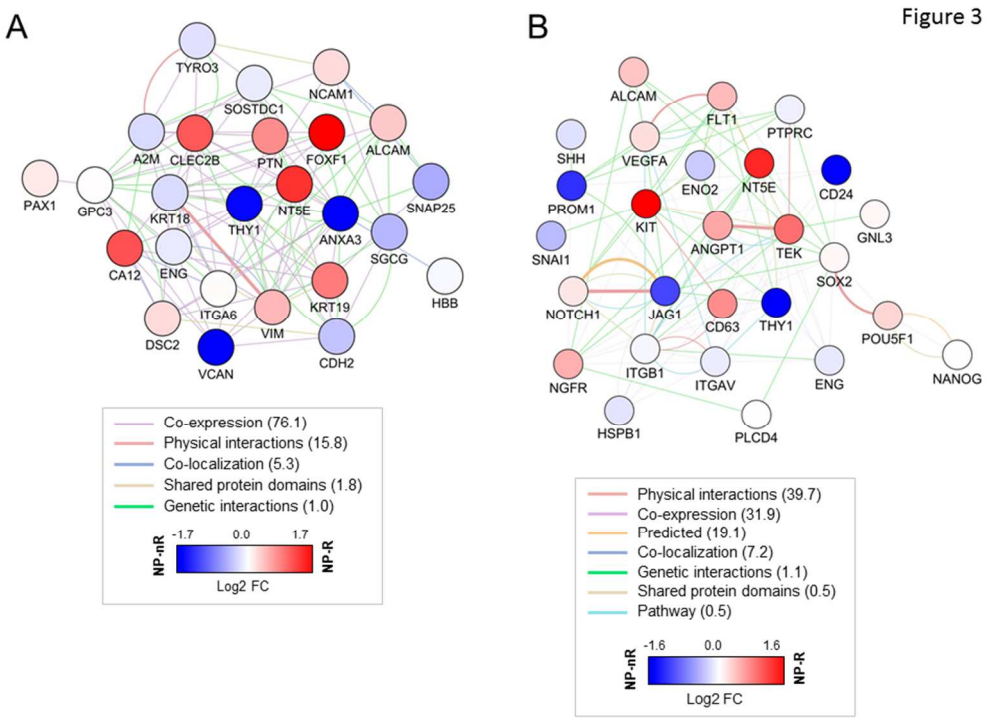

Figure 3: Differential expression of published NP markers among NP cellular subtypes. A) Selected marker genes (NP cell markers; Supplementary Table S1, S2) were used as input for GeneMANIA network building using Cytoscape visualization software. Gene expression differences between NP-R (red) and NP-nR (blue) clones are depicted as log2-based fold change (FC) in colored circles. Colored lines in the network represent established biological connections between markers; parentheses: percentage representation. The network is strongly inter-connected based on co-expression and genetic interactions studies. Of note: established NP markers are not equally expressed by isolated cellular sub-populations derived from the NP. B) Progenitor-associated gene expression in NP-Responder clones. Selected genes (NP progenitor markers; Table 1, Supplementary Table S2) were used as input for network building using GeneMANIA and Cytoscape visualization software. Gene expression differences between NP-R (red) and NP-nR (blue) clones are depicted as log2-based fold change (FC) in colored circles. Colored lines in the network represent established biological connections between markers; parentheses: percentage representation.

254x190mm (96 x 96 DPI)

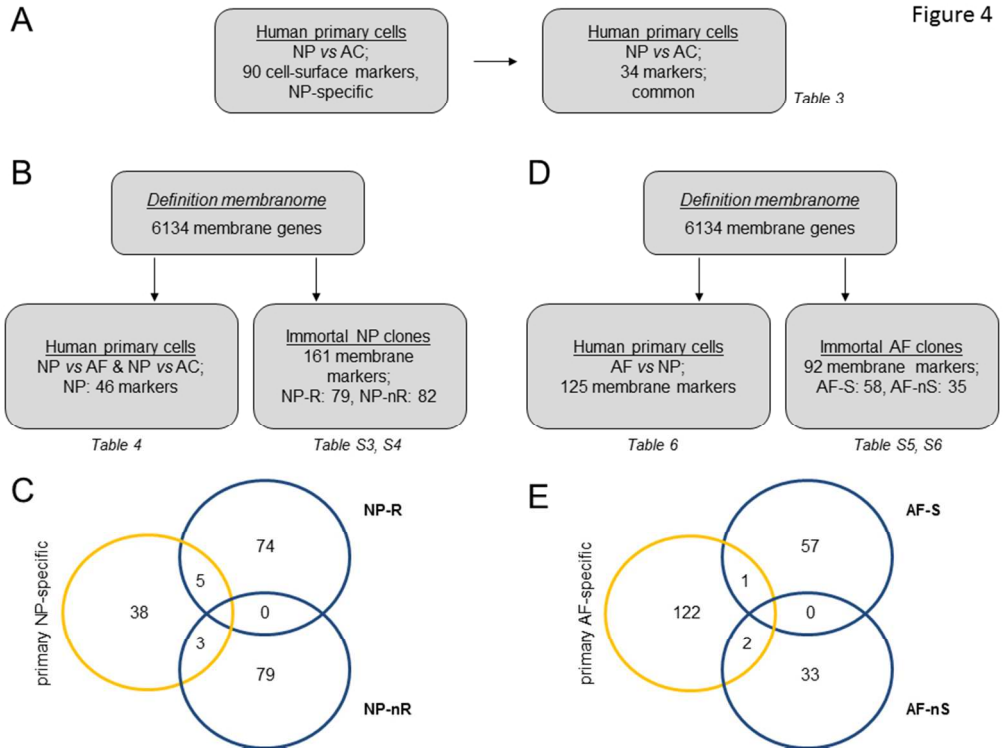

Figure 4

Figure 4: Identification of membrane-associated NP marker genes. Comparison strategy of whole transcriptome datasets of published primary IVD cellular isolates and clonal cell lines. A) Marker genes significantly more highly expressed in primary NP compared to AC from two independent published datasets (26, 27), were compared. B) Definition of primary NP-specific membrane marker genes compared to AF and AC (26) and definition of NP cell clone (R/nR) specific marker genes (18). C) Subsequently both primary and clonal NP membrane marker gene sets were compared to each other. D) Definition of primary AF-specific membrane marker genes compared to NP and definition of AF cell clone (S/nS) specific marker genes. E) Subsequent comparison of primary and clonal AF membrane marker genes.

254x190mm (96 x 96 DPI)

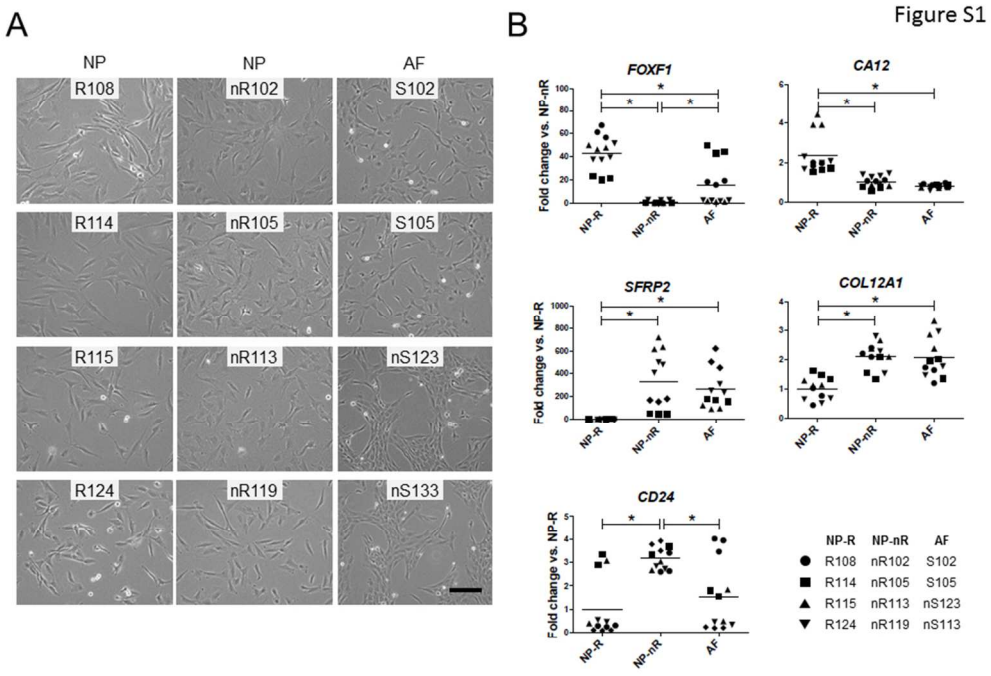

Figure S1: Morphological characteristics of immortalized cell clones. A) Phase contrast images of NP-R, NP-nR and AF exponentially growing cell cultures prior to RNA isolation; bar represents 20 micron. B) Gene expression analysis of FOXF1, CA12, COL12A1, SFRP2 and CD24 in indicated cell clones (n=3 biological repeats per clone). Consistent with earlier observations, FOXF1 and CA12 expression was highest in NP-R clones, while CD24 mRNA levels were highest in NP-nR clones (9). SFRP2 was strongly expressed in AF and NP-nR clones and was nearly undetectable in NP-R clones. COL12A1 was significantly higher expressed in both AF and NP-nR clones compared to NP-R clones. Single dots represent a single measurement; horizontal lines represent the mean of all measurements. Correspondence of symbols to clone identity is indicated in the table (bottom right). Gene expression was normalized to Cyclophillin A. Statistical significant differences are indicated by an asterisk (two-tailed Student's t-test; p value <0.05).

254x190mm (96 x 96 DPI)

Figure S2

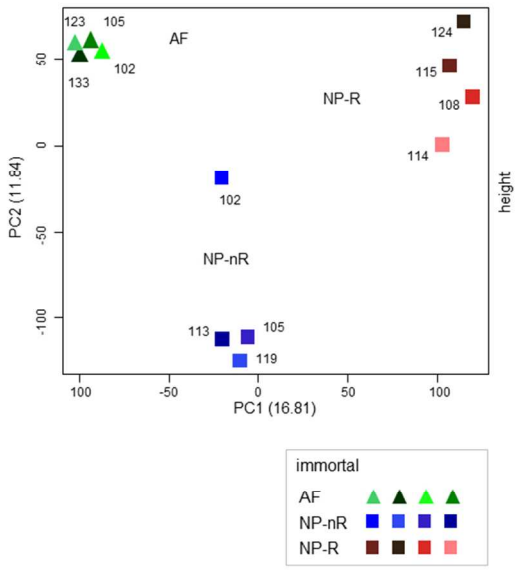

Figure S2: Principle component analysis clusters immortalized NP and AF clones. A) PCA plot of immortalized NP-Responder (n=4; red square symbols), NP non-Responder (n=4; blue square symbols) and AF (n=4; green triangular symbols) cell clones. NP-nR clone 102 was considered an outlier based on these analyses and additional gene expression measurements and was not included in subsequent statistical analyses.

254x190mm (96 x 96 DPI)

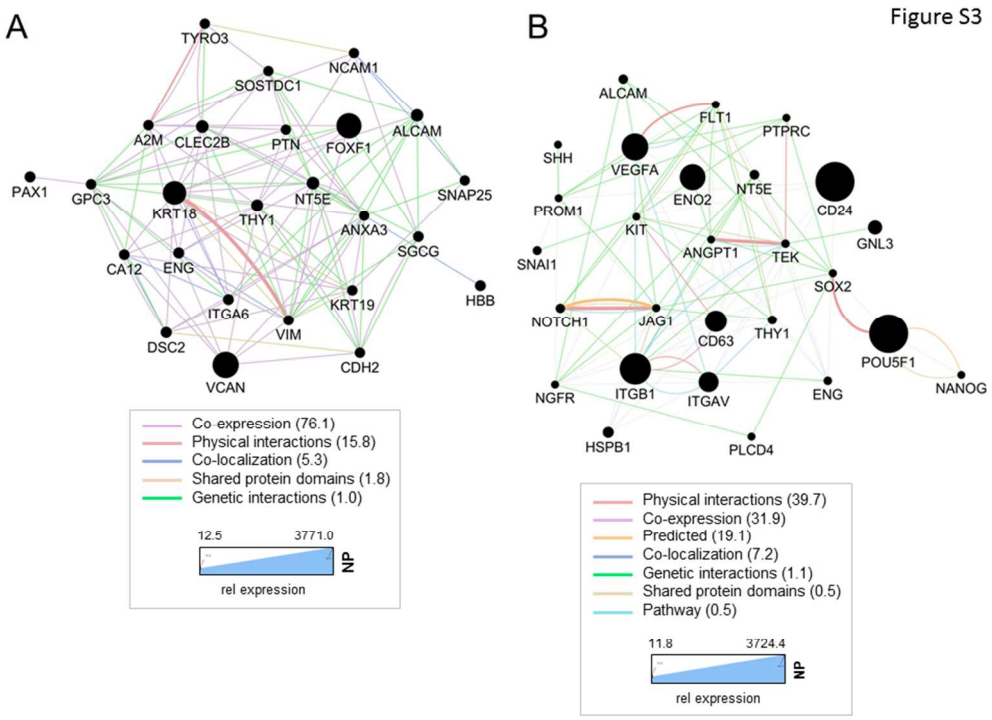

Figure S3: Published NP marker expression in primary NP cells. The gene networks depicted in Figure 3 were used to visualize expression levels of corresponding genes in primary NP cell isolates for either NP marker genes (A) or NP progenitor-associated genes (B). A) Selected NP cell marker gene network. Circle-size indicates an arbitrary log2-based expression value (i.e. based on fluorescence intensity on the micro-array); a log2-based expression value below 6.0 indicates low/absent gene expression (i.e. not expected to be biologically relevant). Colored lines in the network represent established biological connections between markers; parentheses: percentage representation. The highly expressed genes (KRT18, FOXF1) are among the best known NP cell marker genes. B) NP progenitor-associated gene network. Gene network showing log2-based expression values of the indicated genes in primary NP cells (cf. Figure 3B). Circle-size indicates an arbitrary log2-based expression value (i.e. based on fluorescence intensity on the micro-array. VEGFA, ENO2, CD24, CD63, ITGB1, ITGAV and POU5F1 are strongly expressed by primary NP cells. Of note: expression of genes strongly linked to stem cells and NP progenitor cells is low/not detectable in primary isolates.

254x190mm (96 x 96 DPI)

Figure S4

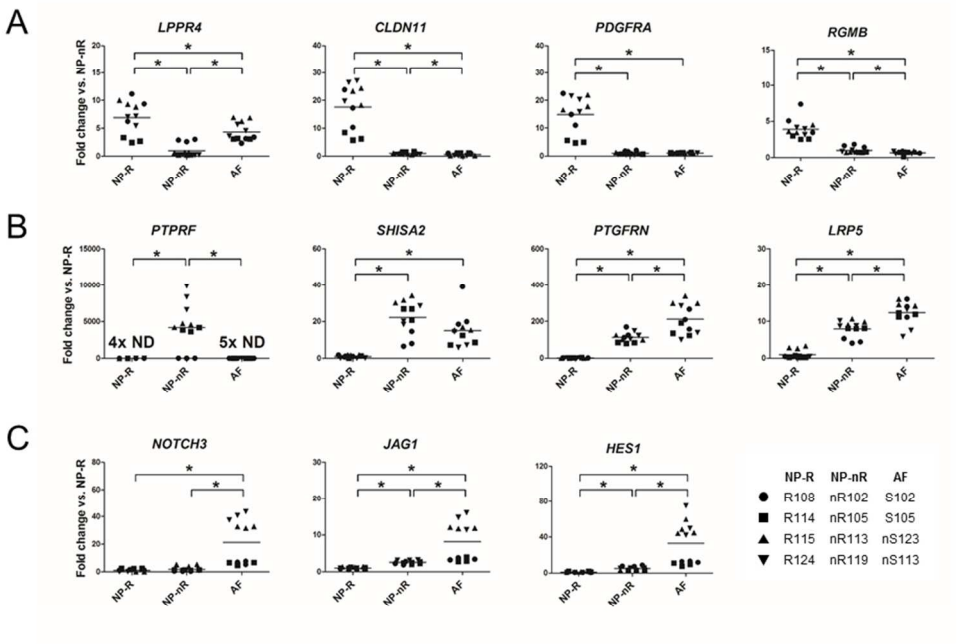

Figure S4: Validation of membrane-associated NP and AF cellular heterogeneity markers. A) rtPCR measurements for NP-R marker genes LPPR4, CLDN11, PDGFRA, and RGMB in four NP-R, NP-nR and AF clones under normal proliferation conditions. B) rtPCR measurements for NP-nR marker genes PTPRF, SHISA2, PTGFRN and LRP5 in four NP-R, NP-nR and AF clones. C) rtPCR measurements for AF-S and AF-nS marker genes NOTCH3, JAG1 and HES1. Gene selection was arbitrarily based on the fold expression difference and previous reports in literature. Correspondence of symbols to clone identity is indicated in the table (bottom right). Gene expression was normalized to the housekeeper gene Cyclophillin A. Statistical significant differences are indicated by asterisks (two-tailed Student's t-test; p value <0.05).

254x190mm (96 x 96 DPI)

Tables

tissue origin (ref)	ref	isolation method	markers
Human[A]; NP, AF	54	flow cytometry	**CD105**, CD166, CD63, **CD90**, **CD73** (NT5E), P75, CD133
Human; NP	48	IHC	HSP27, HSP72
Lapine, ovine, human[A]; NP	49	IHC	Notch1, Jag1, Delta4, Stro1, C-kit (CD117), **CD105**, **CD166**
Human; AF	50	flow cytometry	**CD24**, CD49, CD51, NSE
Human; CEP	52	CFU assay; agarose	CD133, **CD105**, **CD90**, **CD73** (NT5E), Stro1, CD166, **CD44**
Canine; NP	53	CFU assay	Sox2, Oct3/4, Nanog, CD133 (mRNA only)
Murine, human; NP	55	CFU assay; agarose	**CD24**, Tie2, **GD2**, **CD44**, (CD49f, CD56, CD73 (NT5E), CD90, CD105, CD166)[B]; (CD271, Flt1)[C]
Lapine; AF	51	CFU assay; fibroblast	**CD29**, **CD44**, CD166, Oct-4, Gnl3, SSEA-4

Table 1: Overview of literature reports on cell heterogeneity and progenitor populations in the intervertebral disc. Various markers have been used to isolate multipotent cell populations from IVD tissues (as indicated in the corresponding colomn). Typical MSC markers (CD73, CD90 and CD105) are often used in addition to CD24 and CD44. We measured gene and protein expression of genes printed in bold in immortal cell clones of a single donor (NP and AF). [A]: degenerate tissue; [B]: expressed on all mouse and human NP cells; [C]: correlated to progenitor phenotype.

	Progenitor		Differentiation			Immortal clones	
marker	dormant	activated	1	2	3	NP-R	NP-nR
Tie2	+	+	-	-	-	-	-
GD2	-	+	+	+	-	-	+
CD24	-	-	-	+	+	-	+
CD44	+/-	+	+	+	+	+	+
CD271	+	+	+/-	-	-	ND	ND
FLT1	+	+	+/-	-	-	ND	ND

Table 2: Positioning of immortal NP-phenotypes in NP-based progenitor/differentiation model. A recently postulated progenitor/differentiation model based on expression (+) or absence (-) of markers (55), is represented in the columns marked: *Progenitor* and *Differentiation*. The right most columns show the measurements we performed on our immortal NP cell clones; *ND*: analysis not done.

Gene	ID[A]	AC Mean[B]	NP Mean[B]	FC[C]
ABCG1	9619	20,3	739,7	36,4
AGPAT9	84803	675,2	2611,6	3,9
C1orf9	51430	1166	3352,1	2,9
CA12	771	71,3	2726,9	38,2
CDH19	28513	761,7	2760,6	3,6
CDS1	1040	76,9	249,5	3,2
CLDN11	5010	19,3	204,2	10,6
CLEC2B	9976	53,4	303,3	5,7
DNER	92737	365,6	3671,4	10
FAM62B	57488	78,5	196,5	2,5
FKBP11	51303	1102,8	2438,5	2,2
FLRT2	23768	21,1	240,5	11,4
FNDC3B	64778	385,9	1598,9	4,1
HS6ST3	266722	54,2	494,7	9,1
INSIG1	3638	885,9	2010,5	2,3
JPH3	57338	50,3	252,9	5
KCNS3	3790	61,2	175,8	2,9
MFAP3L	9848	49,2	128,8	2,6
MXRA7	439921	238,6	844,3	3,5
NETO2	81831	49,3	323	6,6
ODZ3	55714	36,3	217,6	6
ORAI3	93129	431,5	2239,2	5,2
PPAPDC1B	84513	774,6	2823,4	3,6
REEP1	65055	42,2	129,8	3,1
SEL1L	6400	190,6	1166,5	6,1
SGCG	6445	24,3	75,5	3,1
SLC27A2	11001	19,2	350	18,3
SLC44A1	23446	86,6	477,2	5,5
SLC7A1	6541	99,9	237	2,4
TMEM27	57393	19,3	44,2	2,3
TMEM71	137835	29,5	336,6	11,4
TSPAN31	6302	574	1177,6	2,1
WFS1	7466	484,1	1093,4	2,3
WNT11	7481	179,9	406,2	2,3

Table 3: Common human NP and AC markers. A comparison between 90 transmembrane cell surface markers (27) and primary tissue-specific markers (26) extracted from human micro-array data (NP *vs*. AC). [A]: ENTREZ Gene ID; [B]: adjusted p value <0.05, fold change >2.0 and Expression in NP >100 arbitrary units; see Methods section for details); [C]: Fold Change (NP/AC).

Gene	ID[A]	AC Mean[B]	AF Mean[B]	NP Mean[B]	FC[C]	FC[D]
ANKRD10	55608	687,4	909,1	2267,0	3,3	2,5
ANXA2	302	18,5	23,5	380,8	20,6	16,2
ATP6V0E1	8992	35,4	37,7	124,4	3,5	3,3
C11orf1	64776	50,3	48,5	178,3	3,5	3,7
C9orf91	203197	68,0	75,3	276,3	4,1	3,7
CA12	771	71,3	134,4	2726,9	38,2	20,3
CLDN11	5010	19,3	47,2	204,2	10,6	4,3
CLIP4	79745	98,1	63,9	207,2	2,1	3,2
DKFZp451A211	400169	505,2	989,2	2828,0	5,6	2,9
DUSP6	1848	533,6	680,5	2172,2	4,1	3,2
DYRK2	8445	238,8	247,3	746,3	3,1	3,0
EFNA1	1942	909,3	2144,1	6752,4	7,4	3,1
F3	2152	68,8	55,4	263,1	3,8	4,8
FAM134A	79137	173,5	188,9	408,8	2,4	2,2
FGFR2	2263	25,9	26,5	109,7	4,2	4,1
FLRT1	23769	79,5	90,6	243,3	3,1	2,7
FUS	2521	43,2	41,3	105,2	2,4	2,5
GPR153	387509	319,2	310,1	686,7	2,2	2,2
GRB10	2887	179,9	154,4	651,0	3,6	4,2
HS6ST3	266722	54,2	183,9	494,7	9,1	2,7
KHDRBS3	10656	164,8	289,8	1349,9	8,2	4,7
KIAA1217	56243	39,9	184,3	529,1	13,2	2,9
LDLR	3949	140,3	148,1	339,9	2,4	2,3
LRRFIP1	9208	398,6	499,9	1150,3	2,9	2,3
NCAM1	4684	49,2	43,7	116,8	2,4	2,7
NETO2	81831	27,7	48,6	122,4	4,4	2,5
NNMT	4837	89,5	52,7	263,0	2,9	5,0
ORAI3	93129	431,5	1058,3	2239,2	5,2	2,1
PGK1	5230	39,1	54,9	471,6	12,1	8,6
PSMB4	5692	1084,0	800,1	2256,8	2,1	2,8
RP1-21O18.1	23254	198,8	153,9	422,4	2,1	2,7
RPL38	6169	1114,4	863,3	6333,9	5,7	7,3
RPLP2	6181	234,9	194,1	739,6	3,1	3,8
RRAD	6236	234,2	247,8	745,4	3,2	3,0
SLC16A3	9123	33,5	42,1	169,9	5,1	4,0
SLC2A1	6513	353,9	851,8	2506,6	7,1	2,9
SLC2A3	6515	410,9	437,5	1330,5	3,2	3,0
SLC44A1	23446	86,6	208,1	477,2	5,5	2,3
SLC6A6	6533	264,9	185,9	569,8	2,2	3,1
STX3	6809	356,4	327,7	746,2	2,1	2,3
TMEFF2	23671	271,5	321,8	2501,6	9,2	7,8
TMEM200A	114801	61,7	53,1	183,0	3,0	3,4
TNS1	7145	2142,0	3537,6	7840,3	3,7	2,2
TRPM8	79054	31,6	65,0	269,2	8,5	4,1

ZNF395	55893	1956,3	4013,4	9607,8	4,9	2,4
ZNF791	163049	711,1	1397,7	3029,9	4,3	2,2

Table 4. NP-specific membrane-associated markers compared to AF and AC in primary cell expression datasets. [A]: ENTREZ Gene ID; [B]: Differentially expressed genes were determined by: adjusted p value <0.05, fold change >2.0 and Expression in NP >100 (arbitrary units; see Methods section for details) (26). Marker genes overlapping with those in Table 4 are printed in bold. [C]: Fold Change (NP/AC); [D]: Fold change (NP/AF).

Gene	ID[A]	FC[B]	*p*[C]	Gene description
CLDN11	5010	6,90	0,011	claudin 11 (oligodendrocyte transmembrane protein)
TMEFF2	23671	5,46	0,005	transmembrane protein with EGF-like and two follistatin-like domains 2
CA12	771	2,72	0,025	carbonic anhydrase XII
ANXA2	302	2,51	0,009	annexin A2
CD44	960	2,16	0,042	CD44 antigen (homing function and Indian blood group system)
EFNA1	1942	-3,24	0,003	ephrin-A1
NETO2	81831	-3,64	0,002	neuropilin (NRP) and tolloid (TLL)-like 2
SLC2A1	6513	-4,55	0,022	solute carrier family 2 (facilitated glucose transporter), member 1

Table 5. Differentially expressed NP-specific membrane-associated markers in NP-R and NP-nR cell clones. [A]: ENTREZ Gene ID [B]:Fold Change (NP-R/NP-nR); [C]: adjusted p value <0.05, fold change >2.0 and expression in NP >100 (arbitrary units; see Methods section for details).

Gene	Gene ID[A]	AF mean[B]	NP mean[B]	FC[C]
SLC40A1	30061	1241,8	148,2	8,4
RNFT1	51136	271,2	38,0	7,1
C18orf55	29090	369,5	66,6	5,5
COLEC12	81035	257,7	47,9	5,4
DAD1	1603	3141,1	584,2	5,4
LPAR1	1902	1546,4	294,7	5,2
IFITM2	10581	1411,7	297,8	4,7
ALG10B	144245	307,2	68,5	4,5
IFITM1	8519	995,7	222,5	4,5
THBD	7056	388,0	88,4	4,4
HEATR5A	25938	286,0	66,2	4,3
LRP6	4040	704,8	173,6	4,1
SLC38A9	153129	249,9	62,3	4,0
ATP11C	286410	397,1	100,6	3,9
FLRT2	23768	688,3	175,1	3,9
GPR125	166647	374,9	96,5	3,9
SLC30A5	64924	115,6	30,1	3,8
GPR137B	7107	232,0	61,0	3,8
TMEM133	83935	331,6	88,7	3,7
CD55	1604	3992,5	1074,7	3,7
EFHA2	286097	149,8	41,6	3,6
LRIG1	26018	1224,4	340,5	3,6
SLC35A5	55032	486,7	135,8	3,6
ANKRD46	157567	778,0	220,9	3,5
CD58	965	237,3	68,8	3,4
TMEM27	57393	151,3	44,2	3,4
EMP3	2014	868,2	254,0	3,4
CD59	966	333,2	97,9	3,4
FAM69A	388650	348,7	102,9	3,4
TACSTD2	4070	2031,0	603,1	3,4
C5orf28	64417	195,1	58,0	3,4
SCNN1A	6337	513,8	153,3	3,4
SLC10A7	84068	388,9	116,1	3,4
LRIG2	9860	524,8	158,6	3,3
SMC3	9126	957,8	296,4	3,2
ATP2C1	27032	122,2	38,4	3,2
FLJ20160	54842	236,1	76,0	3,1
LIFR	3977	447,5	144,3	3,1
FLJ11171	55783	216,9	72,3	3,0
OMA1	115209	327,8	111,3	2,9
C9orf123	90871	212,5	72,9	2,9
ITM2C	81618	1483,2	514,1	2,9
ABHD2	11057	854,9	297,1	2,9
EMP2	2013	999,4	350,1	2,9
CD99	4267	606,0	213,8	2,8
IL13RA1	3597	169,0	59,7	2,8
C20orf199	441951	1288,8	466,2	2,8

LOC440993	440993	276,4	100,4	2,8
ADAM9	8754	229,3	85,2	2,7
MGC24039	160518	131,7	49,0	2,7
TNFSF10	8743	290,6	109,2	2,7
PCNX	22990	168,4	63,7	2,6
TMEM200B	399474	300,5	113,7	2,6
ZMYM6	9204	220,5	83,9	2,6
SEPSECS	51091	131,7	50,6	2,6
FAT4	79633	213,8	82,8	2,6
LPAR4	2846	212,2	82,3	2,6
LOC493869	493869	206,8	80,9	2,6
AGTRAP	57085	136,9	54,0	2,5
CHIC1	53344	385,5	152,1	2,5
CLCN3	1182	174,6	69,1	2,5
STX7	8417	194,0	77,0	2,5
SLC39A7	7922	782,0	312,4	2,5
ACVR1	90	3337,6	1337,0	2,5
DISP1	84976	178,9	73,0	2,4
REEP1	65055	316,9	129,8	2,4
SLC35B3	51000	1395,7	572,9	2,4
SLC39A14	23516	344,9	142,4	2,4
ROBO1	6091	606,3	251,1	2,4
SLC41A2	84102	272,7	113,4	2,4
FDFT1	2222	593,3	247,6	2,4
MIA3	375056	362,0	151,4	2,4
HIGD1A	25994	2291,2	960,2	2,4
ITGB1	3688	221,4	92,9	2,4
SLC33A1	9197	1725,7	731,3	2,4
UPK1B	7348	160,2	68,0	2,4
SLC16A3	9123	183,0	77,6	2,4
FAS	355	192,3	82,1	2,3
C20orf30	29058	217,5	92,9	2,3
SLC30A6	55676	365,7	156,2	2,3
PTPLAD2	401494	108,2	46,3	2,3
LRP12	29967	129,7	55,8	2,3
STEAP2	261729	312,0	134,9	2,3
FAM73A	374986	365,2	158,3	2,3
LAG3	3902	330,2	144,4	2,3
KCNA1	3736	167,5	73,6	2,3
ULBP2	80328	184,1	81,2	2,3
ATL3	25923	1223,0	540,4	2,3
ATP11B	23200	236,3	104,9	2,3
PROCR	10544	568,0	252,3	2,3
C7orf44	55744	237,1	105,6	2,2
ASAM	79827	195,6	87,3	2,2
C14orf101	54916	230,7	103,0	2,2
TSPAN4	7106	115,5	51,7	2,2
SLC31A2	1318	545,7	246,3	2,2

SACM1L	22908	1277,7	577,6	2,2
FZD2	2535	197,5	89,6	2,2
TM2D2	83877	1033,3	469,6	2,2
USMG5	84833	3457,8	1573,8	2,2
MFSD8	256471	161,8	73,7	2,2
JAM2	58494	318,1	144,9	2,2
ARMCX5	64860	273,9	125,2	2,2
AXL	558	1232,8	566,4	2,2
STX17	55014	421,2	194,1	2,2
LNPEP	4012	215,2	99,5	2,2
DLEU2	8847	406,6	188,7	2,2
LOC219854	219854	395,2	183,6	2,2
C9orf105	401505	649,6	303,2	2,1
GHITM	27069	649,2	303,7	2,1
SLC16A7	9194	237,6	112,6	2,1
TMEM154	201799	172,1	81,9	2,1
EGFR	1956	253,3	121,0	2,1
TMEM19	55266	113,9	55,0	2,1
CAPRIN1	4076	308,4	149,0	2,1
SEL1L	6400	889,5	433,1	2,1
IFNAR1	3454	197,7	96,8	2,0
SPA17	53340	125,9	61,8	2,0
GLIPR1	11010	850,1	417,6	2,0
CNIH4	29097	2894,7	1424,1	2,0
KCNK1	3775	582,0	287,8	2,0
LOC203547	203547	1509,4	747,3	2,0
IL17RD	54756	149,7	74,3	2,0
INTS7	25896	102,8	51,2	2,0
DPY19L4	286148	283,4	141,6	2,0
SLC13A4	26266	1083,4	541,6	2,0

Table 6: Expression of AF-specific membrane-associated markers genes compared to NP. [A]: ENTREZ Gene ID; [B]: adjusted p value <0.05, fold change >2.0 and expression: >100 (arbitrary units; see Methods section for details); [C]: Fold Change (AF/NP).

Gene	Gene ID[A]	FC[B]	*p*[C]	Gene description	FC[D]
CHIC1	53344	2,3	0,0017	cysteine-rich hydrophobic domain 1	2,5
COLEC12	81035	-2,3	0,0432	collectin sub-family member 12	5,4
LPAR1	1902	-2,0	0,0232	lysophosphatidic acid receptor 1	5,2

Table 7: Differentially expressed AF-specific membrane-associated markers in AF-S and AF-nS cell clones. [A]: ENTREZ Gene ID; [B]: Fold Change (AF2/AF1); [C]: adjusted p value <0.05, fold change >2.0 and expression in AF >100 (arbitrary units; see Methods section for details). [D]: Fold Change (AF/NP).