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