

Supplemental Information

Primary Mesenchymal Stem and Progenitor Cells from Bone Marrow Lack Expression of CD44

Hong Qian^{1,2*}, Katarina Le Blanc², Mikael Sigvardsson¹

Supplemental data:

Figure S1: related to Figure 3;

Figure S2: related to Figure 4A;

Figure S3: related to Figure 4B;

Figure S4: related to Figure 7;

Figure S5: related to Figure 7;

Table S1-S3, related to Figure 4B.

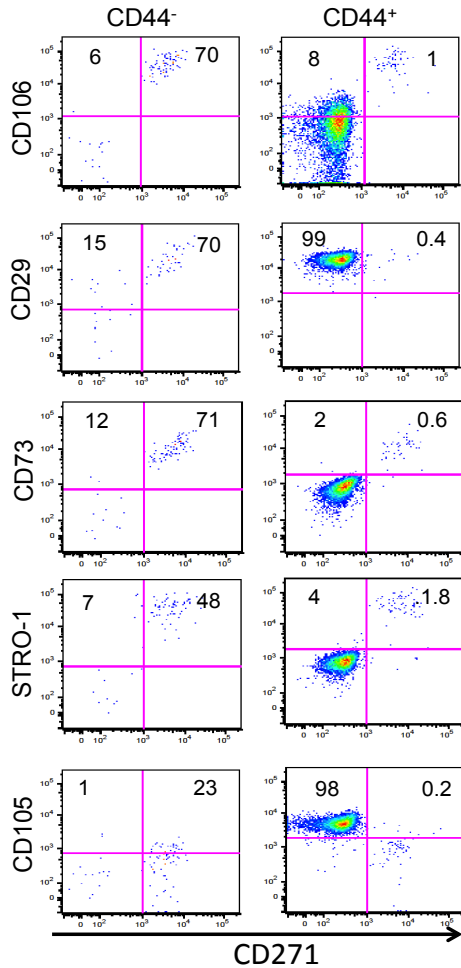
Supplemental materials:

Antibodies used for flow cytometry.

Probes used for Q-PCR.

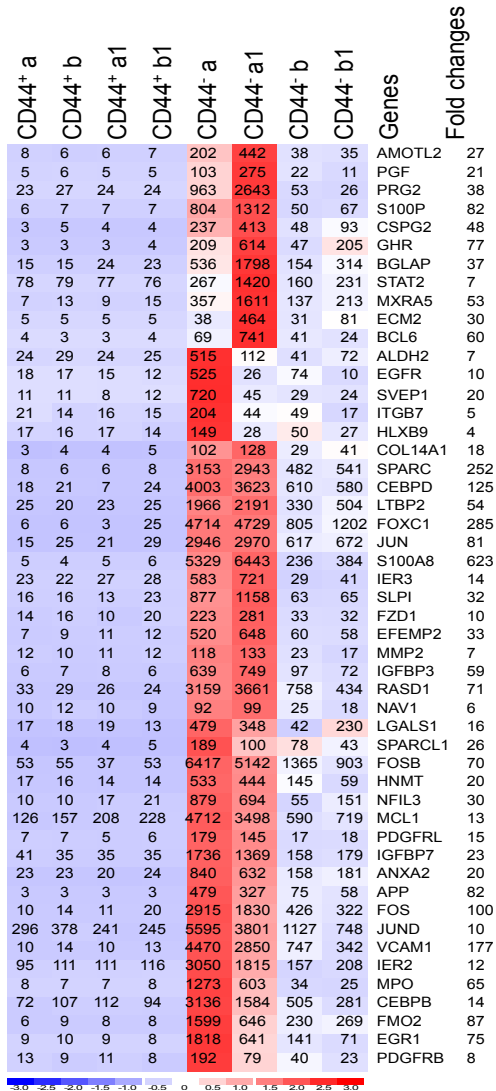
Supplemental Data:

Fig. S1. (Qian, et al)



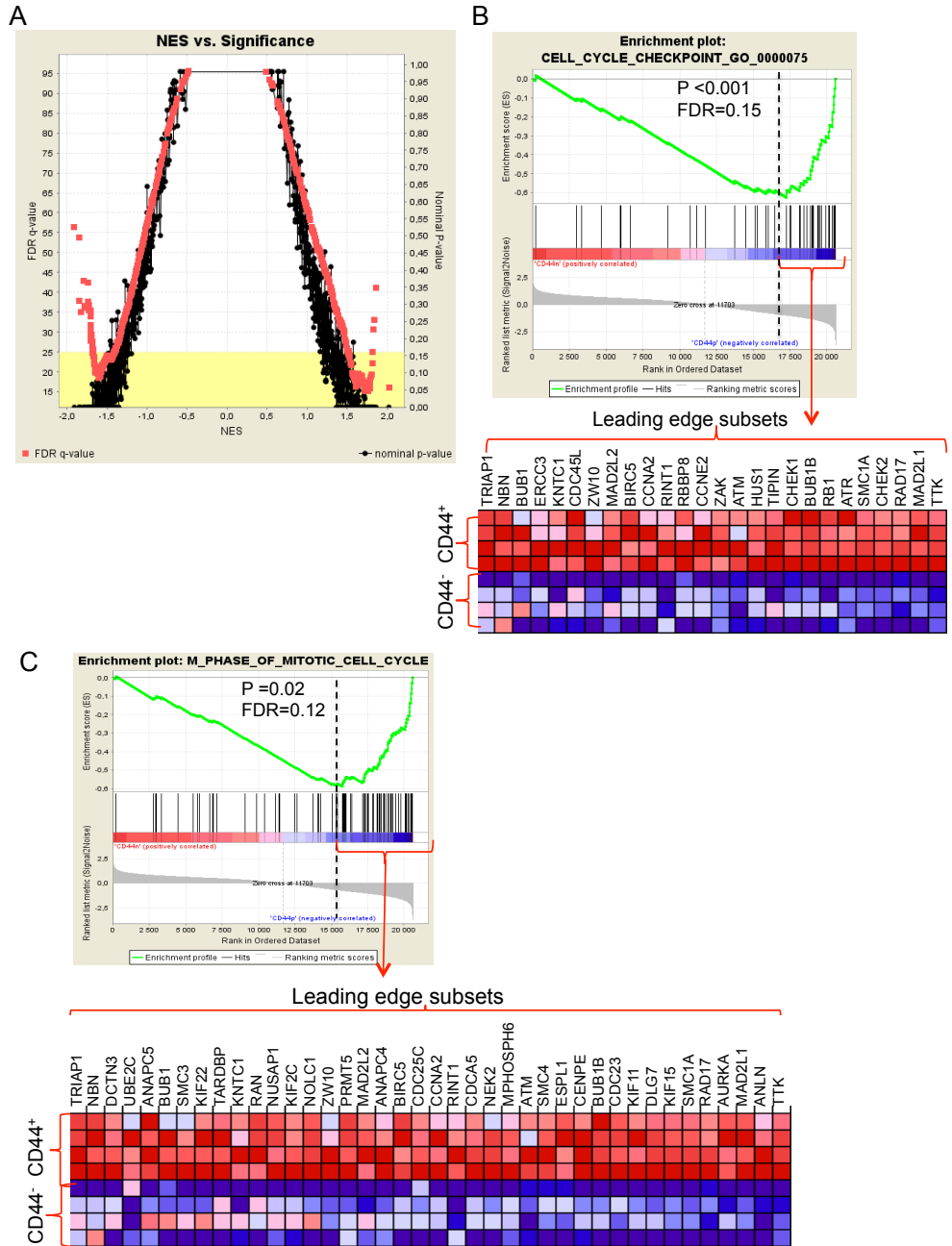
Supplemental Figure 1. FACS analysis of STRO1, CD73, CD29, CD105 and CD106 expression on the CD44⁻ mesenchymal cells of human bone marrow. The stromal cells were first enriched by MACS using CD45 and CD235 microbeads. The enriched cells were stained with fluorochrome conjugated antibodies against CD45, CD235, CD31, CD44, CD271 and the respective cell surface markers. The CD44⁺ and CD44⁻ cells were gated within CD45⁻CD235⁻CD31⁻ live cells. The data were from 3 independent experiments with 3 healthy young adult donors. The numbers in the panels are mean % of the CD44⁺ or CD44⁻ cells.

Fig. S2. (Qian, et al)



Supplemental Figure 2. Up-regulation of MSC-related genes, oncogenes and anti-apoptotic genes in the CD44⁻ cells. dCHIP analysis of global gene expression in the freshly sorted CD44⁺ cells and CD44⁻ cells from normal human BM. Clustering shows some of the genes being up-regulated more than 4-fold in the CD44⁻ cells compared to the CD44⁺ cells. Red represents high and blue low expression. Data were from 2 independent sorting experiments, 4 microarray experiments. The numbers in the heat map are expression values of each indicated genes, normalized by dCHIP software. Related to Figure 4.

Fig. S3 (Qian, et al.,)



Supplemental Figure 3. GSEA analysis of cell-cycle regulator genes in the CD44⁺ cells.

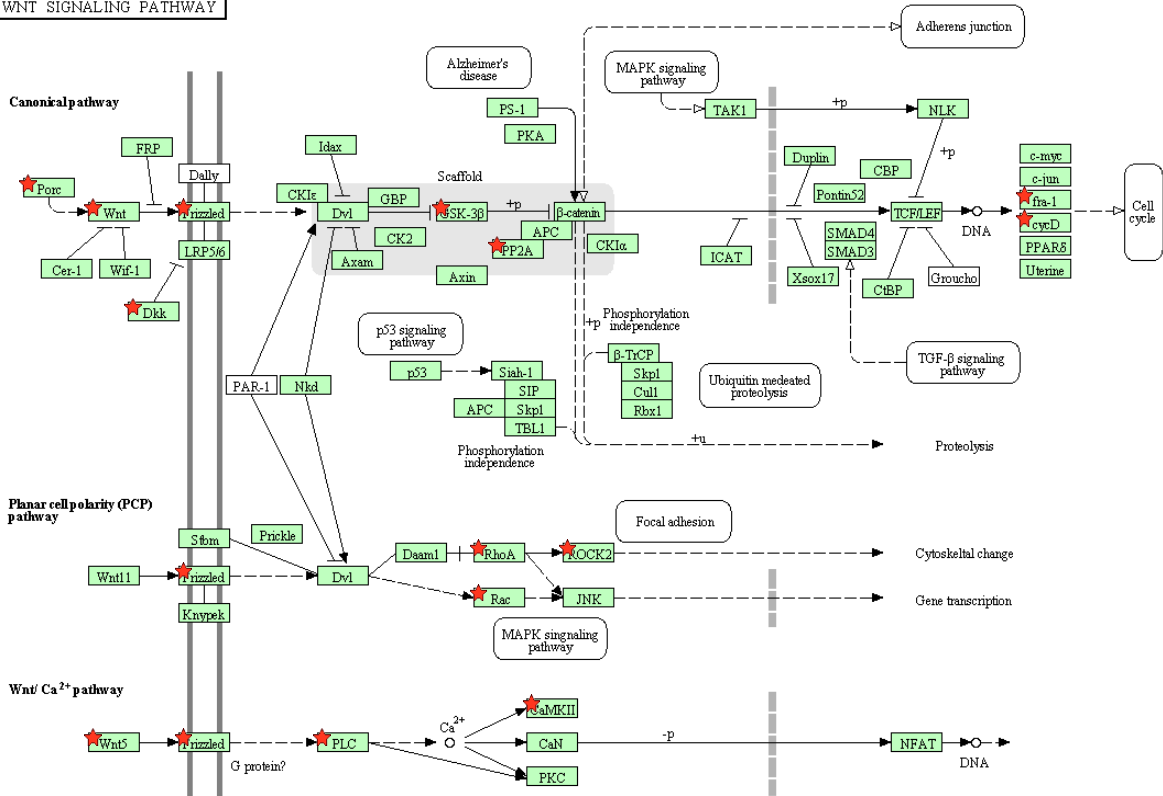
(A) Plot of p values versus normalized enrichment scores (NES) shows the number of enriched gene sets that are significantly different between the two cell populations. False discovery rate (FDR)-q value is the estimated probability that the normalized enrichment score represents a false positive finding. NES is the enrichment score for the gene set after it has been normalized across analyzed gene sets. (B-C) Enrichment of cell cycle genes in the CD44⁺ cells. (B) Enrichment of the genes in the gene set for M phase mitotic cell cycle. (C) Enrichment of genes in the gene set

of cell cycle checkpoint in the CD44⁺ cells. The green curves plot the ES. Black vertical dashed lines specify the maximum ES scores. Within each gene set, the farther the position of a gene to the left (red) implies a higher correlation with CD44 negative phenotype, and the farther to the right (blue) implies a higher correlation with genes down-regulated upon CD44 expression. Significantly enriched data sets are defined according to GSEA default settings ($p < 0.001$ and $FDR < 0.25$). The heatmap shows expressions of the genes in the leading edge subsets (only the top 40 genes were shown for clarity if more than 40 genes in the leading edge subsets). Data were from 8 microarray platforms and 2 independent experiments. The data were normalized by RMAExpress software. Related to Figure 4.

Fig. S4 (Qian, et al.,)

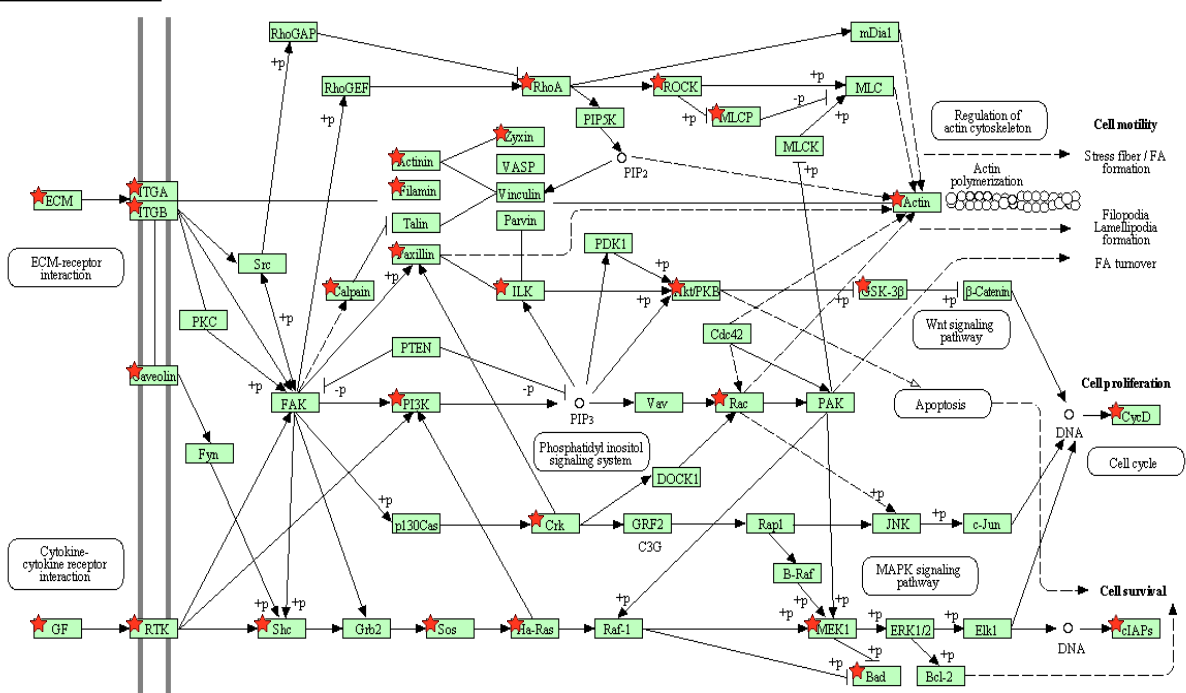
A

WNT SIGNALING PATHWAY



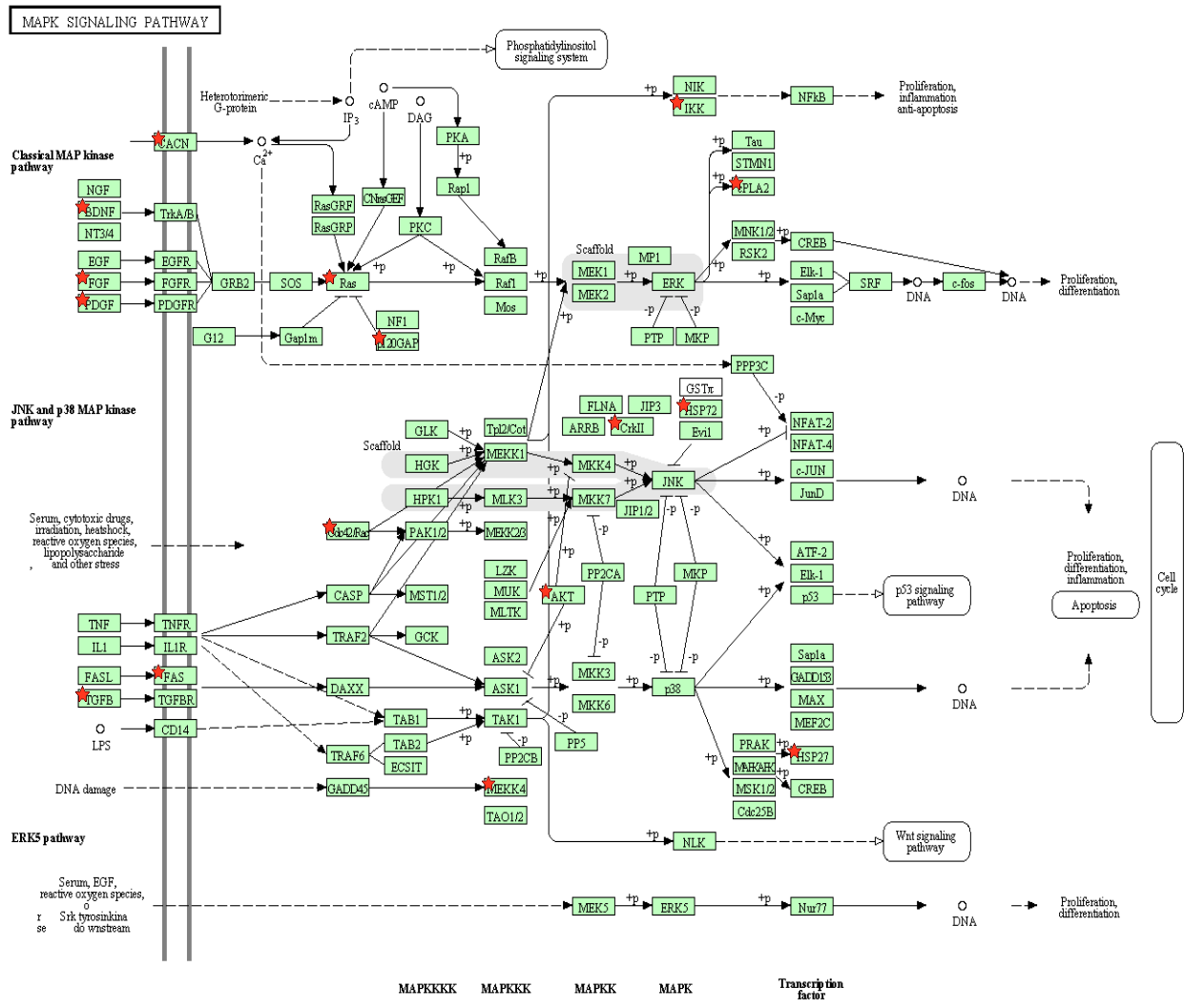
B

FOCAL ADHESION



Supplemental Figure 4. The biological pathways analysis of the genes in culture-expanded human BM MSCs. The genes from microarray data that are more than 5-fold upregulated in the culture-expanded MSCs compared to the freshly sorted CD44⁺ cells were used for mapping KEGG pathways using the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics resources v6.7 (<http://david.abcc.ncifcrf.gov/tools.jsp>). (A), WNT signal pathway. (B) Focal adhesion pathway. Red stars indicate the genes are more than 5 fold increase in the cultured MSCs compared to the freshly sorted CD44⁺ cells.

Fig. S5 (Qian, et al.,)



Supplemental Figure 5. Upregulation of the genes in MAPK pathway in culture-expanded human BM MSCs. The genes from microarray data that are more than 5-fold upregulated in the culture-expanded MSCs compared to the freshly sorted CD44⁻ cells were used for mapping MAPK signal pathway analyzed using the DAVID Bioinformatics resources v6.7. Red stars indicate the genes are more than 5 fold increase in the cultured MSCs compared to the freshly sorted CD44⁻ cells.

Supplemental Table 1. GSEA analysis of enrichment of gene sets that are positively correlated to CD44⁺ phenotype. Related to Figure 4.

NAME	SIZE	ES	NES	NOM p-value	FDR q-value
RESPONSE_TO_WOUNDING	180	0.28	1.74	0.00	0.15
HSA04510_FOCAL_ADHESION	194	0.20	1.72	0.00	0.15
EXTRACELLULAR_REGION	425	0.30	1.77	0.00	0.15
EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT	25	0.49	1.74	0.00	0.15
POSITIVE_REGULATION_OF_METABOLIC_PROCESS	213	0.11	1.75	0.00	0.16
CATION_HOMEOSTASIS	103	0.23	1.78	0.00	0.16
HSA00641_3_CHLOROACRYLIC_ACID_DEGRADATION	15	0.54	2.02	0.00	0.16
AMINE_RECEPTOR_ACTIVITY	34	0.46	1.74	0.00	0.16
TRANSMEMBRANE_RECEPTOR_PROTEIN_KINASE_ACTIVITY	51	0.38	1.72	0.00	0.16
POSITIVE_REGULATION_OF_CYTOKINE_BIOSYNTHETIC_PROCESS	25	0.37	1.77	0.00	0.16
EXTRACELLULAR_REGION_PART	321	0.30	1.74	0.00	0.17
CENTRAL_NERVOUS_SYSTEM_DEVELOPMENT	112	0.21	1.61	0.00	0.17
CELL_JUNCTION	76	0.30	1.61	0.00	0.17
CHEMICAL_HOMEOSTASIS	146	0.19	1.67	0.00	0.17
NERVOUS_SYSTEM_DEVELOPMENT	360	0.21	1.62	0.00	0.17
ANATOMICAL_STRUCTURE_MORPHOGENESIS	359	0.19	1.79	0.00	0.17
WOUND_HEALING	53	0.31	1.77	0.00	0.17
POSITIVE_REGULATION_OF_CELLULAR_PROTEIN_METABOLIC_PROCESS	69	0.18	1.62	0.00	0.17
REGULATION_OF_CYTOKINE_BIOSYNTHETIC_PROCESS	37	0.35	1.67	0.00	0.17
RESPONSE_TO_BACTERIUM	27	0.38	1.62	0.00	0.17
TRANSMEMBRANE_RECEPTOR_ACTIVITY	402	0.30	1.67	0.00	0.17
HSA04742_TASTE_TRANSDUCTION	45	0.27	1.66	0.00	0.18
NEURITE_DEVELOPMENT	53	0.30	1.62	0.00	0.18
HOMEOSTATIC_PROCESS	196	0.14	1.67	0.00	0.18
EXTRACELLULAR_SPACE	234	0.29	1.62	0.00	0.18
POSITIVE_REGULATION_OF_CELLULAR_METABOLIC_PROCESS	206	0.11	1.70	0.00	0.18
BASOLATERAL_PLASMA_MEMBRANE	31	0.32	1.60	0.00	0.18
HSA04940_TYPE_I_DIABETES_MELLITUS	41	0.36	1.79	0.00	0.18
CELLULAR_MORPHOGENESIS_DURING_DIFFERENTIATION	49	0.32	1.60	0.00	0.18
RHODOPSIN_LIKE_RECEPTOR_ACTIVITY	129	0.31	1.60	0.00	0.18
SECOND_MESSENGER_MEDIATED_SIGNALING	151	0.27	1.67	0.00	0.18
ENDOPEPTIDASE_ACTIVITY	110	0.19	1.63	0.00	0.18
TRANSMEMBRANE_RECEPTOR_PROTEIN_TYROSINE_KINASE_ACTIVITY	43	0.37	1.59	0.00	0.19
HSA04080_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	250	0.35	1.69	0.00	0.19
ADHERENS_JUNCTION	22	0.46	1.67	0.00	0.19
NEGATIVE_REGULATION_OF_CELL_DIFFERENTIATION	27	0.34	1.59	0.00	0.19
G_PROTEIN_SIGNALING_COUPLED_TO_CYCLIC_NUCLEOTIDE_SECOND_MESSENGER	100	0.32	1.65	0.00	0.19
ION_HOMEOSTASIS	121	0.22	1.80	0.00	0.19
CELL_MATRIX_JUNCTION	17	0.43	1.59	0.00	0.19
PHOSPHOINOSITIDE_MEDIATED_SIGNALING	46	0.30	1.59	0.00	0.19
NEURON_DIFFERENTIATION	73	0.30	1.59	0.00	0.19
G_PROTEIN_COUPLED_RECEPTOR_ACTIVITY	183	0.30	1.68	0.00	0.19
CYTOKINE_METABOLIC_PROCESS	40	0.32	1.64	0.00	0.19
G_PROTEIN_SIGNALING_COUPLED_TO_IP3_SECOND_MESSENGER_PHOSPHOLIPASE	43	0.29	1.58	0.00	0.19
CYTOKINE_BIOSYNTHETIC_PROCESS	39	0.34	1.68	0.00	0.19
HSA01430_CELL_COMMUNICATION	123	0.32	1.67	0.00	0.19
AXONOGENESIS	43	0.38	1.64	0.00	0.19
CYCLIC_NUCLEOTIDE_MEDIATED_SIGNALING	102	0.31	1.64	0.00	0.19
POSITIVE_REGULATION_OF_PROTEIN_METABOLIC_PROCESS	71	0.17	1.58	0.00	0.19
HSA04610_COMPLEMENT_AND_COAGULATION_CASCADES	66	0.34	1.58	0.00	0.19
CELL_MIGRATION	93	0.26	1.81	0.00	0.19

FDR-q value: the estimated probability that the normalized enrichment score represents a false positive finding; NES: Normalized enrichment score; the enrichment score for the gene set after it has been normalized across analyzed gene sets; Size: Number of genes in the gene set after filtering out those genes not in the expression dataset; The more interesting gene sets achieve the maximum enrichment score near the top or bottom of the ranked list;

Supplemental Table 2. Enrichment of gene sets that are negatively correlated to CD44⁺ phenotype. Related to Figure 4.

NAME	SIZE	ES	NES	NOM p-value	FDR q-value
CELL_CYCLE_ARREST_GO_0007050	55	-0.22	-1.66	0.00	0.23
NUCLEOBASE_NUCLEOSIDE_AND_NUCLEOTIDE_METABOLIC_PROCESS	49	-0.35	-1.49	0.00	0.23
CELLULAR_CATABOLIC_PROCESS	198	-0.27	-1.46	0.00	0.23
REGULATION_OF_CYCLIN_DEPENDENT_PROTEIN_KINASE_ACTIVITY	42	-0.32	-1.66	0.00	0.23
CATABOLIC_PROCESS	210	-0.26	-1.46	0.00	0.23
MICROTUBULE_CYTOSKELETON_ORGANIZATION_AND_BIOGENESIS	34	-0.46	-1.49	0.00	0.23
CELLULAR_CARBOHYDRATE_METABOLIC_PROCESS	123	-0.12	-1.44	0.00	0.23
HSA00240_PYRIMIDINE_METABOLISM	85	-0.43	-1.51	0.00	0.23
NUCLEOTIDYLTRANSFERASE_ACTIVITY	45	-0.46	-1.48	0.00	0.23
MICROTUBULE_BINDING	29	-0.37	-1.45	0.00	0.23
CELLULAR_PROTEIN_COMPLEX_ASSEMBLY	31	-0.34	-1.51	0.00	0.23
BIOSYNTHETIC_PROCESS	433	-0.21	-1.45	0.00	0.23
UNFOLDED_PROTEIN_BINDING	41	-0.39	-1.49	0.00	0.23
ESTABLISHMENT_OF_ORGANELLE_LOCALIZATION	17	-0.51	-1.46	0.00	0.23
NUCLEAR_CHROMOSOME_PART	32	-0.35	-1.52	0.00	0.23
RNA_BINDING	236	-0.46	-1.48	0.00	0.23
RNA_SPLICING_VIA_TRANSESTERIFICATION_REACTIONS	32	-0.57	-1.52	0.00	0.23
ATPASE_ACTIVITY_COUPLED	87	-0.30	-1.66	0.00	0.23
SECRETION_BY_CELL	103	-0.16	-1.53	0.00	0.23
HSA00450_SELENOAMINO_ACID_METABOLISM	28	-0.37	-1.53	0.00	0.23
SPINDLE_MICROTUBULE	16	-0.70	-1.53	0.00	0.23
RNA_SPLICING_FACTOR_ACTIVITY_TRANSESTERIFICATION_MECHANISM	18	-0.72	-1.53	0.00	0.22
RESPONSE_TO_DNA_DAMAGE_STIMULUS	151	-0.39	-1.66	0.00	0.22
UBIQUITIN_CYCLE	46	-0.32	-1.54	0.00	0.22
DNA_DIRECTED_RNA_POLYMERASE_II_HOLOENZYME	55	-0.53	-1.54	0.00	0.22
PHOSPHOINOSITIDE_METABOLIC_PROCESS	27	-0.41	-1.66	0.00	0.22
GLYCEROPHOSPHOLIPID_BIOSYNTHETIC_PROCESS	26	-0.44	-1.65	0.00	0.22
RESPONSE_TO_ENDOGENOUS_STIMULUS	187	-0.34	-1.65	0.00	0.22
NUCLEOSIDE_TRIPHOSPHATASE_ACTIVITY	195	-0.22	-1.55	0.00	0.21
SMALL_NUCLEAR_RIBONUCLEOPROTEIN_COMPLEX	21	-0.73	-1.55	0.00	0.21
PHOSPHOINOSITIDE_BIOSYNTHETIC_PROCESS	20	-0.47	-1.63	0.00	0.21
CELL_CYCLE_PHASE	158	-0.34	-1.57	0.00	0.21
HSA04110_CELL_CYCLE	110	-0.45	-1.56	0.00	0.21
PURINE_NUCLEOTIDE_BINDING	199	-0.15	-1.57	0.00	0.21
MRNA_PROCESSING_GO_0006397	68	-0.57	-1.56	0.00	0.21
HSA00564_GLYCEROPHOSPHOLIPID_METABOLISM	67	-0.18	-1.63	0.00	0.21
LIPID_BIOSYNTHETIC_PROCESS	89	-0.23	-1.55	0.00	0.21
HYDROLASE_ACTIVITY_ACTING_ON_ACID_ANHYDRIDES	211	-0.22	-1.61	0.00	0.21
NUCLEAR_LUMEN	342	-0.42	-1.64	0.00	0.21
RNA_PROCESSING	161	-0.58	-1.63	0.00	0.21
NUCLEOTIDE_BIOSYNTHETIC_PROCESS	19	-0.38	-1.64	0.00	0.20
DNA_DEPENDENT_DNA_REPLICATION	53	-0.49	-1.65	0.00	0.20
DNA_INTEGRITY_CHECKPOINT	21	-0.48	-1.58	0.00	0.20
SPLICEOSOME	46	-0.65	-1.61	0.00	0.20
COENZYME_METABOLIC_PROCESS	35	-0.53	-1.63	0.00	0.20
HSA00440_AMINOPHOSPHONATE_METABOLISM	15	-0.62	-1.63	0.00	0.20
DAMAGED_DNA_BINDING	20	-0.59	-1.64	0.00	0.20
MRNA_METABOLIC_PROCESS	79	-0.54	-1.57	0.00	0.20
HSA00563_GLYCOSYLPHOSPHATIDYLINOSITOL_ANCHOR_BIOSYNTHESIS	22	-0.44	-1.65	0.00	0.20
CHROMOSOMAL_PART	89	-0.46	-1.62	0.00	0.20
ENDOPLASMIC_RETICULUM	271	-0.22	-1.64	0.00	0.20
ORGANELLE_LUMEN	410	-0.42	-1.61	0.00	0.20
DNA_DEPENDENT_ATPASE_ACTIVITY	19	-0.61	-1.62	0.00	0.20
CELL_CYCLE_PROCESS	181	-0.36	-1.62	0.00	0.20
MEMBRANE_ENCLOSED_LUMEN	410	-0.42	-1.61	0.00	0.20
CHROMATIN	34	-0.37	-1.60	0.00	0.19
PYROPHOSPHATASE_ACTIVITY	209	-0.21	-1.59	0.00	0.19
GENERAL_RNA_POLYMERASE_II_TRANSCRIPTION_FACTOR_ACTIVITY	24	-0.51	-1.59	0.00	0.19
DNA_REPLICATION_INITIATION	16	-0.60	-1.59	0.00	0.19
NUCLEOTIDE_BINDING	212	-0.16	-1.60	0.00	0.19
HSA00100_BIOSYNTHESIS_OF_STEROIDS	24	-0.43	-1.60	0.00	0.19
CHROMOSOME	115	-0.45	-1.60	0.00	0.19
DOUBLE_STRAND_BREAK_REPAIR	23	-0.49	-1.59	0.00	0.19
COFACTOR_METABOLIC_PROCESS	49	-0.51	-1.60	0.00	0.19
PURINE_RIBONUCLEOTIDE_BINDING	194	-0.15	-1.60	0.00	0.19

Supplemental Table 3. A full list of the ECM and growth factor genes in the leading edge subsets enriched in the CD44⁺ cells. FDR-q value: the estimated probability that the normalized enrichment score represents a false positive finding; NES: Normalized enrichment score; the enrichment score for the gene set after it has been normalized across analyzed gene sets; Size: Number of genes in the gene set after filtering out those genes not in the expression dataset; The more interesting gene sets achieve the maximum enrichment score. The ranking metric measures a gene's correlation with a phenotype. A positive value indicates correlation with the CD44 negative phenotype. Related to Fig.4B.

Supplemental materials:

Antibodies used for flow cytometry.

Antigen	Conjugates	Clone	Isotype	Sources
Rat TER-119	Purified	TER-119		Biologend
Rat CD45	Purified	30-F11		Biologend
Rat B220	Purified	RA3-6B2		Biologend
Rat CD4	Purified	GK1.5		Biologend
Rat CD8	Purified	53-6.7		Biologend
Rat CD11B	Purified	M1/70		Biologend
Rat GR1	Purified	RB6-8C5		Biologend
Rat CD19	APC	1D3		Imgenex
Rat B220	PECY7	RA3-6B2		Biologend
Rat CD45	PECy5	30-F11		eBioscience
Rat CD31	PECY7	390		eBioscience
Rat CD44	APCCY7	IM7	Rat IgG2b, k	eBioscience
Rat TER-119	PECY5	TER-119		eBioscience
Rat CD51	PE	RMV-7	Rat IgG1,k	Biologend
Rat SCA1	PE	D7	Rat IgG2a, k	BD
Rat SCA1	Pacific blue	D7	Rat IgG2a, k	eBioscience
Rat KI67	PE	B56	Mouse IgG1	BD
Rat CD32/16	Purified	93		eBioscience
Goat anti-rat IgG (H+L) polyclonal	PECY5	CLCC40106		Cedar Lane laboratory

Antibodies used for flow cytometry (continued).

Antigen	Conjugates	Clone	Isotype	Sources
Rat CD140/PDGFRa	APC	APA5	Rat IgG2a, k	eBioscience
Rat CD19	PECY5	1D3		Biolegend
Rat B220	PECY5	RA3-6B2		Biolegend
Mouse anti-rat/mouse CD90.1	Alexa Fluor® 647	OX-7	Mouse IgG1, κ	Ebioscience
Mouse CD45	eFluor® 450	2D1	Mouse IgG1, κ	Ebioscience
Mouse CD235	eFluor® 450	6A7M	Mouse IgG1, κ	Ebioscience
Mouse CD31	APC	WM59	Mouse IgG1, κ	Ebioscience
Mouse CD146	PE	SHM-57	Mouse IgG2a, κ	Biolegend
Mouse CD271	FITC	ME20.4	Mouse IgG1, κ	Biolegend
Mouse CD19	eFluor® 450	HIB19	Mouse IgG1, κ	Ebioscience
Rat CD105	PE	MJ7/18	Rat IgG2a, k	eBioscience
Rat CD106	APC	429	Rat IgG2a, k	Biolegend
Mouse CD31	PE-CY7	WM-59	Mouse IgG1, k	eBioscience
Mouse STRO-1	Alexa Fluor® 647	STRO-1	Mouse IgM, λ	Biolegend
Mouse CD105	Alexa Fluor® 647	43A3	Rat IgG1, k	Biolegend
Mouse CD106	PECY5	STA	Rat IgG1, k	Biolegend
Mouse CD29	APC	TS2/16	Rat IgG1, k	Biolegend
Mouse CD73	APC	AD2	Rat IgG1, k	Biolegend

List of the probes used for Q- RT-PCR.

Gene name	Assay ID	Full name
Fmod	Mm00491215_m1	Fibromodulin
Angptl1	Mm00472259_m1	Angiopoietin-Like 1
Colla1	Mm00801666_g1	Collagen, Type I, alpha 1
Igf1	Mm00439560_m1	Insulin-Like Growth Factor 1
p27	Mm00438168_m1	p27/cdkn1b
p21 CIP1	Mm00432448_m1	p21/cdkn1a
Nov	Mm00456855_m1	Nephroblastoma overexpressed gene
Nes	Mm00450205_m1	nestin
Cdk6	Mm00438163_m1	Cyclin-dependent kinase 6
Hprt	Mm00446968_m1	hypoxanthine guanine phosphoribosyl transferase