Available online at www.sciencedirect.com

SciVerse ScienceDirect

Taiwanese Journal of Obstetrics & Gynecology 50 (2011) 447–457

www.tjog-online.com

Original Article

Comparative transcriptome analysis reveals a fetal origin for mesenchymal stem cells and novel fetal surface antigens for noninvasive prenatal diagnosis

Shun-Long Weng [a,e,1], Shing-Jyh Chang [a,b,1], Yi-Chieh Cheng [c], Hua-Yong Wang [c], Tao-Yeuan Wang [d,e], Margaret Dah-Tsyr Chang [b], Hsei-Wei Wang [c,f,g,*]

[a] *Department of Obstetrics and Gynecology, Hsinchu Mackay Memorial Hospital, Hsinchu 300, Taiwan*
[b] *Institute of Molecular and Cellular Biology, National Tsing Hua University, Hsinchu 300, Taiwan*
[c] *Institute of Microbiology and Immunology, National Yang-Ming University, Taipei 112, Taiwan*
[d] *Department of Pathology, Mackay Memorial Hospital, Taipei 104, Taiwan*
[e] *Mackay Medicine, Nursing and Management College, Taipei 112, Taiwan*
[f] *VGH-YM Genome Research Center, National Yang-Ming University, Taipei 112, Taiwan*
[g] *Department of Education and Research, Taipei City Hospital, Taipei 103, Taiwan*

Accepted 10 March 2011

Abstract

Objective: Mesenchymal stem cells (MSCs) are an attractive source for providing the cells necessary for regenerating damaged tissues. Fetal MSCs (fMSCs) are known to proliferate fast and have an excellent osteogenic capacity, yet the underlying mechanisms need to be explored. A better understanding of MSCs from different anatomic origins and ages will eventually benefit cell-based therapies, as well as subsequent mechanistic studies in the field of stem cell biology.
Materials and Methods: We identified the molecular signatures of fetal and adult MSCs via a meta-analytic strategy and compared the enriched canonical pathways and genetic networks within each signature.
Results: Fetal MSCs were found to express more cell cycle genes, which is consistent with the results of wetlab functional assays. In addition, the genes involved in vasculogenesis, neurogenesis, Wnt, MAPKKK pathways, and RNA splicing were found to be enriched in fMSCs. Correlating with the overexpression of multilineage differentiation genes, fMSCs share more genes with embryonic stem cells (ESCs) and are, therefore, more primitive. Further exploration into the transcriptome similarities revealed that MSCs from umbilical cord blood (UCB) express dominant fMSC genes, but not adult genes, suggesting a fetal origin for UCB MSCs. Novel surface proteins that were dominantly expressed in fetal and UCB MSCs, but not in adult MSCs or maternal PBMCs, were also identified.
Conclusion: Our results systematically revealed the underlying genes and regulatory networks of two MSCs from unique origins, the resulting phenotypes, as well as the origin of UCB MSCs. The novel membrane proteins on the fetal MSC surface are promising candidate biomarkers for positively isolating fetal MSCs from maternal blood for noninvasive prenatal diagnosis.
Copyright © 2011, Taiwan Association of Obstetrics & Gynecology. Published by Elsevier Taiwan LLC. All rights reserved.

Keywords: Genetic networks; Mesenchymal stem cells; Noninvasive prenatal diagnosis; Surface antigens; Transcriptome analysis

Introduction

Stem cells have drawn immense research interest because they hold great promise for future therapeutic applications. Mesenchymal stem cells (MSC) are an attractive source of somatic stem cells for the regeneration of damaged tissues. Bone marrow (BM) is the primary source of multipotent

* Corresponding author. Institute of Microbiology and Immunology, National Yang-Ming University, 155, Section 2, Li-Nong Street, Taipei, Taiwan.
E-mail address: hwwang@ym.edu.tw (H.-W. Wang).
[1] Both authors contributed equally to this work.

1028-4559/$ - see front matter Copyright © 2011, Taiwan Association of Obstetrics & Gynecology. Published by Elsevier Taiwan LLC. All rights reserved.
doi:10.1016/j.tjog.2011.10.009

MSCs, but the procurement of BM is a highly invasive procedure. The number, differentiation potential, frequency, and proliferative potential of MSCs decline with increasing age [1]. Alternative sources from which MSCs could be isolated are, therefore, the subject of intensive investigations. MSCs have also been isolated from a variety of tissues, including BM, umbilical cord blood (UCB) [2–4], adipose tissue, and muscle [5–8].

Although the MSCs of different origins share common characteristics, discrepancies in terms of gene expression patterns, differentiation potentials, and proliferation rates exist among them. A comparative analysis of MSCs derived from BM, UCB, and adipose tissue has been conducted [8,9]. Gene expression profiles and comparisons of various human MSCs have also been reported in quite a few studies [7,10–14]. Recently, using systems biology tools, we have also shown that perinatal MSCs isolated from Wharton's jelly matrix of the umbilical cord express more pluripotency markers and, thereby, are able to grow faster than adult MSCs [15]. Regarding MSCs obtained from fetuses, these stem cells possess a superior osteogenic capacity, proliferate faster, and have longer telomeres compared with perinatal and adult MSCs [16–18]. However, the underlying mechanisms have not yet been studied. Further study of the differences in the gene expression profiles of human fetal and adult MSCs should provide more clues.

In this study, we closely examined the transcriptomes of fetal and adult MSCs and applied systems biology tools to provide insights into their differential biological activities and responsible genes. We further compared the relationships between UCB, fetal, and adult MSCs using a molecular fingerprinting approach. Novel surface biomarkers that are dominantly expressed in fetal and UCB MSCs, but not adult MSCs or maternal PBMCs, were also identified.

Materials and methods

Microarray data sets

All array data were gathered using the Affymetrix™ HG-U133 Plus 2.0 GeneChip system (Affymetrix, Santa Clara, CA, USA). Array data for embryonic stem cells (ESC) and adult MSCs were obtained from our previous study [15]. Adult MSC array data were also downloaded from the Gene Expression Omnibus database (GEO; http://www.ncbi.nlm.nih.gov/geo/; datasets GSE9451, GSE10315, GSE12274, and GSE12843). UCB MSC array data were downloaded from GEO dataset GSE13491. Fetal MSC array data were obtained from datasets GSE13975 and GSE18934. Maternal peripheral blood mononuclear cell (PBMC) array data were also obtained from GSE18934.

Array data processing

Feature selection was performed as previously described [19]. Briefly, the default robust multichip average (RMA) settings were used to correct for background noise and normalize and summarize all expression values using the 'affy' package of the Bioconductor suite of software (www.bioconductor.org/) for the R statistical programming language. t-Statistics were calculated as normal for each gene, and a p-value was then calculated using a modified permutation test in the "LIMMA" package [19]. To control for multiple testing errors, a false discovery rate (FDR) algorithm was then applied to the p-values to calculate a set of q-values; i.e., thresholds for the expected proportion of false positives or false rejections of the null hypothesis [19,20]. Gene annotations were performed using the ArrayFusion web tool (http://microarray.ym.edu.tw/tools/arrayfusion/) [21].

Heat maps were created using dChip software (http://www.dchip.org/). Principle component analysis (PCA) was performed using Partek Genomics Suite (Partek, St. Louis, Missouri, USA) in order to provide a visual impression of how the various sample groups are related. Gene enrichment analysis was performed using the Gene Ontology (GO) and KEGG databases and DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/) [22]. The Euclidean distance between the two groups of samples was calculated by average linkage measurements (the mean of all pair-wise distances [i.e., linkages] between members of the two concerned groups) [23]. The standard error of the average linkage distance between two groups (the standard deviation of pair-wise linkages divided by the square root of the number of linkages) is quoted when intergroup distances are compared in the text.

Genetic network analysis

Differential gene expression profiles were imported into the Ingenuity Pathways Analysis (IPA) software (Ingenuity Systems, Redwood City, CA, USA, http://www.ingenuity.com/) in order to determine any functional regulatory networks and canonical pathways. IPA uses a curated database to construct functional regulatory networks from a list of individual genes and determine a statistical score for each network according to the best fit of the network to the set of focused genes. The score is the negative log of p and denotes the likelihood of the focused genes in the network being found together due to chance. The biological functions assigned to each network are ranked according to the significance of that biological function to the network. The Fisher's exact test was used to calculate p-values in order to determine the probability that the biological function assigned to a network is explained by chance alone. The significance of the p-value calculations were based on the hypergeometric distribution calculated using the right-tailed Fisher's exact test for 2x2 contingency tables. "Right-tailed" refers to the fact that IPA only shows overrepresented functional pathway annotations.

Results

Increased embryonic stem cell traits in fetal MSCs

It is well known that fetal MSCs possess a superior osteogenic capacity, grow faster, and have longer telomeres

compared with perinatal and adult MSCs, and it has been suggested that fetal MSCs are more primitive than adult MSCs [16–18]. To analyze whether this hypothesis could be verified at the genomic level, we analyzed the gene expression profiles of fetal and adult MSCs and then compared both transcriptomes to that of ESCs. To eliminate the effects of individual differences and batch effects, our own array data, as well data obtained from four additional adult MSC array datasets from the public data repository, were also used. Also, two fetal MSC array datasets that had been implemented by two different groups, GSE13975 and GSE18934 [24,25], were included in our panel.

Genes that were differentially expressed in fetal and adult MSCs were first filtered to determine the molecular signature of each type of MSC according to the statistical pipeline used [19,23]. There were 1909 probe sets that were abundantly overexpressed in fetal MSCs and another 276 probe sets that were overexpressed in adult MSCs (with a positive false discovery rate (pFDR) threshold of $q < 0.0001$. A principle component analysis (PCA) plot using these genes illustrates their grouping power (Fig. 1A). The top 50 known genes most strongly expressed in each MSC are listed in Tables 1 and 2. In the fetal MSCs, known stem cells and proliferating genes, such as S100A4, CDK6, and SOX11, are present (Table 1, labeled by the asterisks). In the adult MSCs, immune response-associated genes, including HLA-B, HLA-C, and MR1, are present (Table 2). These probe sets were used to compare the relationships between different MSCs and ESCs. As shown in Fig. 1B, fetal MSCs and ESCs share a close relationship, and this impression was verified by average linkage distance analysis (Fig. 1C). Unsupervised hierarchical clustering was also able to group the fetal MSCs and ESCs into the same leaf (Fig. 1D).

It has been suggested that fetal MSCs are more primitive than adult MSCs [16]. One of the possible mechanisms is that fetal stem cells express more ESC genes and are, therefore,

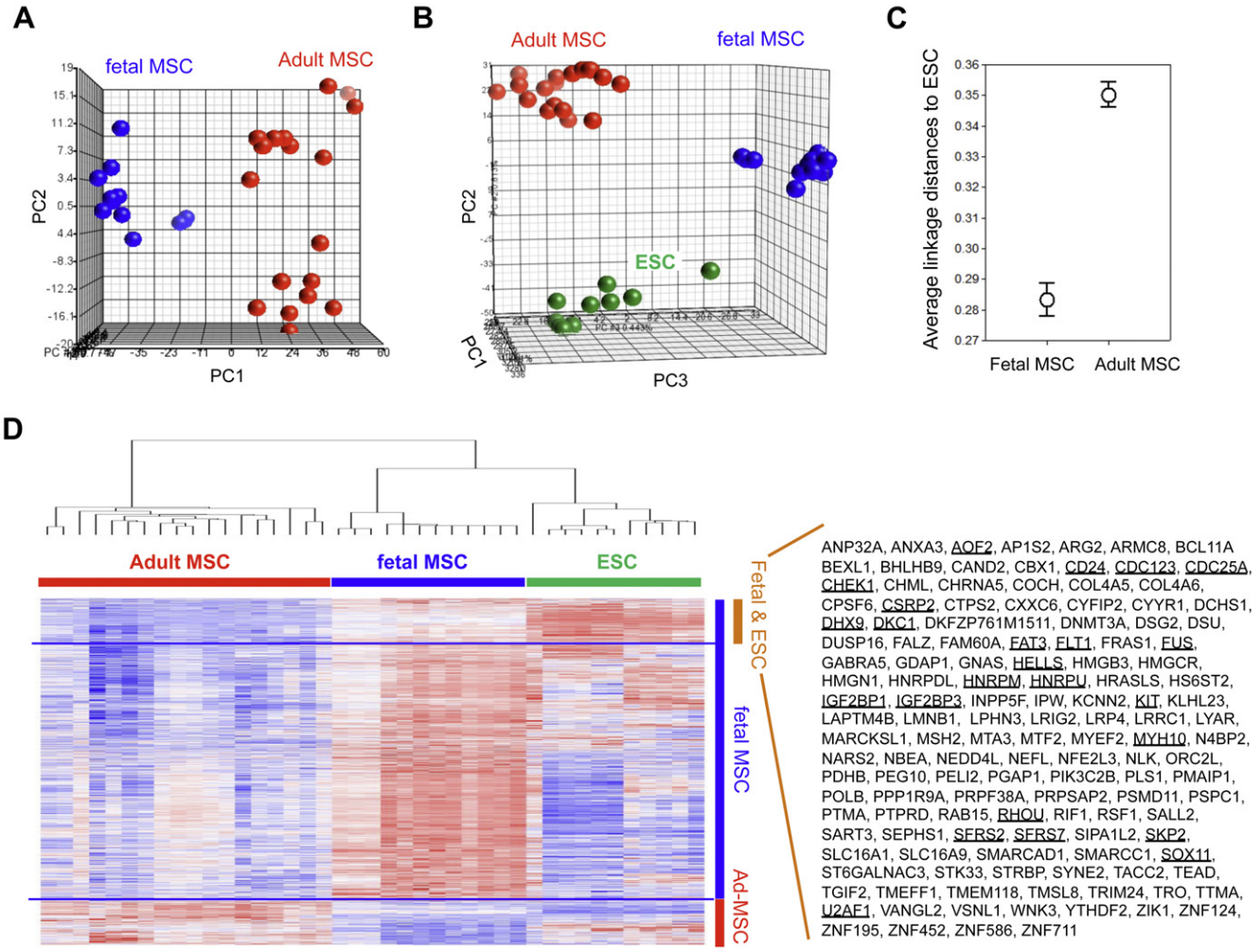

Fig. 1. Distinct gene expression patterns and ESC traits in fetal MSCs. (**A**) A principal component analysis (PCA) plot of genes that are differentially expressed in adult and fetal MSCs. Each spot represents a single array. (**B**) Comparison of the transcriptome traits between ESCs and different MSCs using the same gene set as above. (**C**) Relationships between ESCs and two different types of MSCs. Average linkage distances between MSCs and ESCs were calculated using genes that distinguish these two types of MSC. The confidence limits shown represent the standard error. (**D**) A heat map showing the genes that are expressed in ESCs and different MSC groups.

Table 1
Top 50 known genes in fetal MSCs.

Probe Set ID	UniGene ID	Gene Title	Gene Symbol	Location
202274_at	Hs.516105	actin, gamma 2, smooth muscle, enteric	ACTG2	chr2p13.1
227198_at	Hs.444414	AF4/FMR2 family, member 3	AFF3	chr2q11.2-q12
208016_s_at	Hs.709185	angiotensin II receptor, type 1	AGTR1	chr3q21-q25
209369_at	Hs.480042	annexin A3	ANXA3	chr4q13-q22
219087_at	Hs.435655	asporin	ASPN	chr9q22
222891_s_at	Hs.370549	B-cell CLL/lymphoma 11A (zinc finger protein)	BCL11A	chr2p16.1
209030_s_at	Hs.370510	cell adhesion molecule 1	CADM1	chr11q23.2
208650_s_at	Hs.694721	CD24 molecule	CD24	chr6q21
235287_at	Hs.119882	cyclin-dependent kinase 6	*CDK6	chr7q21-q22
229831_at	Hs.12723	contactin 3 (plasmacytoma associated)	CNTN3	chr3p26
213110_s_at	Hs.369089	collagen, type IV, alpha 5 (Alport syndrome)	COL4A5	chrXq22
201116_s_at	Hs.712551	carboxypeptidase E	*CPE	chr4q32.3
222101_s_at	Hs.199850	dachsous 1 (Drosophila)	DCHS1	chr11p15.4
1553105_s_at	Hs.412597	desmoglein 2	DSG2	chr18q12.1
206710_s_at	Hs.213394	erythrocyte membrane protein band 4.1-like 3	EPB41L3	chr18p11.32
215664_s_at	Hs.654492	EPH receptor A5	EPHA5	chr4q13.1
206795_at	Hs.42502	coagulation factor II (thrombin) receptor-like 2	F2RL2	chr5q13
203184_at	Hs.519294	fibrillin 2 (congenital contractural arachnodactyly)	FBN2	chr5q23-q31
1565717_s_at	Hs.513522	fusion (involved in t(12;16) in malignant liposarcoma)	FUS	chr16p11.2
208025_s_at	Hs.505924	high mobility group AT-hook 2	HMGA2	chr12q15
203819_s_at	Hs.700696	insulin-like growth factor 2 mRNA binding protein 3	IGF2BP3	chr7p11
1554757_a_at	Hs.709221	inositol polyphosphate-5-phosphatase, 40kDa	INPP5A	chr10q26.3
213496_at	Hs.13245	plasticity related gene 1	LPPR4	chr1p21.2
206163_at	Hs.584776	mab-21-like 1 (C. elegans)	MAB21L1	chr13q13
210302_s_at	Hs.584852	mab-21-like 2 (C. elegans)	MAB21L2	chr4q31
205347_s_at	Hs.56145	thymosin-like 8	TMSL8	chrXq21.33-q22.3
220615_s_at	Hs.298851	male sterility domain containing 1	MLSTD1	chr12p11.22
213067_at	Hs.16355	myosin, heavy chain 10, non-muscle	MYH10	chr17p13
237206_at	Hs.567641	myocardin	MYOCD	chr17p11.2
227971_at	Hs.209527	Nik related kinase	NRK	chrXq22.3
219978_s_at	Hs.615092	nucleolar and spindle associated protein 1	NUSAP1	chr15q15.1
228635_at	Hs.192859	protocadherin 10	PCDH10	chr4q28.3
205534_at	Hs.479439	protocadherin 7	PCDH7	chr4p15
205501_at	Hs.584856	phosphodiesterase 10A	PDE10A	chr6q26
212092_at	Hs.147492	paternally expressed 10	PEG10	chr7q21
204285_s_at	Hs.96	phorbol-12-myristate-13-acetate-induced protein 1	PMAIP1	chr18q21.32
235874_at	Hs.98381	protease, serine, 35	PRSS35	chr6q14.2
1554997_a_at	Hs.196384	prostaglandin-endoperoxide synthase 2	PTGS2	chr1q25.2-q25.3
211737_x_at	Hs.371249	pleiotrophin (heparin binding growth factor 8, neurite growth-promoting factor 1)	PTN	chr7q33-q34
1555233_at	Hs.656339	ras homolog gene family, member J	RHOJ	chr14q23.2
203186_s_at	Hs.654444	S100 calcium binding protein A4	*S100A4	chr1q21
210432_s_at	Hs.435274	sodium channel, voltage-gated, type III, alpha subunit	SCN3A	chr2q24
203626_s_at	Hs.23348	S-phase kinase-associated protein 2 (p45)	SKP2	chr5p13
207053_at	Hs.31961	solute carrier family 8 (sodium/calcium exchanger), member 1	SLC8A1	chr2p23-p22
204913_s_at	Hs.432638	SRY (sex determining region Y)-box 11	*SOX11	chr2p25
227705_at	Hs.21861	transcription elongation factor A (SII)-like 7	TCEAL7	chrXq22.1
223557_s_at	Hs.144513	transmembrane protein with EGF-like and two follistatin-like domains 2	TMEFF2	chr2q32.3
210654_at	Hs.213467	tumor necrosis factor receptor superfamily, member 10d, decoy	TNFRSF10D	chr8p21
219937_at	Hs.199814	thyrotropin-releasing hormone degrading enzyme	TRHDE	chr12q15-q21
219778_at	Hs.431009	zinc finger protein, multitype 2	ZFPM2	chr8q23

more multipotent. We examined this hypothesis by investigating the genes shared by ESCs and fetal MSCs, as well by checking the possible outcomes of the preferred ESC gene patterns in fetal MSCs. A total of 191 probe sets showed increased levels of both cell types when compared with adult MSCs (Fig. 1D). To gain more insights into the functional consequences of differential gene expression and to provide quantitative evidence, signature genes were searched for in the Gene Ontology (GO) database [26] in order to find statistically overrepresented functional groups within the gene lists. DAVID 6.7, a web-based tool, and the graph theory evidence-based method for agglomerating gene/protein identifiers [22], were applied to this task. A p-value was assigned to each category based on the EASE co-occurrence scoring system, a Fisher's exact-based method that identifies functional categories that are overrepresented in a gene list [22]. Given that the whole human transcriptome was represented in the microarray that was applied, this analysis was not biased

Table 2
Top 50 known genes in adult MSCs.

Probe Set ID	UniGene ID	Gene Title	Gene Symbol	Location
214454_at	Hs.591725	ADAM metallopeptidase with thrombospondin type 1 motif, 2	ADAMTS2	chr5qter
202888_s_at	Hs.1239	alanyl (membrane) aminopeptidase (CD13, p150)	ANPEP	chr15q25-q26
230925_at	Hs.310421	amyloid beta (A4) precursor protein-binding, family B, member 1 interacting protein	APBB1IP	chr10p12.1
212884_x_at	Hs.654439	apolipoprotein E	APOE	chr19q13.2
213429_at	Hs.158745	bicaudal C homolog 1 (Drosophila)	BICC1	chr10q21.1
224414_s_at	Hs.200242	caspase recruitment domain family, member 6	CARD6	chr5p13.1
224794_s_at	Hs.495230	cerebral endothelial cell adhesion molecule	CERCAM	chr9q34.11
205713_s_at	Hs.1584	cartilage oligomeric matrix protein	COMP	chr19p13.1
201200_at	Hs.5710	cellular repressor of E1A-stimulated genes 1	CREG1	chr1q24
206315_at	Hs.114948	cytokine receptor-like factor 1	CRLF1	chr19p12
204971_at	Hs.518198	cystatin A (stefin A)	CSTA	chr3q21
226911_at	Hs.20103	EGF-like, fibronectin type III and laminin G domains	EGFLAM	chr5p13.2-p13.1
200878_at	Hs.468410	endothelial PAS domain protein 1	EPAS1	chr2p21-p16
205756_s_at	Hs.654450	coagulation factor VIII, procoagulant component (hemophilia A)	F8	chrXq28
205782_at	Hs.567268	fibroblast growth factor 7 (keratinocyte growth factor)	FGF7	chr15q15-q21.1
214702_at	Hs.203717	fibronectin 1	FN1	chr2q34
226930_at	Hs.520525	fibronectin type III domain containing 1	FNDC1	chr6q25
205119_s_at	Hs.753	formyl peptide receptor 1	FPR1	chr19q13.4
201724_s_at	Hs.514806	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 1	GALNT1	chr18q12.1
216526_x_at	Hs.549053	major histocompatibility complex, class I, B	*HLA-B	chr6p21.3
208812_x_at	Hs.654404	major histocompatibility complex, class I, C	*HLA-C	chr6p21.3
209581_at	Hs.502775	HRAS-like suppressor 3	HRASLS3	chr11q12.3-q13.1
213620_s_at	Hs.431460	intercellular adhesion molecule 2	ICAM2	chr17q23-q25
229638_at	Hs.499205	iroquois homeobox 3	IRX3	chr16q12.2
202803_s_at	Hs.375957	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	ITGB2	chr21q22.3
1557080_s_at	Hs.696554	integrin, beta-like 1 (with EGF-like repeat domains)	ITGBL1	chr13q33
212192_at	Hs.644125	potassium channel tetramerisation domain containing 12	KCTD12	chr13q22.3
231031_at	Hs.536967	keratinocyte growth factor-like protein 2	KGFLP2	chr9p12
203543_s_at	Hs.150557	Kruppel-like factor 9	KLF9	chr9q13
201721_s_at	Hs.371021	lysosomal associated multispanning membrane protein 5	LAPTM5	chr1p34
200923_at	Hs.514535	lectin, galactoside-binding, soluble, 3 binding protein	LGALS3BP	chr17q25
206483_at	Hs.591865	leucine rich repeat containing 6	LRRC6	chr8q24.22
206584_at	Hs.660766	lymphocyte antigen 96	LY96	chr8q21.11
207565_s_at	Hs.101840	major histocompatibility complex, class I-related	*MR1	chr1q25.3
37005_at	Hs.466662	neuroblastoma, suppression of tumorigenicity 1	NBL1	chr1p36.13-p36.11
200632_s_at	Hs.372914	N-myc downstream regulated gene 1	NDRG1	chr8q24.3
215870_s_at	Hs.319438	phospholipase A2, group V	PLA2G5	chr1p36-p34
224506_s_at	Hs.134292	phosphatidic acid phosphatase type 2 domain containing 3	PPAPDC3	chr9q34.13
211663_x_at	Hs.446429	prostaglandin D2 synthase 21kDa (brain)	PTGDS	chr9q34.2-q34.3
221840_at	Hs.127022	protein tyrosine phosphatase, receptor type, E	PTPRE	chr10q26
211600_at	Hs.160871	protein tyrosine phosphatase, receptor type, O	PTPRO	chr12p13.3-p13.2|12p13-p12
225188_at	Hs.471162	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1	RAPH1	chr2q33
200986_at	Hs.384598	serpin peptidase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	SERPING1	chr11q12-q13.1
222784_at	Hs.497349	SPARC related modular calcium binding 1	SMOC1	chr14q24.2
201150_s_at	Hs.701968	TIMP metallopeptidase inhibitor 3 (Sorsby fundus dystrophy, pseudoinflammatory)	TIMP3	chr22q12.1-q13.2|22q12.3
213293_s_at	Hs.501778	tripartite motif-containing 22	TRIM22	chr11p15
202546_at	Hs.714302	vesicle-associated membrane protein 8 (endobrevin)	VAMP8	chr2p12-p11.2
204619_s_at	Hs.695930	versican	VCAN	chr5q14.3
228617_at	Hs.441975	XIAP associated factor 1	XAF1	chr17p13.2
205383_s_at	Hs.655108	zinc finger and BTB domain containing 20	ZBTB20	chr3q13.2

toward the coverage of the microarray. The included genes are involved in RNA splicing (including DHX9, FUS, HNRNPM, HNRNPU, SFRS2, SFRS7, and U2AF1; $p = 0.0018$; genes are bolded in Fig. 1D), cell proliferation (AOF2, CDC25A, CSRP2, DKC1, HELLS, KIT, MYH10, SKP2, SOX11, and TACC2; $p = 0.011$), and the interphase of the mitotic cell cycle (CDC123, CDC25A, CHEK1, RHOU, and SKP2; $p = 0.011$). These results correlate with the concept that fetal MSCs have a more multipotent and grow faster than adult MSCs [16–18].

A

KEGG Term	Count	%	P Value	
Cell adhesion molecules (CAMs)	12	5.50	1.20E-06	
Antigen processing and presentation	6	2.74	0.005092	*(adult MSC)*
Arachidonic acid metabolism	5	2.28	0.006710	
Natural killer cell mediated cytotoxicity	7	3.20	0.008894	
GnRH signaling pathway	5	2.28	0.043271	
Spliceosome	31	2.08	8.08E-08	
Ubiquitin mediated proteolysis	28	1.88	1.74E-05	
Cell cycle	25	1.68	7.95E-05	
Steroid biosynthesis	8	0.54	2.52E-04	
Wnt signaling pathway	24	1.61	0.003211	
Terpenoid backbone biosynthesis	6	0.40	0.005910	*(fetal MSC)*
Adherens junction	14	0.94	0.010531	
Inositol phosphate metabolism	11	0.74	0.012640	
Neurotrophin signaling pathway	19	1.28	0.014029	
Phosphatidylinositol signaling system	13	0.87	0.018586	
mTOR signaling pathway	10	0.67	0.026568	

B

	Count	%	PValue
Biological process			
Immune system process	20	9.13	8.23E-03
Antigen processing and presentation	5	2.28	1.12E-02
Antigen processing and presentation of peptide antigen	5	2.28	1.95E-04
Antigen processing and presentation of peptide antigen via MHC class I	5	2.28	2.48E-05
HLA-A, HLA-B, HLA-E, B2M, MR1			
Biological adhesion	21	9.59	4.32E-05
Cell adhesion	21	9.59	4.23E-05
Cell-substrate adhesion	7	3.20	5.72E-04
Cell-matrix adhesion	7	3.20	3.41E-04
ITGBL1, ITGB2, ITGB5, ITGB7, PTK2B, FN1, RAPH1			
Developmental process			
Anatomical structure development	36	16.44	4.58E-02
System development	34	15.53	4.00E-02
Organ development	27	12.33	3.83E-02
Tissue development			
Epidermis development	6	2.74	4.44E-02
FGF7, CSTA, ADAMTS2, RELA, KGFLP2, SPRR2G			
Locomotion	10	4.57	3.78E-02
DNAH1, ITGB2, PTK2B, VCAN, FPR1, SPN, OPHN1, PVRL1, FN1, CCL28			

C

	Count	%	PValue
Biological process			
Metabolic process	723	48.55608	1.21E-08
Macromolecule metabolic process	593	39.82539	5.60E-14
cellular macromolecule metabolic process	577	38.75084	2.81E-20
DNA metabolic process	74	4.969778	1.61E-06
DNA replication	27	1.813298	0.007265
DNA repair	43	2.887844	1.50E-04
RNA metabolic process	152	10.20819	2.40E-16
mRNA metabolic process	74	4.969778	1.38E-12
mRNA processing	70	4.701142	7.96E-14
nuclear mRNA splicing, via spliceosome	41	2.753526	2.82E-11
Gene expression	348	23.37139	4.82E-13
Cellular process	1003	67.36064	3.30E-23
Mitotic cell cycle	65	4.365346	9.23E-09
Cell cycle	113	7.588986	2.28E-09
Cell cycle process	85	5.708529	6.90E-08
Cell cycle phase	69	4.633983	2.71E-08
M phase	54	3.626595	1.72E-06
M phase of mitotic cell cycle	42	2.820685	9.63E-07
Mitosis	41	2.753526	1.55E-06
Cellular metabolic process	678	45.53392	1.46E-15
Cellular catabolic process	107	7.186031	0.00803
Cellular macromolecule catabolic process	94	6.312962	1.07E-05
Cellular protein catabolic process	82	5.507052	7.45E-06
Proteolysis involved in cellular protein catabolic proces	81	5.439893	1.12E-05
Developmental process			
Embryonic development	61	4.096709	0.02911
In utero embryonic development	23	1.544661	0.032773
Neural tube formation	8	0.537273	0.043207
Vasculogenesis	9	0.604433	0.017294
Leukocyte differentiation	18	1.208865	0.041244
Signaling			
Signaling pathway			
Intracellular signaling pathway	122	8.193418	0.038689
Regulation of MAPKKK cascade	17	1.141706	0.016848
Wnt receptor signaling pathway	22	1.477502	0.002886
Ras protein signal transduction	15	1.007388	0.049708

Fig. 2. Altered functional modules of different MSCs. (**A**) KEGG pathways that are significantly enriched in each of the MSCs. The number of genes, their percentages in terms of the total number of genes, and the *p*-values for pathways that are significantly overrepresented ($p < 0.05$) are listed. (**B-C**) Gene set enrichment analysis according to the Gene Ontology (GO) classification. Probe sets that are unique in adult or fetal MSCs were searched for in the GO database using the DAVID 6.7 interface. The number of genes, gene symbols, their percentages, and *p*-values for each category that shows significance ($p < 0.05$) and are enriched in either adult (**B**) or fetal (**C**) MSCs are listed.

Canonical pathway and functional module analysis as a framework for the interpretation of MSC biology

To provide a more comprehensive understanding of fetal and adult MSCs, we also analyzed the enhanced canonical pathways of each subtype. The genes of the Wnt and mTOR pathways were enriched in fetal MSCs, while genes associated with the GnRH signaling pathway were enriched in the adult stem cells (Fig. 2A). Cell cycle genes were overexpressed in the fetal MSCs, correlating with their faster proliferation capacities [16]. On the other hand, antigen processing and presentation were more active in the adult stem cells (Fig. 2A).

We further applied the GO database to find enriched biological activities in each type of cell. Adult MSCs significantly overexpressed the genes involved in antigen processing and the presentation of MHC class I peptide antigens ($p = 0.0000248$; Fig. 2B), which is consistent with the results of the KEGG analysis (Fig. 2A). Five BM-MSC genes (HLA-B, HLA-C, HLA-E, MR1, and B2M) are in this category (Fig. 2B), suggesting that fetal MSCs are less immunogenic and may induce less allograft rejection. On the other hand, adult MSCs also overexpressed cell adhesion genes ($p = 0.000341$; Fig. 2B), implying that they are less motile.

In contrast, the principal functions of the upregulated genes in fetal MSCs include those related to RNA splicing, cell proliferation, the mitotic cell cycle, and DNA replication and repair (Fig. 2C). The differential splicing genes expressed in different MSCs indicate that alternative RNA splicing may occur in MSCs, thereby influencing their biological activities. Genes involved in *in utero* embryonic development, leukocyte differentiation, vasculogenesis, and neurogenesis are more abundant in fetal MSCs (Fig. 2C), correlating with the above observations that these MSCs are more primitive, more similar to ESC, and, therefore, may hold multipotent differentiation abilities. Wnt receptor signaling, Ras protein signaling, and MAPKKK pathway genes are enriched in fetal MSCs and may, therefore, be responsible for their active proliferation and diverse differentiation capabilities (Fig. 2C). The genes involved in the cell cycle, Wnt pathway, and RNA splicing are illustrated in Supplementary Files 1–4, which are available online.

MSC genetic networks

The gene list outlined above gave us preliminary insights into the functional consequences of the detected differential gene expression. An increasing amount of evidence shows that genes do not act as individuals but are collaborative in genetic networks. To better understand how the genes that are enriched in different stem cells are related and, most importantly, how those interactions may be related to cellular functions, we performed a genetic network analysis to indentify the signature genes. Filtrated probe sets were input into the IPA software to construct network modules. For the fetal MSCs, a major network consisting of 547 genes was identified (not shown). Again, genes involved in cell cycle development, RNA posttranscriptional modification (including RNA splicing), DNA replication, recombination, and repair were found in this major network. Interestingly, we found that the targets of several microRNAs, including miR-124 (Fig. 3A), miR-1 (Fig. 3B), miR-199a1, miR-17, and miR-181b2 (Fig. 3C), were enriched in the fetal MSCs, suggesting the reduced expression of these miRNAs in fetal MSCs. Similar miRNA modules could not be found in the adult MSCs (not shown).

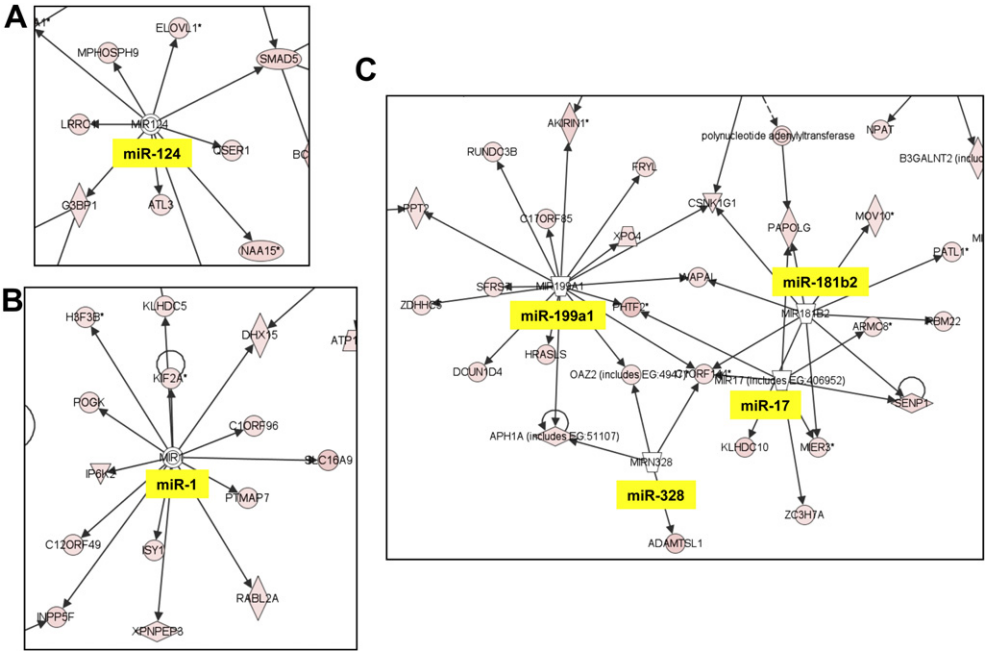

Fig. 3. Functional genetic network analysis reveals microRNA targets enriched with fetal MSCs. These networks are graphically displayed as nodes (gene products) and edges (biological relationships between nodes) that were mapped using the Ingenuity Pathway Analysis (IPA) tool. The intensity of the node color indicates the degree of upregulation. MicroRNAs are highlighted.

The molecular classification approach used to assign the possible source of UCB MSCs

Now that we have groups of genes that can be used to represent the two MSCs at the genomic level, the next issue we addressed was which MSCs are more similar to the UCB-derived MSCs, i.e., the possible origin of UCB MSCs. We answered this by including UCB MSCs in a PCA analysis and calculated the group distances between UCB MSCs, fetal MSCs, and adult MSCs. As shown in Fig. 4A, clearly the UCB MSCs have the closest relationship with the fetal MSCs. This impression was verified by average linkage distance analysis (Fig. 4B). A heat map based on the fetal and adult MSC signature genes showed common gene expression patterns between UCB MSCs and fetal MSCs (Fig. 4C). These two MSC subtypes also shared more common genes with the ESCs when compared with the adult MSCs (Fig. 4C).

Novel positive selection biomarkers used to isolate circulating fetal MSCs from maternal blood

It has been shown that fetal MSCs hold the potential for noninvasive prenatal diagnosis [25]. Isolating fetal MSCs from posttermination maternal blood using positive selection strategies relies largely on the identification of the novel surface antigens that are present on the circulating human fetal MSCs but not on maternal cells, including peripheral blood mononuclear cells (PBMC) and the maternal adult MSCs. We compared the transcriptome profiles of fetal MSCs and paired maternal PBMC and then searched for genes that are rarely expressed in adult MSCs. A total of 90 genes were found to fit the above criteria (Fig. 5A). Four cell membrane proteins were found to be unique in the fetal MSCs, including CPE, FZD2, ADAM23, and PVR (Fig. 5B). Among these, the expression of CPE (carboxypeptidase E) has recently been verified by

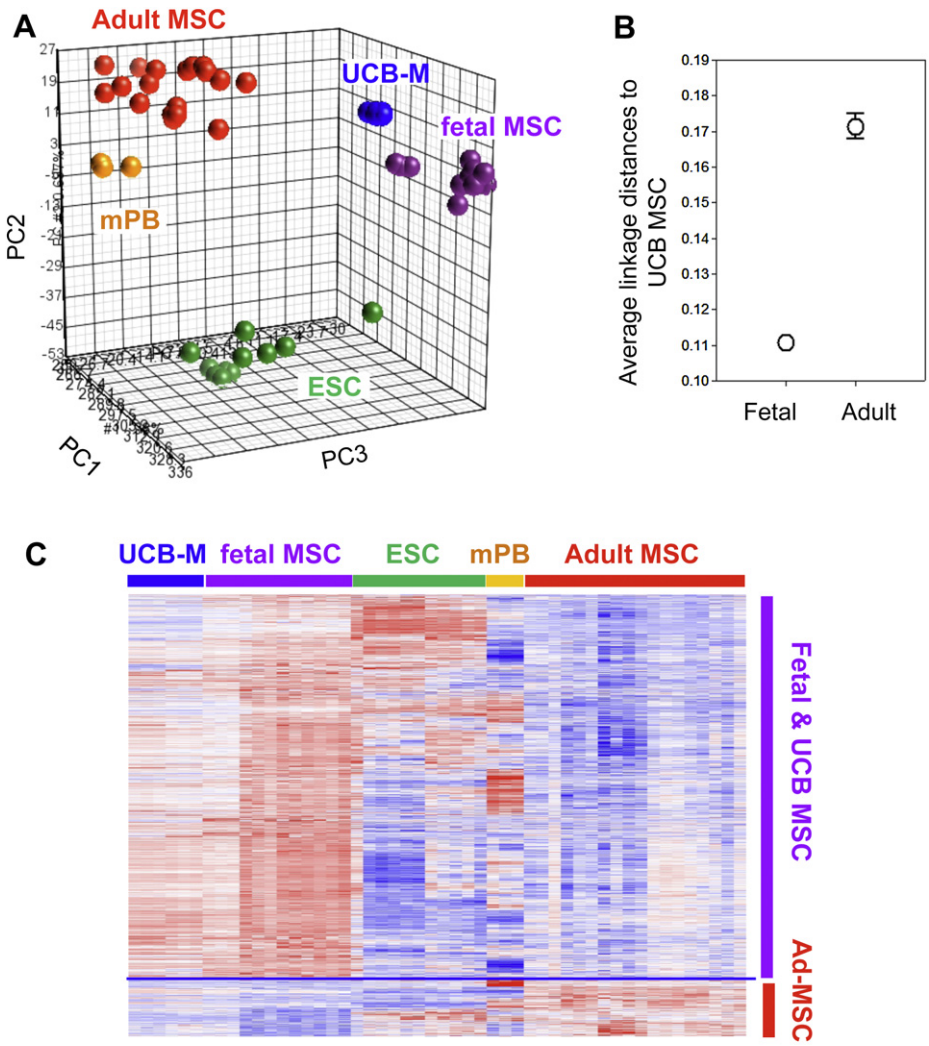

Fig. 4. Relationships between UCB MSCs, fetal MSCs, and adult MSCs. (**A**) A PCA plot using genes that are differentially expressed in adult and fetal MSCs. mPB: maternal PBMC. (**B**) Average linkage Euclidean distances between MSCs. (**C**) A heat map showing the gene distribution patterns among the different MSC groups.

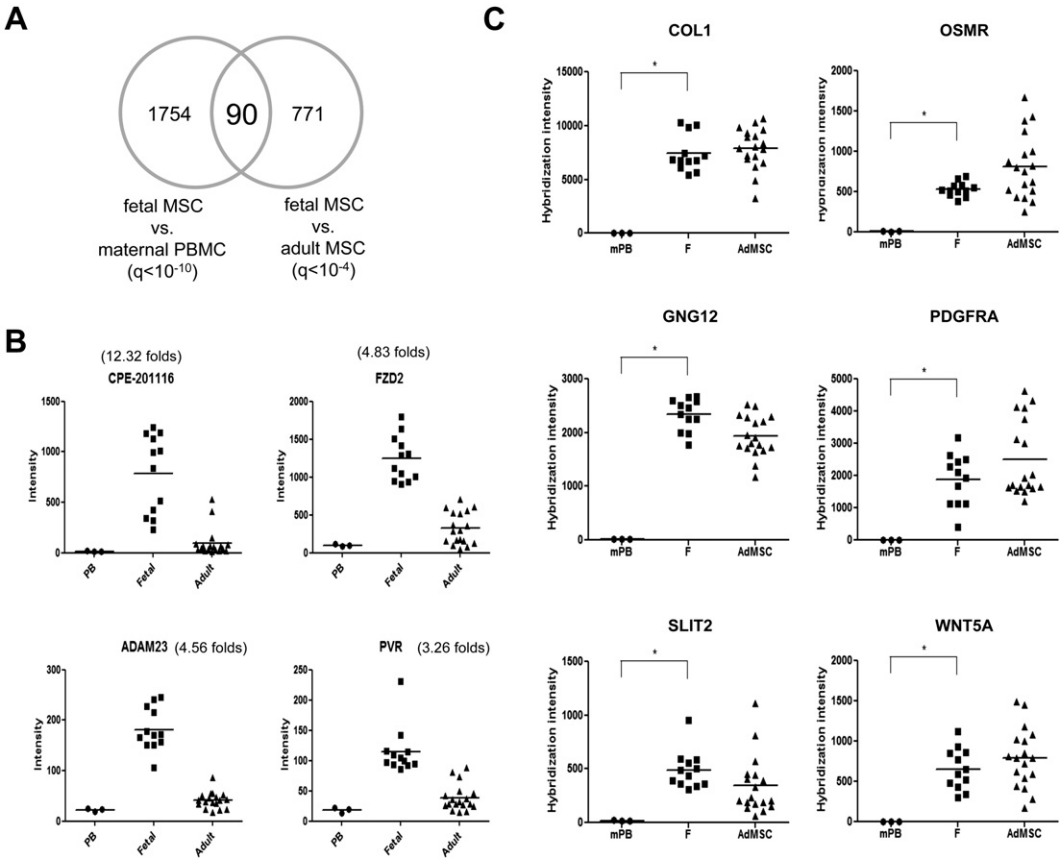

Fig. 5. Potential surface antigens for noninvasive prenatal diagnosis. A Venn diagram showing 90 unique genes in fetal MSCs. (**B**) Scatter plots illustrating cell-surface transcripts that are expressed in fetal MSCs but not in maternal PBMCs (mPBMC) or adult MSCs. (**C**) Genes not in mPBMC, but commonly expressed by both fetal and adult MSCs.

Gotherstrom et al to be specific to fetal MSCs but not to maternal PBMC [25]. Other surface markers described by Gotherstrom et al were not differentially expressed in the adult and fetal MSCs (Fig. 5C).

Discussion

MSCs have attracted a great deal of attention due to their unique stemness and differential abilities, as well as their potential uses in regenerative medicine. Much progress has been made in MSC research, yet the clinical application of these cells is stymied by an insufficient understanding of how to functionally integrate their complex, interactive genetic processes. Further characterization of stem cells via a global, genomics-based approach will be critical to both better understand the molecular events involved in the properties of stem cells and how to utilize these cells. It is now clear that the functional differences between MSC populations could be a function of their transcriptomes.

So far, the origin of UCB MSCs is still subject to debate. Gene signature analyses have showed that UCB MSCs have a distinct gene expression pattern and biological activities and are, therefore, different from MSCs that come from an adult origin [2,14]. Comparative gene expression profiling of MSCs from paired umbilical cord units have also revealed that umbilical cord MSCs are different from UCB MSCs [27]. It has been postulated that both fetal hematopoietic cells and MSCs travel, via the fetal blood, from early fetal hematopoietic sites to newly formed bone marrow [2,13]. It is, therefore, possible that fetal MSCs will circulate in fetal blood and then enter the umbilical cord to become UCB MSCs. The confirmation of transplacental trafficking of other fetal cells via cord blood into the maternal circulation, albeit at a low frequency, further strengthens this hypothesis: various fetal cell types, such as trophoblasts, leukocytes, nucleated erythrocytes, platelets, and hematopoietic progenitors, could be isolated from maternal blood [28–30]. Since the origin of MSCs will determine their functional characteristics, and since hematopoietic stem cells (HSCs) and MSCs have been clinically used in cell-based therapies, the characteristics and classification of UCB MSCs will provide further insights and benefit the selection of cell sources.

Fetal MSCs have also been observed in maternal circulation [31]. The transplacental trafficking of fetal MSCs represents a promising candidate for noninvasive prenatal diagnosis. However, fetal MSCs are rarely present in posttermination maternal blood ($\sim$1 cell per $10^{6\text{-}7}$ maternal PBMCs [28,29]). The identification of the surface antigens specific to fetal MSCs will benefit from the development of novel and noninvasive prenatal diagnosis methods based on positive

selection strategies. A recent genomics study compared the gene expression profiles of fetal MSCs and maternal PBMCs (mPBMCs) as possible surface antigens for noninvasive prenatal diagnosis [25]. Eighteen unique surface proteins, including CPE and OSMR, were filtrated from the array data and confirmed by RT-PCR and immunocytochemistry [25]. However, it has also been demonstrated that adult MSCs can be mobilized into peripheral blood [32,33]. To exclude the contamination of maternal MSCs during positive purification, marker identification of fetal MSCs in maternal blood could be improved by further comparing fetal MSC genetic profiles with those of adult MSCs.

In summary, our findings shed light on common myths surrounding different types of MSCs, provide evidence of the underlying mechanisms, and contribute to the development of cell-based therapies. Studying the genes or pathways that are differentially expressed in various MSCs may eventually decipher the mechanisms of stemness, neurogenesis, angiogenesis, and osteogenesis. Thus, this study provides novel strategies for further functional or mechanistic studies on stem cells. The approaches taken here can also serve as a model for the comparison of other complex biological systems or the functions of a given gene.

Acknowledgments

The authors acknowledge the efforts of International Genomics Consortium (IGC) and expO (Expression Project for Ontology) for GEO database. This work was supported by the Mackay General Hospital (MMH-HB-98-04 and MMH-98106), Taipei Veterans General Hospital (DOH100-TD-C-111-007, Center of Excellence for Cancer Research at TVGH), National Science Council (NSC98-2320-B-010-020-MY3, NSC 99-2627-B-010-010, and NSC98-3111-B-010-004), and Yang-Ming University (Ministry of Education, Aim for the Top University Plan). The authors have no competing interests to declare.

Supplementary material

Supplementary material related to this article can be found online at doi:10.1016/j.tjog.2011.10.009.

References

[1] Mueller SM, Glowacki J. Age-related decline in the osteogenic potential of human bone marrow cells cultured in three-dimensional collagen sponges. J Cell Biochem 2001;82:583−90.
[2] Erices A, Conget P, Minguell JJ. Mesenchymal progenitor cells in human umbilical cord blood. Br J Haematol 2000;109:235−42.
[3] Romanov YA, Svintsitskaya VA, Smirnov VN. Searching for alternative sources of postnatal human mesenchymal stem cells: candidate MSC-like cells from umbilical cord. Stem Cells 2003;21:105−10.
[4] Lee OK, Kuo TK, Chen WM, Lee KD, Hsieh SL, Chen TH. Isolation of multipotent mesenchymal stem cells from umbilical cord blood. Blood 2004;103:1669−75.
[5] Mitchell KE, Weiss ML, Mitchell BM, Martin P, Davis D, Morales L, et al. Matrix cells from Wharton's jelly form neurons and glia. Stem Cells 2003;21:50−60.
[6] Wang HS, Hung SC, Peng ST, Huang CC, Wei HM, Guo YJ, et al. Mesenchymal stem cells in the Wharton's jelly of the human umbilical cord. Stem Cells 2004;22:1330−7.
[7] Panepucci RA, Siufi JL, Silva Jr WA, Proto-Siquiera R, Neder L, Orellana M, et al. Comparison of gene expression of umbilical cord vein and bone marrow-derived mesenchymal stem cells. Stem Cells 2004;22:1263−78.
[8] Chang YJ, Shih DT, Tseng CP, Hsieh TB, Lee DC, Hwang SM. Disparate mesenchyme-lineage tendencies in mesenchymal stem cells from human bone marrow and umbilical cord blood. Stem Cells 2006;24:679−85.
[9] Kern S, Eichler H, Stoeve J, Kluter H, Bieback K. Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood or adipose tissue. Stem Cells 2006;24:1294−301.
[10] Wagner W, Wein F, Seckinger A, Frankhauser M, Wirkner U, Krause U, et al. Comparative characteristics of mesenchymal stem cells from human bone marrow, adipose tissue, and umbilical cord blood. Exp Hematol 2005;33:1402−16.
[11] Song L, Webb NE, Song Y, Tuan RS. Identification and functional analysis of candidate genes regulating mesenchymal stem cell self-renewal and multipotency. Stem Cells 2006;24:1707−18.
[12] Wagner W, Feldmann Jr RE, Seckinger A, Maurer MH, Wein F, Blake J, et al. The heterogeneity of human mesenchymal stem cell preparations: evidence from simultaneous analysis of proteomes and transcriptomes. Exp Hematol 2006;34:536−48.
[13] Tsai MS, Hwang SM, Chen KD, Lee YS, Hsu LW, Chang YJ, et al. Functional network analysis on the transcriptomes of mesenchymal stem Cells derived from amniotic fluid, amniotic membrane, cord blood, and bone marrow. Stem Cells 2007;25:2511−23.
[14] Jansen BJ, Gilissen C, Roelofs H, Schaap-Oziemlak A, Veltman JA, Raymakers RA, et al. Functional differences between mesenchymal stem cell populations are reflected by their transcriptome. Stem Cells Dev 2010;19:481−90.
[15] Hsieh JY, Fu YS, Chang SJ, Tsuang YH, Wang HW. Functional module analysis reveals differential osteogenic and stemness potentials in human mesenchymal stem cells from bone marrow and Wharton's jelly of umbilical cord. Stem Cells Dev 2010;19:1895−910.
[16] Guillot PV, Gotherstrom C, Chan J, Kurata H, Fisk NM. Human first-trimester fetal MSC express pluripotency markers and grow faster and have longer telomeres than adult MSC. Stem Cells 2007;25:646−54.
[17] Guillot PV, De Bari C, Dell'Accio F, Kurata H, Polak J, Fisk NM. Comparative osteogenic transcription profiling of various fetal and adult mesenchymal stem cell sources. Differentiation 2008;76:946−57.
[18] Zhang ZY, Teoh SH, Chong MS, Schantz JT, Fisk NM, Choolani MA, et al. Superior osteogenic capacity for bone tissue engineering of fetal compared with perinatal and adult mesenchymal stem cells. Stem Cells 2009;27:126−37.
[19] Huang TS, Hsieh JY, Wu YH, Jen CH, Tsuang YH, Chiou SH, et al. Functional network reconstruction reveals somatic stemness genetic maps and dedifferentiation-like transcriptome reprogramming induced by GATA2. Stem Cells 2008;26:1186−201.
[20] Chang TY, Li YY, Jen CH, Yang TP, Lin CH, Hsu MT, et al. easyExon: a Java-based GUI tool for processing and visualization of Affymetrix exon array data. BMC Bioinformatics 2008;9:432.
[21] Yang TP, Chang TY, Lin CH, Hsu MT, Wang HW. ArrayFusion: a web application for multi-dimensional analysis of CGH, SNP and microarray data. Bioinformatics 2006;22:2697−8.
[22] Dennis Jr G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.
[23] Wang HW, Trotter MW, Lagos D, Bourboulia D, Henderson S, Makinen T, et al. Kaposi sarcoma herpesvirus-induced cellular reprogramming contributes to the lymphatic endothelial gene expression in Kaposi sarcoma. Nat Genet 2004;36:687−93.
[24] Lee ES, Chan J, Shuter B, Tan LG, Chong MS, Ramachandra DL, et al. Microgel iron oxide nanoparticles for tracking human fetal mesenchymal stem cells through magnetic resonance imaging. Stem Cells 2009;27:1921−31.

[25] Gotherstrom C, Chan J, O'Donoghue K, Fisk NM. Identification of candidate surface antigens for non-invasive prenatal diagnosis by comparative global gene expression on human fetal mesenchymal stem cells. Mol Hum Reprod 2010;16:472−80.
[26] Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res 2004;32:D258−61.
[27] Secco M, Moreira YB, Zucconi E, Vieira NM, Jazedje T, Muotri AR, et al. Gene expression profile of mesenchymal stem cells from paired umbilical cord units: cord is different from blood. Stem Cell Rev 2009;5:387−401.
[28] Bianchi DW, Shuber AP, DeMaria MA, Fougner AC, Klinger KW. Fetal cells in maternal blood: determination of purity and yield by quantitative polymerase chain reaction. Am J Obstet Gynecol 1994;171:922−6.
[29] Little MT, Langlois S, Wilson RD, Lansdorp PM. Frequency of fetal cells in sorted subpopulations of nucleated erythroid and CD34+ hematopoietic progenitor cells from maternal peripheral blood. Blood 1997;89:2347−58.
[30] Krabchi K, Gadji M, Forest JC, Drouin R. Quantification of all fetal nucleated cells in maternal blood in different cases of aneuploidies. Clin Genet 2006;69:145−54.
[31] O'Donoghue K, Choolani M, Chan J, de la Fuente J, Kumar S, Campagnoli C, et al. Identification of fetal mesenchymal stem cells in maternal blood: implications for non-invasive prenatal diagnosis. Mol Hum Reprod 2003;9:497−502.
[32] Rochefort GY, Delorme B, Lopez A, Herault O, Bonnet P, Charbord P, et al. Multipotential mesenchymal stem cells are mobilized into peripheral blood by hypoxia. Stem Cells 2006;24:2202−8.
[33] Chen Y, Xiang LX, Shao JZ, Pan RL, Wang YX, Dong XJ, et al. Recruitment of endogenous bone marrow mesenchymal stem cells towards injured liver. J Cell Mol Med 2010;14:1494−508.