STEM CELLS®
Original Article

The Profile of Gene Expression of Human Marrow Mesenchymal Stem Cells

WILSON A. SILVA, JR., DIMAS T. COVAS, RODRIGO A. PANEPUCCI, RODRIGO PROTO-SIQUEIRA, JORGE L.C. SIUFI, DALILA L. ZANETTE, ANEMARI R.D. SANTOS, MARCO A. ZAGO

Center for Cell Therapy and Regional Blood Center, Department of Clinical Medicine, Faculty of Medicine, Ribeirão Preto, Brazil

Key Words. *Mesenchymal stem cells · SAGE · Gene expression · Hematopoesis*

ABSTRACT

Mesenchymal stem cells (MSCs) are multipotent precursors present in adult bone marrow, that differentiate into osteoblasts, adipocytes and myoblasts, and play important roles in hematopoiesis. We examined gene expression of these cells by serial analysis of gene expression, and found that collagen I, secreted protein acidic and rich in cysteine (osteonectin), transforming growth factor beta- (TGF-β) induced, cofilin, galectin-1, lamininreceptor 1, cyclophilin A, and matrix metalloproteinase-2 are among the most abundantly expressed genes. Comparison with a library of CD34⁺ cells revealed that MSCs had a larger number of expressed genes in the categories of cell adhesion molecule, extracellular and development. The two types of cells share abundant transcripts of many genes, some of which are highly expressed in myeloid progenitors (thymosin-β4 and β10, fos and jun). Interleukin-11 (IL-11), IL-15, IL-27 and IL-10R, IL-13R and IL-17R were the most expressed genes among the cytokines and their receptors in MSCs, and various interactions can be predicted with the CD34⁺ cells. MSCs express several transcripts for various growth factors and genes suggested to be enriched in stem cells. This study reports the profile of gene expression in MSCs and identifies the important contribution of extracellular protein products, adhesion molecules, cell motility, TGF-β signaling, growth factor receptors, DNA repair, protein folding, and ubiquination as part of their transcriptome. ***Stem Cells*** *2003;21:661-669*

INTRODUCTION

In addition to hematopoietic cells, bone marrow (BM) comprises a heterogeneous population of cells that plays a key role in hematopoiesis, referred to as marrow stromal cells, including endothelial cells, adipocytes, osteoblasts and fibroblasts. Mesenchymal stem cells (MSCs) are multipotent precursors present in adult BM, capable of differentiating into osteoblasts, adipocytes, and myoblasts [1-3]. Although they represent only about 0.01%-0.001% of the marrow cells, they can be separated from the hematopoietic stem cells (HSC) because they adhere to glass and plastic [4]. Once in culture they proliferate to originate spindle-shaped cells in confluent cultures, but exhibit a variable morphology and differentiation potential under appropriate conditions. These cells are present in adult BM and peripheral blood, and in the fetal BM and liver [5]. Besides their ability to give rise to cells that constitute the BM stroma, they have been reported to originate glial and neuronal cells, whereas protein and mRNA expression have

Correspondence: Marco A. Zago, M.D., Ph.D., Hemocentro, R. Tenente Catao, Roxo, 2501, 14051-140 Ribeirao Preto, Brazil. Telephone: 55-16-3963-9361; Fax: 55-16-3963-9309; e-mail: marazago@usp.br *Received March 19, 2003; accepted for publication July 25, 2003.*

demonstrated epithelial, endothelial, and neuronal markers. A mesoderm progenitor that gives rise to mesenchymal and to endothelial cells has been identified [6]. Although part of these diverse results is probably caused by the fact that the majority of studies have dealt with heterogeneous cell populations that differ because of the preparation protocols or the time in culture, *Tremain et al*. have demonstrated that markers of various differentiation lineages are concomitantly present in a colony derived from a single MSC [7] providing evidence for their stem-cell (SC) nature.

The wide therapeutic potential of these cells has attracted much attention to them [2, 8], and in vitro and in vivo functional studies and therapeutic trials have been started [9]. However, the transcriptome and broad gene expression profile of a well-defined MSC population has not been described in detail. In addition to the study of *Tremain et al*. [7], the reports have been limited to analyzing the expression of gene families under particular experimental conditions. We have employed serial analysis of gene expression (SAGE) to examine the gene expression of MSC obtained from normal human BM and compared it with the expression profile of $CD34^+$ hematopoietic precursors.

Materials and Methods

Isolation and Culture of Human MSCs

Human MSCs were obtained by aspiration from the iliac crest of a BM donor who gave consent after full information. The mononuclear cells were separated by ficol gradient (Ficoll-Paque; Amersham Biosciences; Peapack, NJ; http://www.bioprocess.amershambiosciences.com), washed in Hank's balanced salt solution, and then cultured in a 25 cm^2 tissue culture flask (Falcon; BD Biosciences; Franklin Lakes, NJ; http://www.bdbiosciences.com) in alpha-minimal essential medium supplemented with 100 μg/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine, and 20% fetal calf serum (Atlanta Biological; Norcross, GA; http://www.atlantabio.com) [10, 11]. After 24 hours the nonadherent cells were removed and the adherent layer cultured until it reached 50%-70% confluence. Cells were then harvested by incubation with 0.2% trypsin-EDTA. Cells collected after the fourth culture passage were stimulated to osteoblast and to adipocyte differentiation as described [2, 10] and used for RNA extraction for SAGE analysis.

Flow Cytometry Analysis

The cells harvested were washed in phosphate-buffered saline, counted pelleted by centrifugation, and resuspended in 100 μl of the appropriate monoclonal antibody and corresponding isotype controls (Pharmingen; San Diego, CA; http://www.bdbiosciences.com/pharmingen). The labeled cells were analyzed on a FACSort by collecting 10,000 events with the Cell Quest software program (Becton Dickinson; San Jose, CA; http://www.bd.com). The antibodies used were CD90-PE, CD51/61-PE, CD29-PE, CD49e-PE, CD49d-PE, CD44-FITC, CD45-FITC, CD13-FITC, HLADR-FITC, HLAclassI-FITC.

SAGE Procedure

Total RNA was prepared from 4×10^7 cells obtained from a fresh cell culture using TRIzol®LS Reagent (Cat No. 10296010; Invitrogen Corporation; Carlsbad, CA; http://www.invitrogen.com) and treated with RQ1 RNase-Free Dnase (Cat. No. M6101; Promega Corporation; Madison, WI; http://www.promega.com) according to manufacturer's instructions. Absence of DNA contamination was ascertained by Southern blot analysis with a mitochondrial DNA marker (D-loop) as a probe, using the treated RNA as template in a polymerase chain reaction (PCR). Thirty μg of total RNA were then used for the SAGE procedure. SAGE was carried out using the I-SAGE™ Kit (Cat. No. T5001-01; Invitrogen) based on the original SAGE [12]. Amplified inserts were sequenced with forward M13 primer in a MegaBACE™ 1000 sequencer and the DYEnamic ET Dye Terminator Sequencing Kit (Cat. No. US81090; Amersham Biosciences; Piscataway, NJ; http://www.amershambiosciences.com).

SAGE Analysis

Tag frequency tables were obtained from sequences by the SAGE™ analysis software, with minimum tag count set to one, maximum ditag length set to 28 bp, and other parameters set as default. The annotation was based on two specific tools, SAGEmap (http://www.ncbi.nlm.nih.gov/SAGE/) and CGAP SAGE Genie (http://cgap.nci.nih.gov/SAGE). We downloaded a SAGE library of $CD34^+$ HSCs purified from BM [13] available as supplemental material in the Proceedings of the National Academy of Science (PNAS) website http://www.pnas.org. When the two libraries were compared, the number of tags was normalized to a total count of 200,000 tags.

Semiquantitative Evalution by Real Time-PCR (RT-PCR)

Total RNA was obtained from seven human tissues. The transcription reaction was performed with 2 μg of total RNA, 0.5 μg of Oligo (dT) primer and 200 U of Superscript II Rnase H Reverse Transcriptase (Invitrogen) in a total volume of 20 μl, and one-tenth of the volume of the cDNA was used in the semiquantitative PCR. The specific primers used are listed in Table 1. When the reaction was positive in the undiluted samples, the cDNA was serially diluted (1:2 to 1:128) before performing the PCR. Secreted protein acidic and rich in cysteine (SPARC) expression was measured by RT-PCR with the *Taqman* approach (Applied Biosystems; Foster City, CA; http://www.appliedbiosystems.com).

Table 1. Sequences of the primer used for RT-PCR amplification of selected genes to corroborate the results obtained by SAGE

UniGene	Primer	Sequence
172928	Col1A1-F	5′-CGCTACTACCGGGCTGATGAT
	Col1A1-R	5′-GTCCTTGGGGTTCTTGCTGATGTA
179573	Col1A2-F	5′-AGGGCAACAGCAGGTTCACTTACA
	Col1A2-R	5′-AGCGGGGGAAGGAGTTAATGAAAC
382367	Lgal1S-F	5′-CCACGGCGACGCCAACACCAT
	Lgal1S-R	5′-TGGGCTGGCTGATTTCAGTCAAAG
297753	VIM-F	5′-TCTATCTTGCGCTCCTGAAAAACT
	VIM-R	5′-AAACTTTCCCTCCCTGAACCTGAG
401448	TPT1-F	5′-ATCCAGATGGCATGGTTGCTCTAT
	TPT1-R	5′-TGCCTCCACTCCAAATAAATCACA
433399	TAGLN-F	5′-CTTTGGGCAGCTTGGCAGTGACCA
	TAGLN-R	5′-CCAGCCCGCTTCTCCCTGCTTAG
406504	TAGLN2-F	5′-AGCGGACGCTGATGAATCTGG
	TAGLN2-R	5′-TGGCTATGGGGAAGGGAATGTATT
111301	MMP2-F	5′-CAGGCACTGGTGTTGGGGGAGAC
	MMP2-R	5′-CCATCGCTGCGGCCAGTATCAGTG
380674	ANXA2-F	5′-GGTCTCCCGCAGTGAAGTGGACAT
	ANXA2-R	5′-GGCCAGGCAATGCTTAGGCAACTA
416073	S100A8-F	5′-GAATTTCCATGCCGTCTACAGG
	S100A8-R	5′-GCCACGCCCATCTTTATCACCAG
169476	GAPDH-F	5′-TTAGCACCCCTGGCCAAGG
	GAPDH-R	5′-CTTACTCCTTGGAGGCCATG
444446	ZNF287-F	5′-CATGGATGGTGATAAAAGAAA
	ZNF287-R	5′-TGAGACATGATTGAAGGTTGC
406409	MYO1F-F	5′-CCAAGCCCCAGCCTCGGACACAT
	MYO1F-R	5′-GGCGGGCGAAAGAGAAGGCAGTAT
79877	MTMR6-F	5′-AAGACTTATTCCCTGTGG
	MTMR6-R	5′-ATTGTTTATTTTGCTCATTC
108885	COL6A1-F	5′-CAAGAGCCTGCAGTGGATGG
	COL6A1-R	5′-CTGGTCTGAGCCTGGGATGAA
238928	HT002-F	5′-GGCCACCTCTAAGACCACACTGAC
	HT002-R	5′-GCCGGCCATTCTAGCGATTTG
75511	CTGF-F	5′-GCCCGGGAAATGCTGCGAGGAGT
	CTGF-R	5′-TGCAGGAGGCGTTGTCATTGGTAA
181163	HMGN2-F	5′-CTGGCAAGGAGGGGAATAACC
	HMGN2-R	5′-GGGAGCCAAGAGGAAGTC
111779	SPARC-F	5′-ACAAGCTCCACCTGGACTACATC
	SPARC-R	5′-GGGAATTCGGTCAGCTCAGA
	SPARC-probe	5′-TTGCAAATACATCCCC

Results

Characteristics of the MSC Population

The cells that assumed a spindle-shaped morphology in confluent wave-like layers at 7 to 14 days of culture were $CD13^+$, $CD29^+$, $CD44^+$, $CD45^-$, $CD49d^-$, $CD49e^+$, $CD90^+$, $CD51/61^-$, $HLAClassI^+$, and $HLADR^-$. When cultured with dexamethasone and ascorbic acid, they underwent osteogenic differentiation, as demonstrated by positive calcium staining by the von Kossa reaction, whereas in culture with insulin, dexamethasone and indomethacin, they originated adipocytes, identified by vacuoles that stained positively with Sudan III. They thus have the distinguishing characteristics of the MSC [2].

Gene Expression of MSC

A total of 102,796 tags were obtained by sequencing. Excluding redundancy, these results correspond to 34,649 unique tags, 22,343 of which matched known genes or expressed sequence tags (ESTs) in the CGAP SAGE Genie mapping (84,364 total tags corresponding to 15,167 UniGene clusters), whereas 12,306 unique tags were no matches (18,432 total tags). The 50 most abundant transcripts are listed in Table 2. Some are known to be highly expressed genes in this type of cell, whereas others are recognized here for the first time.

A list of all the tags found in the MSC can be found at our website http://bit.fmrp.usp.br/msc_tags/.

Comparison of MSC with $CD34^+$ Cells

The 1,000 most abundant tags of each of the two types of progenitor cells (our MSC library and the downloaded library obtained from $CD34^+$ cells) were compared directly with the complete list of tags of the other cell type. This comparison revealed 607 tags exclusively expressed in $CD34^+$ hematopoietic precursors, 602 exclusively expressed in MSCs and 791 tags common to both, 393 of which were more expressed in $CD34^+$ cells and 398 more expressed in MSCs (Table 3). A search of gene ontology (GO) terms was performed for 549 and 489 unique tags among the 1,000 more expressed respectively in MSCs and CD34 cells. The search revealed that MSCs, as compared with $CD34^+$ cells, had a higher percentage of genes in the categories of "cell adhesion" (6.1% × 1.6%), "extracellular" (11.1% × 2.9%) and "development" (11.4 × 7.3%) ($p < 0.05$). When compared with the number of the gene products annotated under a specific term for the whole GO, MSCs had a higher percentage of genes in "cell adhesion" (0.4% × 6.1%), "extracellular" (6.5% × 11.1%), "cell motility" (1.97% × 4.0%), and "metabolism" (48.7% × 65.0%). A comparison of the two types of cells concerning genes expressed for cell adhesion, extracellular, and motility is shown in Table 4.

Corroboration of SAGE Results

Gene expression was measured semiquantitatively by RT-PCR in seven different tissues (MSCs, a sample of partly purified $CD34^+$ cells [66% purity], bulk normal BM, brain, skeletal muscle, leukocytes and liver) for the following 11 genes: COL1A1, COL1A2, matrix metalloproteinase 2 (MMP2), TPT1, LGAL1S, TGLN, TGLN2, SPARC, vimentin (VIM), ANXA2, and S100A8. The results correlated

Table 2. List of the 50 most abundant tags in mesenchymal cell SAGE library

Tag	Count*	UniGene Cluster Hs.		Descriptions**
		CGAP	SAGEmap	
TGGAAATGAC	1319	172928	172928, 444446	**Collagen type I alpha 1**, ESTs weakly similar to zinc finger protein ZNF287
ATGTGAAGAG	770	111779	111779, 406409	**Secreted protein acidic cysteine-rich (osteonectin)**, Myosin IF
GTGTGTTTGT	737	118787	118787	**Transforming growth factor β-induced**
CCCATCGTCC	708	294286	—	EST, No match
GAAGCAGGAC	628	180370	180370	**Cofilin 1 (non-muscle)**
TACCATCAAT	620	169476	169476, 79877	**Glyceraldehyde-3-phosphate dehydrogenase**, Myotubularin related protein 6
GAGGGAGTTT	583	76064	76064, 356342	**Ribosomal protein L27a**, ESTs highly similar to S55914 ribosomal protein L27a
GCCCCCAATA	576	382367	382367	**Galectin 1**
CTGGGTTAAT	569	298262	298262, 97457	**Ribosomal protein S19**, ESTs
GGGCTGGGGT	560	430207	90436, 430207	Ribosomal protein L29, Sperm associated antigen 7
GCATAATAGG	489	431927	431927, 356482	**Ribosomal protein L21**, ESTs highly similar to 2113200B ribosomal protein L21
TTGGTCCTCT	419	356795	356795, 427046	**Ribosomal protein L41**, E1BP1 pseudogene mRNA sequence
GAAAAATGGT	411	181357	181357, 356267	**Laminin receptor 1**, ESTs moderately similar to RSP4_HUMAN 40S ribosomal protein SA (P40)
TAGGTTGTCT	376	401448	401448	**Tumor protein translationally controlled 1**
CCTAGCTGGA	374	401787	401787	**Peptidylprolyl isomerase A (cyclophilin A)**
TTTGGTTTTC	372	179573	179573, 21431	**Collagen type I alpha 2**, Suppressor of fused homolog (Drosophila)
TGTGTTGAGA	369	422118	422118, 356428	**Eukaryotic translation elongation factor 1 alpha 1**, mRNA expressed only in placental villi, clone SMAP83.
GGAAATGTCA	359	111301	111301	**Matrix metalloproteinase 2 (gelatinase A)**
TTGCTGACTT	349	108885	108885, 238928	**Collagen type VI alpha 1**, HT002 protein hypertension-related calcium-regulated gene
AGCACCTCCA	338	75309	75309	**Eukaryotic translation elongation factor 2**
TGTACCTGTA	337	334842	334842, 249922	**Tubulin alpha ubiquitous**, EST
TTGGGGTTTC	334	418650	418650, 356581	**Ferritin heavy polypeptide 1**, ESTs highly similar to ferritin heavy chain
TTCATACACC	327	330871	—	?, No match
GGCTGGGGGC	321	408943	408943, 352407, 265848	**profilin 1**, Chromosome 1 amplified sequence 3, Phosphodiesterase 4D interacting protein (myomegalin)
TGCACGTTTT	320	169793	169793	**ribosomal protein L32**
ACTTTTTCAA	307	328511	133430, 156814	?, ESTs, KIAA0377 gene product
TCCAAATCGA	290	297753	297753	**Vimentin**, Ribosomal protein large P2
GTGTGTTGTA	285	422118	—	**Eukaryotic translation elongation factor 1 alpha**, No match
ATCTTGTTAC	280	287820	287820	**Fibronectin 1**
TAATAAAGGT	278	399720	399720, 355986	**Ribosomal protein S8**, ESTs highly similar to S25022 ribosomal protein S8
TAAGGAGCTG	277	299465	299465, 50651	**Ribosomal protein S26**, Janus kinase 1 (a protein tyrosine kinase)
TTTGCACCTT	269	75511	75511, 181163	**Connective tissue growth factor**, High-mobility group nucleosomal binding domain 2
AGGGCTTCCA	268	412900	412900, 356769	**Ribosomal protein L10**, ESTs moderately similar to RL10_HUMAN 60S ribosomal protein L10

Table 2. List of the 50 most abundant tags in mesenchymal cell SAGE library (continued)

Tag	Count*	UniGene Cluster Hs.		Descriptions**
		CGAP	SAGEmap	
GGCAAGCCCC	262	425293	425293, 187577	**Ribosomal protein L10a**, SRY (sex determining region Y)-box 21
GGATTTGGCC	258	297753	297753	**Ribosomal protein large P2**
GATGAGGAGA	247	179573	179573, 426977	**Collagen type I alpha 2**, Ribosomal protein L19
CTAGCCTCAC	244	14376	14376, 356575	**Actin gamma 1**, ESTs
GGGAAGCAGA	241	12284	—	**F11 receptor**, No match
ATAATTCTTT	238	539	539, 356463	**Ribosomal protein S29**, Farnesyltransferase CAAX box alpha
GGAGTGTGCT	231	9615	9615	**Myosin light polypeptide 9 regulatory**
CACCTAATTG	228	301104	—	?, No match
TTGGTGAAGG	228	75968	75968, 426138	**Thymosin beta 4 X chromosome**, promyelocytic leukemia cell mRNA (clones pHH58 and pHH81)
GTTGTGGTTA	219	48516	48516, 99785	**Beta-2-microglobulin**, cDNA: FLJ21245 fis clone COL01184
TCAGATCTTT	212	389933	389933, 85112	**Ribosomal protein S4 X-linked**, Insulin-like growth factor 1 (somatomedin C)
ACAGGCTACG	209	433399	433399	**Transgelin**
GGACCACTGA	209	119598	119598, 300622	**Ribosomal protein L3**, Plexin A2
TTCAATAAAA	207	424299	424299, 2012	**Ribosomal protein large P1**, Transcobalamin I (vitamin B12 binding protein, R binder family)
CCAGAACAGA	205	334807	334807	**Ribosomal protein L30**
CAATAAATGT	202	337445	337445, 446130	**Ribosomal protein L37**, ESTs
TACCTGCAGA	199	416073	416073	**S100 calcium binding protein A8**

*Non-normalized counts (actual numbers of reads)

**Bold: UniGene cluster annotated by CGAP that coincides with the cluster with the higher score or the single cluster in the SAGEmap annotation. Underlined: gene expression validated by RT-PCR or real time PCR. ?: the cluster indicated was not found.

with gene expression profiles derived from SAGE data obtained from the SAGEmap site for 10 of the 11 genes tested (the only exception was S100A8). Expression of collagen type 1, alpha 1, and alpha 2, transgelin and MMP2 was detected only in MSCs, even when MSC RNA was diluted to 1:128. Except for tumor protein translationally controlled 1 (TCTP), the other genes tested were more markedly expressed in MSC. For instance, galectin-1 was at least 10 times more expressed in MSCs than in the other tissues evaluated. RT-PCR for galectin-1 was positive in MSC RNA diluted 1:128 times (1,125 tags); it was positive up to the 1:8 dilution in CD34 cells (16 tags) and liver (27 tags), up to the 1:16 dilution for bulk BM (36 tags) and brain (61 tags), and up to 1:64 dilution for peripheral blood leukocyte (92 tags) and skeletal muscle (115 tags). Similar results were obtained for VIM, annexin 2, transgelin 2, and TCTP 1. Figure 1 exemplifies these results. Additionally, for five unique tags for which there were two possible SAGEmap annotations, RT-PCR with specific primers confirmed the CGAP annotation: COL1A1 × ZNF287, SPARC × MYO1F, GAPD × MTMR6, COL6A1 × HT002, and connective tissue growth factor (CTGF) × high mobility group protein-N2 (HMGN2).

Discussion

The MSCs and the HSCs are mesoderm-derived and considered to belong to two independent differentiation pathways, each with its own precursor, although there is evidence for a common precursor or for trilineage hematopoietic recovery of totally irradiated dog transplanted with $CD34^-$ fibroblast-like SCs [14-16]. MSCs have been demonstrated to engraft after transplantation, with partial correction of osteogenesis imperfecta [17, 18]. Preliminary results of co-transplantation with HSCs suggest a faster engraftment and a lower incidence of graft-versus-host disease [19, 20].

The transcriptome of MSC reveals both significant differences and similarities with the $CD34^+$ hematopoietic

Table 3. Comparison of MSC and CD34+ marrow cell gene expression, as measured by the number of tags of the corresponding genes (ribosomal proteins excluded)

Exclusively or highly represented in MSC[1]

collagens type I and type VI, SPARC, matrix metalloproteinase 2 (gelatinase), transforming growth factor β1 induced, eukaryotic translation elongation factor 1 alpha 1, fibronectin 1, light peptide 9 of myosin, transgelin, calgranulin A, heat shock protein 47, latent transforming growth factor β binding protein 2, gap junction protein alpha 1, biglycan, annexin A2, IGF binding proteins 4 and 6, hemeoxygenase 1, tropomyosin 2, connective growth factor, brain abundant membrane attached signal protein 1, galectin 1, nexin, integrins alpha 2 and alpha V, endocytic receptor, CD151 antigen

Exclusively or highly represented in CD34+ stem cells[2]

CD2 antigen, kinesin family member 5, MAP2K3, MHC class II DR alpha, eukaryotic translation elongation factor 3, aldolase C fructose-biphosphate, pulmonary surfactant associated protein C, glucose phosphate isomerase, myeloperoxydase, phosphatic acid phosphatase type 2A

Highly represented in both types of cells[3]

filamin alpha, early growth response, annexin 2 ligand (calpactin I), serotonin receptor 1 D, calcyclin, calgizzarin, cofilin 1, COX8, tropomyosin 1, 5′3′ nucleotidase, benzodiazepine receptor, IGF binding protein 7, cysteine and glycine rich protein 1, vigilin, MAP2K2, pyruvate kinase, phosphoglycerate mutase 1, integrin-linked kinase, cyclophilin A and B, vimentin, thymosin β10, milk fat globule EGF 8 protein, zyxin, heavy polypeptide 1 of ferritin, glyceraldehyde-3-phospate dehydrogenase, heat shock 70 kD, light peptide 6 of myosin, high motility group protein 1, Erb-b3, CD74 antigen, cell division cycle-2 like 5, 5′-nucleotidase, transmembrane gamma-carboxyglutamic acid protein 4, B-cell CLL/lymphoma 7A

[1]Tags for these transcripts were at least 50 times more abundant in the MSC as compared with CD34.

[2]Tags for these transcripts were at least 50 times more abundant in the CD34 as compared with MSC.

[3]Tags correspond to at least 0.05% of total tags from one of the cells, and the relative abundance in one type of cell does not exceed 50 times the other.

Table 4. Comparison of genes expressed in MSC and in CD34+ cells for selected terms of gene ontology. Derived from the 1,000 most abundant tags in each type of cell.

Cell adhesion molecules

MSC: Laminin receptor 1, Integrin β1 (fibronectin receptor), Integrin alpha V (vitronectin receptor), Collagens (type I alpha 1, type III alpha 1, type IV alpha 1, type V alpha 1, type VI alpha 1, type VI alpha 2, type VI alpha 3, type VII alpha 1, type XVI alpha 1), Transforming growth factor β-induced, Connective tissue growth factor, Chondroitin sulfate proteoglycan 2 (versican), Fibronectin 1, Activated leukocyte cell adhesion molecule (ALCAM), Milk fat globule-EGF factor 8 protein, Lysyl oxidase-like 2, MIC2, CD151 antigen, RAC1, RAB13, Protein tyrosine kinase 7, Ninjurin 1, Vinculin, Osteoblast specific factor 2, Syndecan 2, Zyxin, Cadherin 11

CD34: Laminin receptor 1, Collagen type I alpha 1, CD164 antigen (sialomucin), Selectin L, Ninjurin 2, Selectin L, Macrophage erythroblast attacher, Integrin cytoplasmic domain-associated protein 1, Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4

Extracellular

MSC: Collagens (type I alpha 1, type III alpha 1, type IV alpha 1, type V alpha 1, type VI alpha 1, type VI alpha 2, type VI alpha 3, type VII alpha 1, type XVI alpha 1), SPARC (osteonectin), Insulin-like growth factor binding proteins 3, 4 and 6, Biglycan, Fibrillin 1, Fibronectin 1, Lysozyme, Macrophage migration inhibitory factor, Calgranulins A and B, Stanniocalcin 2, Lumican, Chondroitin sulfate proteoglycan 2 (versican), Granulin, Prosaposin, Connective tissue growth factor, 5-hydroxytryptamine receptor 1D, Tissue factor pathway inhibitor 2, Transforming growth factor β1, Lysyl oxidase-like 1 and 2, Follistatin-like 1, Nucleobindin 1, Matrix metalloproteinases 2 and 19, Galectin 3, Transforming growth factor β-induced, Dickkopf homolog 3, Amyloid β (A4) precursor protein, Tissue inhibitor of metalloproteinase 1 and 3, Microfibrillar-associated protein 2, Tumor protein translationally controlled 1, Cysteine-rich angiogenic inducer 61, PRSS11 (IGF binding), CRI1, ACPP, TM4SF7

CD34: Collagen type I alpha 1, Tumor protein, translationally controlled 1, Prosaposin, Calgranulin B, Nucleobindin 2, Lysozyme, Macrophage migration inhibitory factor, Tissue factor pathway inhibitor 2, Chorionic somatomammotropin hormone 1, Surfactant, pulmonary-associated protein C, DEF6, Ribonuclease RNase A family 2, Chondroitin sulfate proteoglycan 6 (bamacan)

Cell motility

MSC: Defensin alpha 1, Actins alpha 1, alpha 4 and β, Actin related protein 2/3 complex subunit 1B and subunit 2, Tumor necrosis factor receptor superfamily member 12 A, Annexin A1, Crystallin alpha B, Connective tissue growth factor, Moesin, Gap junction protein alpha 1, RAC1, Tropomyosin 2, Tropomyosin 1, Calgranulin A, Filamin A alpha, Fibronectin 1, Aldolase A

CD34: Selectin L, Poly(A) binding protein nuclear 1, RalA binding protein 1, Crystallin alpha B, Aldolase A, Actin β, Annexin A1, Actin related protein 2/3 complex subunits 1B and 2, Nebulin-related anchoring protein, Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4

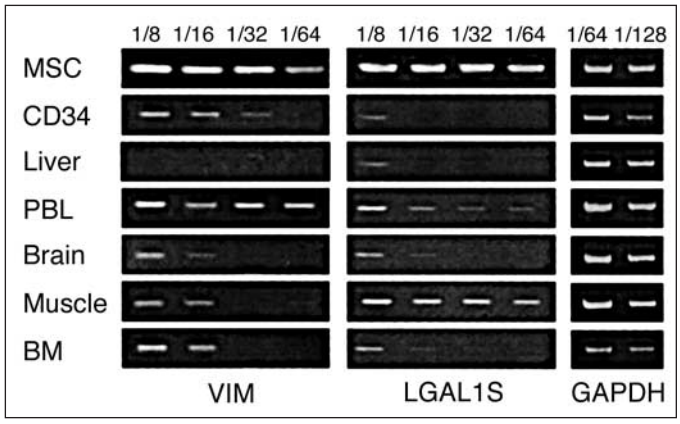

Figure 1. Semiquantitative evaluation of mRNA abundance by RT-PCR. *Total RNA was diluted 1/2 to 1/128 (only 1/8 to 1/64 dilutions are shown), reverse transcribed into cDNA and then a 30-cycle PCR with specific primers located in different exons was performed. At the left is shown the reaction for vimentin (VIM) and at the right the reaction for galectin 1 (LGAL1S). MSC: mesenchymal stem cells; CD34: CD34⁺ hematopoietic progenitor cells; Liver: adult human liver; PBL: peripheral blood leukocytes; Brain: control human normal brain (temporal); Muscle: skeletal muscle; BM: normal human unfractioned bone marrow. A control with GAPDH primers was carried out and gave positive results up to 1/128 dilution for all samples (last two columns).*

precursor. One-third of the most abundant gene products of one cell type is also detected in the other, while about two-thirds are exclusively or significantly overexpressed in one type of cell. Most of the highly expressed genes in MSC are related to extracellular components, receptors to matrix components, and cell adhesion molecules (CAMs), such as collagens, SPARC, galectin-1, laminin receptor, fibronectin and MMP2. SPARC is also found in fibroblasts, in cells derived from the MSC (osteoblasts and condrocytes) [21], and in cells derived from hematopoietic precursor (megakaryoblast and platelet). Galectin-1 (β-galactoside binding protein) is involved in regulation of cell adhesion, cell proliferation, and cell death of T-cells [22], B-cells [23], and the muscular differentiation of dermal fibroblasts [24]. Transforming growth factor beta (TGF-β)-induced is the third most abundant transcript, thus confirming the important role of the TGF-β signaling pathway in this cell population [25, 26], although only two tags specific for endoglin have been detected. The finding that activin A (a cytokine of the TGF-β superfamily) and its receptors are expressed moderately in the MSC agrees with the suggestion that it may influence the growth of stromal cells in an autocrine fashion, whereas only activin receptors were found in CD34 cells [27].

A comparison of Gene Ontology™ (GO) [28] terms between the two libraries and with all gene products in GO revealed that the number of genes expressed in the categories of CAMs and metabolism are over-represented both in MSCs and CD34. Those in the categories of extracellular, cell motility, and cell proliferation are over-represented in MSCs, and genes in the categories of extracellular and development are under-represented in CD34 cells. The most expressed adhesion molecule in the two types of cells is laminin-1 receptor, suggesting that it may contribute to colocalization of the cells in postnatal BM, in addition to other adhesion molecules that may have a homing function, such as CD44 (H-CAM), CD47, and integrins alpha 4 and alpha E. Also highly expressed in MSCs are the genes for integrin (alpha V component of vitronectin receptor and alpha 2 component of VLA or glycoprotein I/II), CD151 antigen TGF-β-induced, osteoblast specific factor 2, milk fat globule-epidermal growth factor (EGF) 8 protein (also known as medin and lactadherin), and activated leukocyte cell adhesion molecule (ALCAM). Some of these molecules, such as laminin receptor and integrins, participate also in cell surface signaling. The most striking difference between the top expressed genes of the two types of SCs is the number of genes related to cell adhesion and extracellular component.

There are also various similarities between MSCs and CD34⁺ cells, which include specialized genes such as filamin, calpactin, calcyclin, cofilin, insulin-like growth factor-binding protein 7, VIM, prosaposin, lysozyme and macrophage migration inhibitory factor. Abundant transcripts were found in the two cell types for genes that are highly expressed in myeloid progenitors (CD15⁺) [29], such as thymosin-β4 and thymosin-β10, which are involved in the differentiation of granulocytes, monocytes and lymphocytes. Recently *Tsai* and *McKay* associated nucleostemin with cell-cycle progression in stem and cancer cells [30]. This protein is present in nucleoli of nervous and embryonic stem cells, in several cancer cell lines, and is preferentially expressed in other SC-enriched populations We found transcripts for its gene in the two cell types (18 tags in MSC and 10 tags in CD34⁺ cells).

Cytokine and growth factor signaling is an important determinant of the functional state of these cells and of the relationship between MSC and CD34⁺ progenitors. We found 30 unique tags for ILs, their receptors, and related proteins that were enriched in the two progenitors cells: 10 were shared by the two types of cells, whereas 10 were exclusive of MSC and 10 were exclusive of CD34. The most abundant transcript in this category was that for the IL-10 receptor both in CD34⁺ and MSCs. IL-1 is produced by the CD34⁺ cells, whereas MSCs have moderate expression of genes for IL-1 receptor and IL-1 receptor-associated kinase 1. Three IL genes were actively expressed in MSC—IL-11, IL-15 and IL-27—whereas CD34 cells have receptors for IL-11. Other IL receptors detected in MSCs are those for IL-9, IL-13 and for IL-17, which was found also in CD34 and plays a role in hematopoietic regulation.

There were also 669 tags for various growth factors, such as stem cell growth factor, TGF-β1, CTGF, hepatoma-derived growth factor, midkine (neurite growth-promoting factor 2), fibroblast growth factor 2, platelet derived growth factor C, and endothelial cell growth factor 1.

Finally, we found at least 6,300 tags related to genes from six of the seven categories indicated by *Ramalho-Santos et al*. [26] as basic characteristics of "stemness": A) Notch, Yes, JAK/STAT and TGF-β pathways; B) seven genes related to interaction with the extracellular matrix; C) ubiquination pathways, protein folding, and DNA repair; D) cell cycle and cell cycle control; E) DNA helicases and histone deacetylases, and F) RNA helicases. The strategy of comparing unfractionated BM cells with the mesenchymal and hematopoietic progenitor cells (results not shown) did not reveal a common set of transcripts enriched in the more primitive cells. These findings seem to strengthen the suggestion that although some similar genes may be active in more than one SC type, there is not a rigid pattern that can be associated with the signature of "stemness" for all the SCs, since related but not identical genes may perform the same function in different SCs, and "stem" or progenitor cells of different tissues probably do not have an equivalent collection of expressed genes.

Thus we report the profile of gene expression in MSC from adult BM in culture and find both similarities and differences with $CD34^+$ progenitors. Although the majority of the results probably reflect the gene expression inherent to this particular cell type, we cannot rule out the possible effect of culture-induced changes on gene expression. The study identifies the important contribution of extracellular protein products, adhesion molecules, cell motility, growth factor receptors, DNA repair, protein folding, and ubiquination as part of the transcriptome of these cells. However, when extrapolating these results to MSCs of other origins, it is necessary to take into consideration possible differences that depend on the anatomical site of the cell [31]. Our results must be viewed from the perspective that large-scale gene expression profiles are more adequate to propose the rationale for future hypothesis-driven studies than to provide a direct explanation for the cell functioning and behavior [32].

ACKNOWLEDGMENTS

The authors would like to thank *Maristela Orellana, Amelia G. Araujo, Marli H. Tavela, Cristiane A. Ferreira, Fernanda G. Barbuzano,* and *Adriana A. Marques* for their assistance with the laboratory techniques, and *Israel T. Silva, Marco V. Cunha,* and *Daniel G. Pinheiro* for their help with the bioinformatic analysis.

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Financiadora de Estudos e Projetos (FINEP), Brazil.

REFERENCES

1 Gronthos S, Graves SE, Ohta S et al. The STRO-1^+ fraction of adult human bone marrow contains the osteogenic precursors. Blood 1994;84:4164-4173.

2 Pittenger MF, Mackay AM, Beck SC et al. Multilineage potential of adult human mesenchymal stem cells. Science 1999;284:143-147.

3 Bianco P, Riminucci M, Gronthos S et al. Bone marrow stromal stem cells: nature, biology, and potential applications. STEM CELLS 2001;19:180-192.

4 Colter DC, Class R, DiGirolamo CM et al. Rapid expansion of recycling stem cells in cultures of plastic-adherent cells from human bone marrow. Proc Natl Acad Sci USA 2000;97:3213-3218.

5 Campagnoli C, Roberts IAG, Kumar S et al. Identification of mesenchymal stem/progenitor cells in human first-trimester fetal blood, liver, and bone marrow. Blood 2001;98:2396-2402.

6 Reyes M, Lund T, Lenvik T et al. Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. Blood 2001;98:2615-2625.

7 Tremain N, Korkko J, Ibberson D et al. MicroSAGE analysis of 2,353 expressed genes in a single cell-derived colony of undifferentiated human mesenchymal stem cells reveals mRNAs of multiple cell lineages. STEM CELLS 2001;19:408-418.

8 Abkowitz JL. Can human hematopoietic stem cells become skin, gut, or liver cells? N Engl J Med 2002;346:770-772.

9 Strauer BE, Brehm M, Zeus T et al. Repair of infarcted myocardium by autologous intracoronary mononuclear bone marrow cell transplantation in humans. Circulation 2002;106:1913-1918.

10 Goodwin HS, Bicknese AR, Chien SN et al. Multilineage differentiation activity by cells isolated from umbilical cord blood: expression of bone, fat, and neural markers. Biol Blood Marrow Transplant 2001;7:581-588.

11 Jaiswal N, Haynesworth SE, Caplan AI et al. Osteogenic differentiation of purified, culture-expanded human mesenchymal stem cells in vitro. J Cell Biochem 1997;64:295-312.

12 Velculescu VE, Zhang L, Vogelstein B et al. Serial analysis of gene expression. Science 1995;270:484-487.

13 Zhou G, Chen J, Lee S et al. The pattern of gene expression in human CD34(+) stem/progenitor cells. Proc Natl Acad Sci USA 2001;98:13966-13971.

14 Singer JW, Chambord P, Keating A et al. Simian virus 40-transformed adherent cells from human long-term marrow

cultures: cloned cell lines produce cells with stromal and hematopoietic characteristics. Blood 1987;70:464-474.

15 Singer JW, Keating A, Cuttner J et al. Evidence for a stem cell common to hematopoiesis and its in vitro microenvironment: studies of patients with clonal hematopoietic neoplasia. Leuk Res 1984;8:535-545.

16 Huss R, Lange C, Weissinger EM et al. Evidence of peripheral blood-derived, plastic-adherent $CD34^{-/low}$ hematopoietic stem cell clones with mesenchymal stem cell characteristics. STEM CELLS 2000;18:252-260.

17 Nilsson SK, Dooner MS, Weier HU et al. Cells capable of bone production engraft from whole bone marrow transplants in nonablated mice. J Exp Med 1999;189:729-734.

18 Horwitz EM, Prockop DJ, Gordon PL et al. Clinical responses to bone marrow transplantation in children with severe osteogenesis imperfecta. Blood 2001;97:1227-1231.

19 Lazarus HM, Curtin P, Devine S. Role of mesenchymal stem cells in allogeneic transplantation: early phase I clinical results. Blood 2000;96:1691a.

20 Frassoni F, Labopin M, Bacigalupo A. Expanded mesenchymal stem cells (MSC), co-infused with HLA identical hematopoietic transplants, reduce acute and chronic graft versus host disease: a matched pair analysis. EBMT Meeting, Montreux, 2002.

21 Chandrasekhar S, Harvey AK, Johnson MG et al. Osteonectin/SPARC is a product of articular chondrocytes/cartilage and is regulated by cytokines and growth factors. Biochim Biophys Acta 1994;1221:7-14.

22 Pace KE, Lee C, Stewart PL et al. Restricted receptor segregation into membrane microdomains occurs on human T cells during apoptosis induced by galectin-1. J Immunol 1999;163:3801-3811.

23 Gauthier L, Rossi B, Roux F et al. Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. Proc Natl Acad Sci USA 2002;99:13014-13019.

24 Goldring K, Jones GE, Thiagarajah R et al. The effect of galectin-1 on the differentiation of fibroblasts and myoblasts in vitro. J Cell Sci 2002;115:355-366.

25 Chen CZ, Li M, de Graaf D et al. Identification of endoglin as a functional marker that defines long-term repopulating hematopoietic stem cells. Proc Natl Acad Sci USA 2002;99:15468–15473.

26 Ramalho-Santos M, Yoon S, Matsuzaki Y et al. "Stemness": transcriptional profiling of embryonic and adult stem cells. Science 2002;298:597-600.

27 Shav-Tal Y, Zipori D. The role of activin A in regulation of hemopoiesis. STEM CELLS 2002;20:493-500.

28 http://www.geneontology.org/

29 Lee S, Zhou G, Clark T et al. The pattern of gene expression in human $CD15^{+}$ myeloid progenitor cells. Proc Natl Acad Sci USA 2001;98:3340-3345.

30 Tsai RY, McKay RD. A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. Genes Dev 2002;16:2991-3003.

31 Charbord P, Oostendorp R, Pang W et al. Comparative study of stromal cell lines derived from embryonic, fetal, and postnatal mouse blood-forming tissues. Exp Hematol 2002;30:1202-1210.

32 Peng CF, Wei Y, Levsky JM et al. Microarray analysis of global changes in gene expression during cardiac myocyte differentiation. Physiol Genomics 2002;9:145-155.