The proteomic expression program of human bone marrow multipotent stromal cells: donor-to-donor heterogeneity and the proteomic changes during cell propagation

Samuel T Mindaye*, **Moonjin Ra***, **Jessica Lo Surdo****, **Steven Bauer****, **Michail Alterman***

*Tumor Vaccines and Biotechnology Branch, **Cellular and Tissue Therapies Branch, Division of Cellular and Gene Therapies, Center for Biologics Evaluation and Research, Food and Drug Administration, 8800 Rockville Pike, Bethesda, Maryland 20892

CBER

Background

Multipotent stromal cells (MSCs) can be isolated from various sources including bone marrow. They hold huge potential for cell-based therapy for a wide array of diseases [1-3].

They are heterogenous cell populations and their isolation to homogeneity for clinical applications and basic research has been quite a challenge. Their *in vitro* differentiation into the three mesenchymal lineages, namely, adipogenic, chondrogenic, and osteogenic cells has been described [1-3].

Despite the intensity of focus to use them in regenerative medicine, their biology has not been fully described.

Advances in the area of genomics, trnscriptomics, and proteomics have offered opportunities to explore MSCs in a systematic and integrated ways.

Studies dealing with MSCs so far have been based only on proteins that are highly abundant and are insufficient to reveal important cellular events [4-6].

The present work aims to:

(1) Significantly improve the large scale proteomic profiling of MSCs derived from human bone marrow (hBM-MSCs)
(2) Describe results of the comparative proteomic screening of cells obtained from six donors. The work highlights the proteomic basis of donor-to-donor variability in hBM-MSCs.
(3) Perform an extensive bioinformatic analyses to gain insight into the highly enriched signaling events and protein-protein interaction networks that can potentially influence the biology of culture-expanded and undefferentiated hBM-MSCs.
(4) Study the change in protein expression program when hBM-MSCs are subjected to in vitro expansion stress

The analytical workflow involves pressure cycling-based protein extraction, an overall 3D proteome fractionation steps, and multiple MS, namely, LC-ESI-MS/MS, nanoLCESI-MS^E, and LC/MALDI-MS/MS for peptide sequencing and protein identification.

Summary of the hBM-MSCs proteome

Cell culturing, protein harvesting, protein identification, and characterization using proteomic approaches

Table 1: Description of the hBM-MSCs source

Donor #	Cell line	Sex	Age	Race
1	PCBM1641	F	23	Hispanic
2	167696	F	22	Hispanic
3	PCBM1632	M	24	Black
4	110877	M	22	Black
5	8F3560	F	24	Hispanic
6	PCBM1662	F	31	Caucasian

Table 2: Proteome coverage of hBM-MSCs

Donor ID	MALDI	ESI	Total
PCBM1641	2120	1186	2768
167696	5181	1220	5667
PCBM1632	3560	838	3960
110877	3370	1758	4404
8F3650	3921	1587	4716
PCBM1662	2861	2248	3464
Total proteins identified			7462

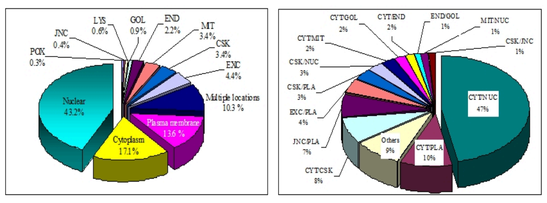


Fig 2: Subcellular classification of proteins identified from six human donors of hBM-MSCs. The cellular location was predicted using ngLOC web-based algorythm. Fig 5A shows proteins predicted to have only one subcellular location (90% of the total proteome) while 5B depicts those expected to appear at more than one location (10% of the total proteome). The notation are POX, peroxisome, JUC, junction, LYS, lysosome, GOL, Golgi, END, endoplasmic, MT, mitochondria, CSK, cytoskeleton, EXC, secreted.

- Proteins harvested from six hBM-MSCs (summarized in Table 1), eithe fractionated using GELFREE system and digested or digested without fractionation, were analyzed using ESI-MS/MS, ESI-MS^E, and MALDI-MS/MS.
- Proteins confidently identified after various statistical testing were summarized in Table 2.
- Proteins are said to be confidently identified if they fulfilled a set of criteria including:Mascot score 40, ID probability 95%, overall false discovery rate (FDR) less than 1%
- On average at least 4000 proteins could be identified from each donor
- Proteins identified using ESI MS and MALDI MS were overlapped for comparison and Fig. 1 is the summary of it. As revealed in Fig 1, MALDI MS identified at least 83% of the total proteome indexed from six donors while ESI MS identified only 52%.
- Fig 2 summarizes the subcellular location of proteins from hBM-MSCs

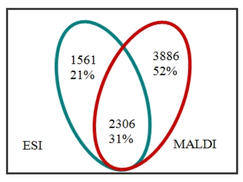
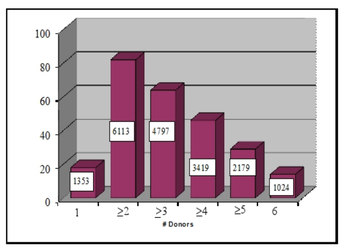


Fig. 1: Summary of proteins identified using electrospray (ESI) and matrix assisted laser desoprtion ionization (MALDI) MS techniques

Comparison between cell lines

Proteomic basis of heterogeneity in hBM-MSCs



- The frequency of protein identification is summarized in Fig 3.
- Overall, a marked proteomic difference between donors is revealed

Fig 3: The frequency distribution of proteins identified from hBM-MSCs from six human donors. Using proteins expressed in 50% donors as

Table 3: The percent variability of protein expression between donors

Donor ID	Total ID	# ID matched with Ref Proteome	% Variability
PCBM1641	2768	2434	28.4
167696	5667	4444	8.7
PCBM1632	3960	3444	23.3
110877	4404	3639	22.2
8F3650	4716	3961	17.1
PCBM1662	3464	3088	25.7

Comparison was made based on a reference proteome constructed using proteins expressed from at least 50% of donors.

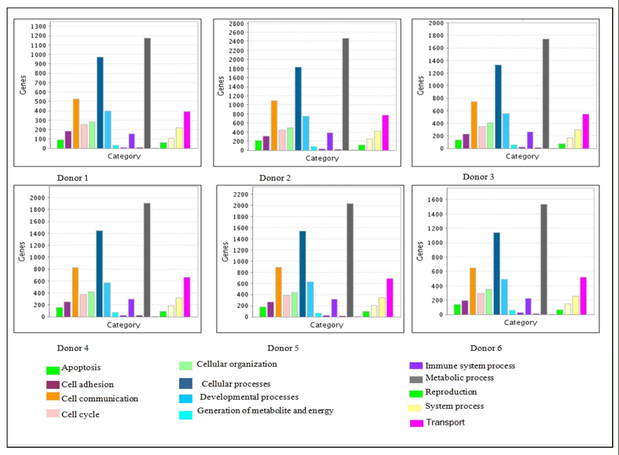


Fig 4: The major biological processes operating in hBM-MSCs. G O t e r m enrichment was performed using the proteins identified from six human donors with the help of the PANTHER program. Fig 5 clearly shows although the proteomes of the six donors are markedly heterogeneous (shown in Fig 3 and Table 3) the variability based on molecular events is not as markedly different as the variability based on protein identity suggests

Results

Quantitative profiling of the expression of proteins across three passages (P3, P5, and P7). Relative quantitation was performed based on the average molar intensities of three most representative peptides of a protein. Known quantity of digested bovine catalase was spiked to protein digests as an internal standard.The proteins whose relative quantities was determined from five hBM-MSCs cell lines are summarized in Table 4

MS-based protein expression profiling

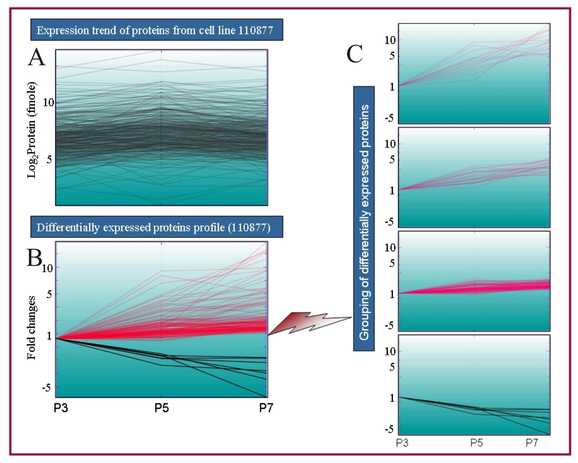


Fig. 5: The expression trend of proteins from cell line 110877 summarized at three passage points (P3, P5, and P7). Fig A shows a plot of the Log2 femtomole amount of proteins when compared across passages. Fig 5B shows differentially expressed proteins while 5C summarizes the classification of differentially expressed trends. Self-organizing maps were generated using Jexpress software..

In total 279 differentially expressed proteins were identified

Validation of differentially expressed proteins using flow cytometry

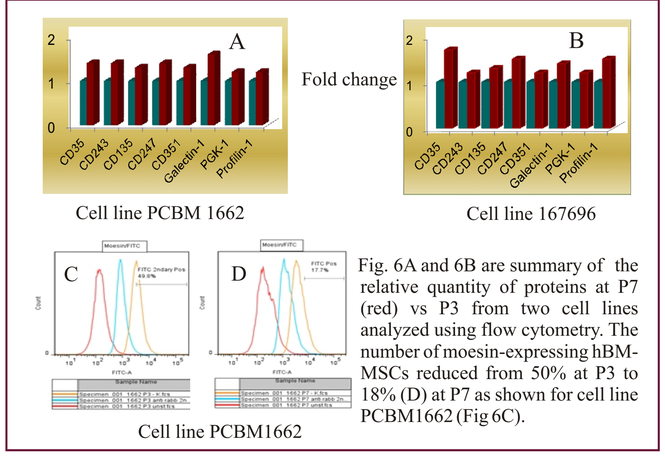


Fig. 6A and 6B are summary of the relative quantity of proteins at P7 (red) vs P3 from two cell lines analyzed using flow cytometry. The number of moesin-expressing hBM-MSCs reduced from 50% at P3 to 18% (D) at P7 as shown for cell line PCBM1662 (Fig 6C).

Table 4: Summary of proteins whose relative quantity was determined and their trend through passaging was profiled. The number of proteins either up- or down-regulated (P3 vs P7) are shown

Cell line	Total proteins	Upreg	Downreg	Total Reg
PCBM1632	765	56	123	179
167696	628	65	45	110
110877	570	122	13	135
8F3650	501	57	13	70
PCBM1662	841	80	21	101

Conclusions

- The present work employed a significantly modified analytical method to sample preparation that involves pressure cycling technology and further protein fractionation using GELFREE protein fractionation system prior to proteomic analysis.
- An overall 3D proteome fractionation with subsequent multiple complementary MS identification techniques, namely ESI-MS/MS, ESI-MS^E, and MALDI-MS/MS, led to comprehensive indexing of proteins expressed by hBM-MSCs.
- The workflow improved the proteomic coverage and 7461 proteins were confidently identified (FDR<1). The majority of proteins are of nuclear origin and some are predicted to appear at more than one subcellular location.
- The proteins are of various molecular classes including 712 transcription and translation regulators, 384 kinases, 248 receptors, and cytokines. Also, the CD marker expression profile of hBM-MSCs have been thoroughly explored with 116 CDs identified and catalogued.
- Comparative screening of the proteomes of cells from six human donors revealed that there is a considerable variability in the proteomic sets between donors. The variability can be estimated to rise above 20% between donors.
- Bioinformatic tools have been used to analyze and identify significantly enriched biomolecular events including signaling pathways and protein-protein interaction networks (interactome) that are likely to operate and influence the biology of hBM-MSCs.
- The Molecular profiles of hBM-MSCs across passages was investigated and the results highlighted the extensive changes in the molecular programming up on passaging.
- The identified proteins could be further investigated as potential markers for hBM-MSCs
- Overall, we present by far the most comprehensive proteomic dataset to date that also contain detailed analyses to help the molecular characterization of undifferentiated and culture-expanded hBM-MSCs. The dataset considerably helps to generate proteomic reference database for these cell types. It also serves as a basis for further exploration of hBM-MSCs with respect to self-renewal, differentiation, and characterization.

References

[1] Charbord P., (2010). Human Gene Therapy, 21:1045-56.

[2] Wagner W., Bork S., Lepperdinger G., Joussen S., Ma N., Strunk D., Koch C., (2010). Aging, 2

[3] Bianco P., Robery PG, Sagio I, Riminucci M., (2010). Hum Gene Therapy, 21: 1057

[4] Roche S., Delrome B., Oostendorp R.A., Barbet R., Caton D., Boumediene K., Papadaki H.A., Cousine B., Crozet C., et al. (2009). Proteomics, 9:223

[5] Niehage C., Steenblock C., Pursche T., Bornhaueser M., Corbeil D., Hoflack B., (2011) . PLOS ONE, 6:e20399

[6] Seshi B., (2007). Proteomics, 7:1984

Sample preparation

HBM-MSCs obtained from six human donors, plated at 60 cells/cm^2, cultured to 80% confluence, harvested at passages 3, 5, and 7

PULSE tube tool

Screw cup

Sample chamber

Barocycler, Pressure Biosciences

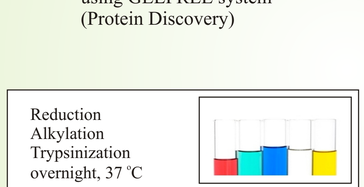
Pressure cycling technology-based protein extraction

Mechanism of enhancement of protein release

1. Thermodynamic: Pressure-induced phase equilibrium shift
   High pressure changes the temperature of membrane phase change (P induced lipid crystallization). Proteins are forced out of lipid crystal. Pressure cycling prohibits phase re-equilibration and improve MP recovery

2. Kinetic: adiabatic cooling and heating

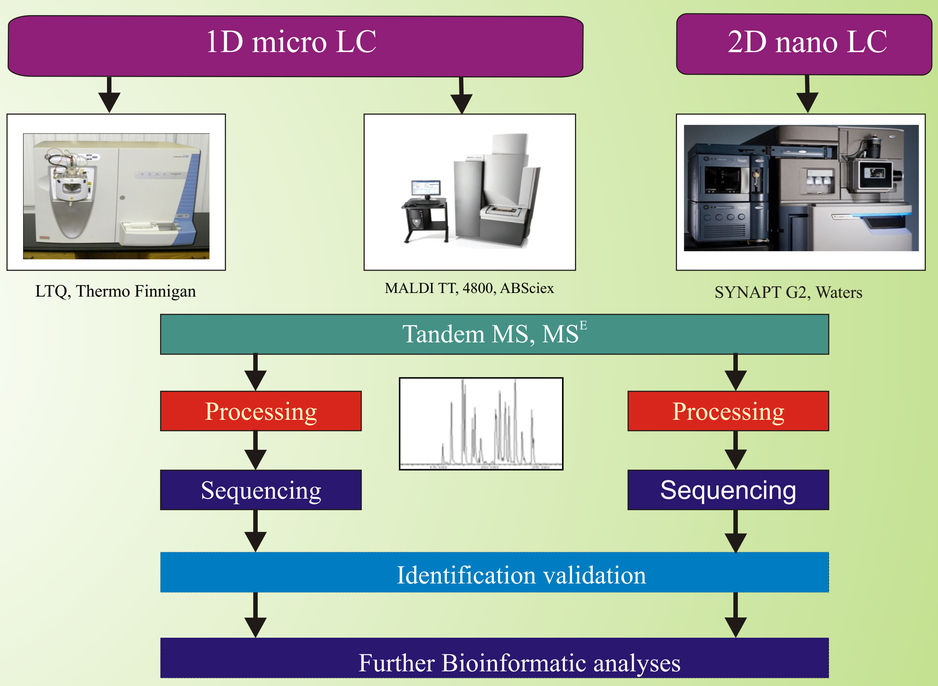
Intact protein fractionation prior to tryptic digestion using GELFREE system (Protein Discovery)



Digest preparation, spiking using known quantity of internal standard

Contact: Samuel.Mindaye@fda.hhs.gov

MS and bioinformatic analyses



1. Gene Ontology, protein loci prediction, network analysis, IPA, ngLOC
2. Protein quantitation, expression trend analysis, self-organizing profile searching, Jexpress
3. Protein functional annotation, PANTHER