A comparative analysis of the MSC transcriptome: Human iPSC-derived MSCs and their tissue-derived counterparts

Margeaux Hodgson-Garms[1,2], James Carthew[1], Mikaël Martino[2], Kilian Kelly[3], Jessica E Frith[1,2]

1. Monash University, Department of Materials Science and Engineering, Melbourne Australia. 2. Australian Regenerative Medicine Institute, Monash University, Melbourne Australia. 3. Cynata Therapeutics, Melbourne Australia.

Margeaux.Hodgson-Garms@monash.edu
MargeauxDjaq
FrithLab.com

Background

- Multipotent mesenchymal stromal cells (MSCs) have considerable therapeutic potential and are one of the most popular and versatile cell therapies[1].
- Traditionally sourced from tissue donations, clinical translation is affected by donor-dependence and significant batch-batch, source-based, and intra-population heterogeneity. This limits the predictability and reproducibility of MSCs in the clinic.
- iPSC-derived MSCs (iMSCs) may bypass many of these problems and can potentially provide a limitless supply of batch-consistent, off-the-shelf cell therapies.
- However, how iMSCs compare to tissue-derived MSCs is not yet fully understood.

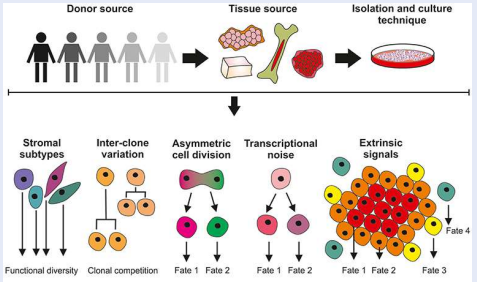

Figure 1. Sources of MSC heterogeneity[1]**Heterogeneity within individual MSC clones**

Aims and project design

Aim: To use single cell sequencing to compare and characterise the MSC transcriptome and quantify inter and intra-population intrapopulation heterogeneity between iPSC and tissue-derived MSCs .

This study used NextGen single cell seq to profile transcriptomes of from **13 MSC populations** including multiple batches of clinical grade and commercially available iMSCs, alongside tissue-derived MSCs from bone marrow, adipose tissue and umbilical cord.

After QC and data processing this data set includes transcriptomes from **72,709 individual MSCs** sequenced at a depth of **>100,000 reads/ cell.**

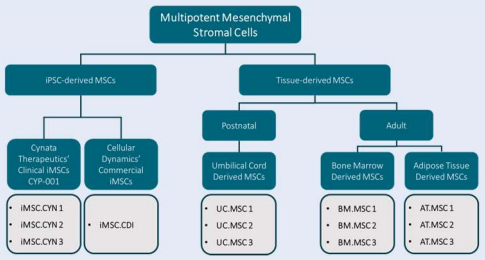

Figure 2. Schematic outline of the source and labelling of MSC populations used.

MSCs cluster primarily by tissue/ source.

UMAP clustering of MSC transcriptomes indicates that tissue/ source of origin accounts for most MSC heterogeneity **(Fig.3A)**. MSC tissue/ sources formed clades within themselves. BM.MSC and AT.MSCs branched latest while iMSC and UC.MSCs branched earlier indicating comparatively less similarity **(Fig.3B).**

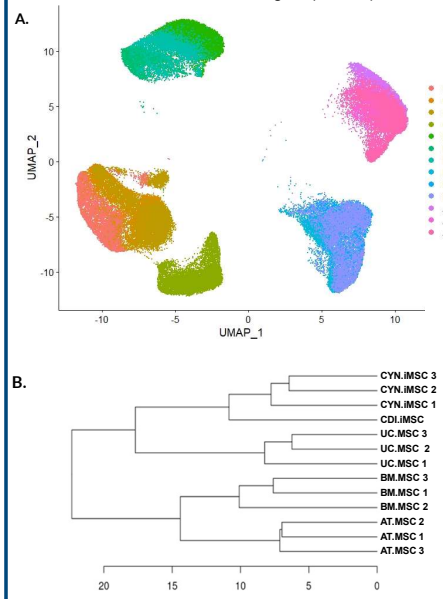

Figure 3. UMAP visualisation of MSC transcriptome clustering (A). Single-cell MSC transcriptomes (N=72,809) from 13 populations (indicated by point colour) visualised using UMAP projection generated in Seurat using function DimPlot. **Dendrogram of MSC population hierarchy (B).** Seurat hclust function was used to generate a hierarchical population (n=13) dendrogram based on single-cell distance-from centroid-values. Branching depth (X) indicates population similarity.

Differentially expressed (DE) genes were identified between iMSC and tissue-derived MSCs.

820 genes were upregulated in tissue-derived MSCs (tMSCs) while 5491 genes were upregulated in iMSCs **(Fig. 4A)**. Gene Ontology (GO) term enrichment analysis was used to query DE genes for enriched Biological Processes (BP). BP including telomere maintenance and RNA catabolism processes were enriched in genes upregulated in iMSCs, while genes downregulated in iMSCs were enriched for humoral immune response and complement processes **(Fig. 4B)**.

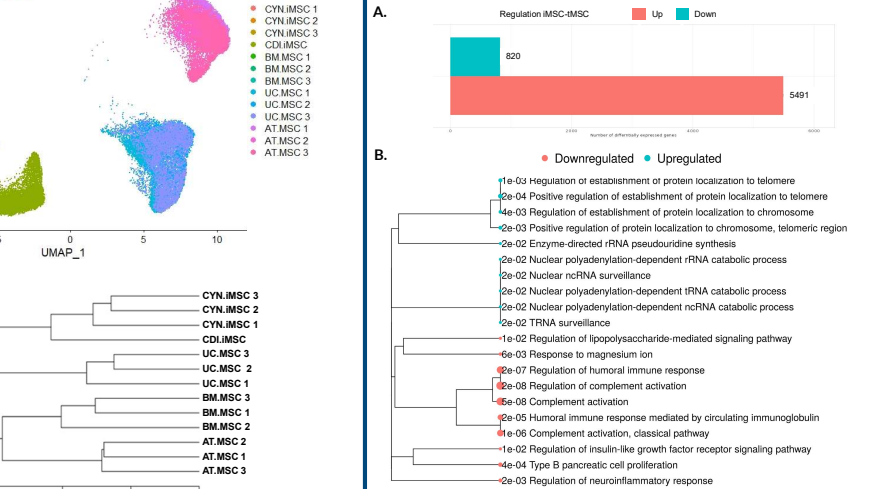

Figure 4. Differential expression iMSC vs tMSCs (A). DeSeq2 was used to identify genes significantly up (N=5491) or downregulated (N-820) between iMSC and tMSCs. **Top biological processes enriched in DE genes (B).** GO term enrichment analysis was used to identify the top 10 most strongly enriched BP both upregulated and downregulated in iMSCs. GO term tree was generated based on shared gene membership. Point size is representative of enrichment score and point colour indicates if gene members are up or down regulated. Plots were generated in iDEP 1.0 [4]

Intrapopulation variance was quantified as a factor of cell-cell gene variance within the top 200 most variable genes.

Mean cell-cell transcriptomic variance was observed to be significantly lower in iMSCs than tMSCs. Furthermore, mean cell-cell variance was comparable between iMSC populations while tMSC populations showed significant donor-donor differences.

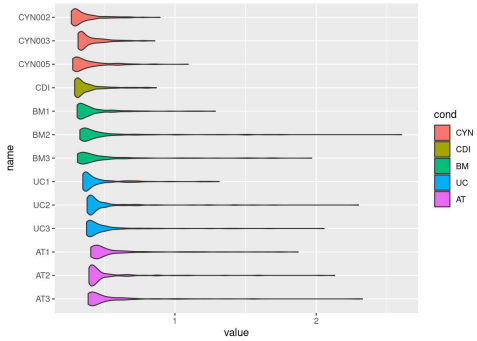

Figure 6. Violin plots of cell-cell variance across top 200 most variable genes. The top 200 most variable genes were identified using DeSeq2. Gene-wise variance from the median was calculated for each cell and single-cell. Variance scores (x) are presented as a violin plot. Tissue/ source is indicated by colour. Plot was produced using Seurat.

Principle Component Analysis (PCA) and PCA loading identify gene markers driving separation of iMSCs.

PCA of MSC transcriptomes identified two components accounting for 40.47% and 20.16% of variation between MSC tissue/sources. PCA loading identified major genes driving component separation **(Fig 5A)**. Expression of CRMP1, DIPK1B and LIN28B was found to drive separation of iMSCs from tMSCs, with highest expression in iMSCs followed by BM.MSCs, UC.MSC and AT.MSCs **(Fig 5B).**

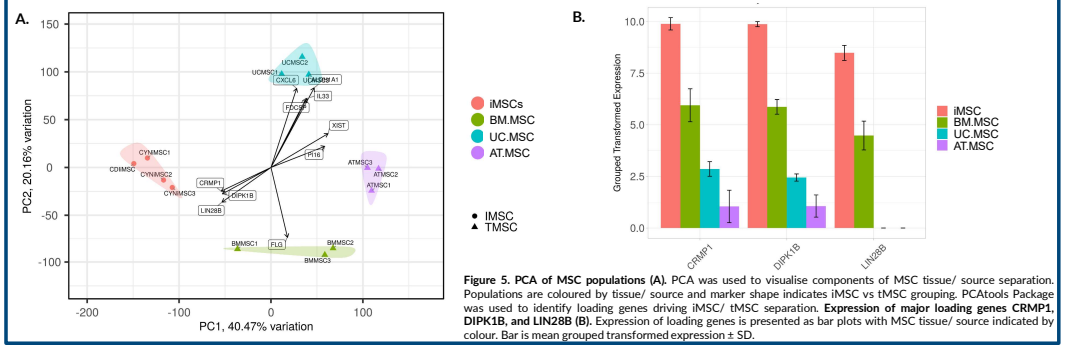

Figure 5. PCA of MSC populations (A). PCA was used to visualise components of MSC tissue/ source separation. Populations are coloured by tissue/ source and marker shape indicates iMSC vs tMSC grouping. PCAtools Package was used to identify loading genes driving iMSC/ tMSC separation. **Expression of major loading genes CRMP1, DIPK1B, and LIN28B (B).** Expression of loading genes is presented as bar plots with MSC tissue/ source indicated by colour. Bar is mean grouped transformed expression ± SD.

Conclusions

Key Findings:

1) Tissue/ source is the primary driver of MSC heterogeneity.
2) iMSCs are most closely related to UC.MSCs, while BM.MSCs and AT.MSCs are more closely related to each other.
3) iMSCs differ from tissue-derived MSCs by the upregulation of biological processes linked to telomere maintenance and RNA catabolism, and the downregulation of humoral immune response and complement processes.
4) iMSCs exhibit less batch-batch heterogeneity than tissue-derived MSCs, furthermore they also exhibit significantly less intra-population variation.

This data set provides a comprehensive profile of MSC transcriptomes at a single-cell level, allowing us to develop a better understanding of the sources of MSC heterogeneity and improve predictability of clinical outcomes. Moreover, this study confirms that iMSCs successfully bypass much of the inherent heterogeneity that affects the clinical application of tissue-derived MSCs, validating their promise as an off-the-shelf cell therapy.

References and Acknowledgments

1. ClinicalTrials.gov. Search of: Mesenchymal Stem Cell - List Results - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/results?cond=Stem+Mesenchymal+Stem+Cell&cntry=&state=&city=&dist=.
2. Wilson, A., Hodgson-Garms, M., Frith, J. E. & Genever, P. Multiplicity of mesenchymal stromal cells: Finding the right route to therapy. *Front. Immunol.* **10**, (2019).
3. Dominici, M. et al. 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, *8*(4), pp.315-317.
4. Ge, S. X., Son, E. W. & Yao, R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinforma. 2018 191* **19**, 1–24 (2018).

This project was supported by:
Australian Govt. RTP Stipend
Monash University Dpt. of Materials Science and Engineering
Monash University Graduate Research Completion Award