

Tesis Doctoral



VNIVERSIDAD
D SALAMANCA

**Caracterización transcriptómica de Células
Estromales/Madre Mesenquimales humanas**

Beatriz Rosón Burgo

Directores

Dr. Javier De Las Rivas

Dr. Fermín Sanchez-Guijo

Dra. M^a Consuelo Del Cañizo



**CENTRO DE INVESTIGACIÓN
DEL CÁNCER**

Salamanca, 2015

Dr. Javier De Las Rivas Sanz, con D.N.I. 15949000H, Investigador Científico del Consejo Superior de Investigaciones Científicas (CSIC), director del grupo de Bioinformática y Genómica Funcional en el Instituto de Biología Molecular y Celular del Cáncer (CiC-IBMCC), y profesor del Programa de Doctorado y del Máster de Biología y Clínica del Cáncer de dicho Instituto y la Universidad de Salamanca (USAL).

Dr. Fermín Sánchez-Guijo Martín, con D.N.I. 07869479Y, Profesor Asociado de Ciencias de la Salud de la Universidad de Salamanca (USAL) y Responsable del Área de Terapia Celular y Hematopoyesis del Servicio de Hematología del Hospital Universitario de Salamanca (HUS).

Y Dra. María Consuelo Del Cañizo Fernández-Roldán, con D.N.I. 03404639H, Profesora Titular de Hematología y Hemoterapia de la Universidad de Salamanca (USAL) y Jefa de Servicio de Hematología y Hemoterapia del Hospital Universitario de Salamanca (HUS).

CERTIFICAN

que han dirigido esta Tesis Doctoral titulada "*Transcriptomic characterization of human Mesenchymal Stromal/Stem Cells*" realizada por Dña. Beatriz Rosón Burgo, que obtuvo su Diploma de Estudios Avanzados en la Universidad de Salamanca (USAL), a través del programa de doctorado *Biología y Clínica del Cáncer* del Instituto de Biología Molecular y Celular del Cáncer (CiC-IBMCC).

Y AUTORIZAN

la presentación de la misma, considerando que reúne las condiciones de originalidad y contenidos requeridos para optar al grado de Doctor por la Universidad de Salamanca.

En Salamanca, a 5 de Noviembre de 2015

Dr. J. De Las Rivas

Dr. F. Sánchez-Guijo

Dra. M. C. Del Cañizo

Para la realización de esta Tesis Doctoral, la doctoranda Beatriz Rosón Burgo obtuvo en concurso público una de las *Ayudas para la Contratación de Personal Investigador* de reciente titulación universitaria, convocadas por la Junta de Castilla y León (JCyL) dentro de la *Estrategia Regional de Investigación Científica, Desarrollo Tecnológico e Innovación 2007-2013*, cofinanciadas por el Fondo Social Europeo (FSE). Posteriormente, pudo continuar su investigación gracias a la contratación a través del Instituto de Investigación Biomédica de Salamanca (IBSAL), y a través de un Proyecto de Investigación con *Celgene Institute for Translational Research Europe* (CITRE) asociado a la Fundación de Investigación del Cáncer (FICUS).

La investigación de esta Tesis Doctoral ha sido desarrollada gracias a la financiación de los siguientes Proyectos concedidos al grupo de investigación del Dr. Javier De Las Rivas en el CiC-IBMCC: (i) "*Biología molecular integrativa de hemopatías malignas: análisis bioinformáticos de datos transcriptómicos y proteómicos para identificar genes marcadores, genes causales y redes reguladoras asociadas a subclases patológicas de dos síndromes proliferativos específicos*", Ministerio de Economía y Competitividad, ISCiii (AES 2012, PI12/00624). (ii) "*Data Representations and Similarity Measures for High-Throughput Clinical Sample Profile*", CITRE-FICUS (2013-2015). (iii) "*Transcriptómica y epigenómica de Células Stem Mesenquimales (MSC) normales y alteradas en hemopatías malignas*", Junta de Castilla y León, Consejería de Sanidad, Proyectos en Biomedicina (BIO/SA68/13). (iv) "*Análisis genómico integrativo y búsqueda de marcadores específicos de Células Stem Mesenquimales (MSC) normales y alteradas en hemopatías malignas*", Junta de Castilla y León, Consejería de Sanidad, Proyectos en Biomedicina (BIO/SA08/14).

Una parte del trabajo de esta Tesis Doctoral fue realizado en el laboratorio del Dr. Paul Bertone en el *European Bioinformatics Institute* (EMBL-EBI) en Cambridge (Reino Unido) gracias a la concesión en 2011 de una *Beca para Estancias de Investigación* de 3 meses de la Junta de Castilla y León (JCyL).

El proceso formativo-investigador que supone una tesis doctoral conlleva la escritura y publicación de resultados científicos en revistas de impacto internacional. Razón por la que se ha optado por el inglés como idioma de escritura de esta tesis doctoral.

Introducción

Células Estromales/Madre Mesenquimales (MSCs)

Células madre adultas: multipotencia y auto-renovación

Las células madre encuentran el sentido de su existencia en los organismos multicelulares. Reciben su nombre por sus capacidades para dividirse en entidades iguales, además de producir otros tipos de células más especializados a través de los procesos de compromiso de linaje y diferenciación celular. Desde los primeros estudios de Becker, Till y McCulloch en la década de 1960, estas características han pasado a ser ampliamente consideradas como auto-renovación y multipotencia de diferenciación [Becker *et al.* 1963, Reya *et al.* 2001].

Se consideran dos tipos principales de células madre a partir de tejidos de mamíferos: las células madre embrionarias, a partir de la masa celular interna del blastocisto; y las células madre adultas, a partir de los tejidos de la vida adulta. Lo que los distingue es la potencia diferenciación. Las embrionarias dan lugar a todos los tipos celulares de un organismo (capacidad llamada pluripotencia), así como las adultas generan varios tipos de células diferenciadas restringidas a un linaje de tejido, llamada multipotencia. El embrión en desarrollo queda tempranamente definido en las tres capas principales: ectodermo, endodermo y mesodermo. Cada capa representa un linaje celular. El linaje mesodérmico se desarrollará en los tejidos óseos y conjuntivos del cuerpo adulto. Las células madre multipotentes en el organismo adulto se almacenan en estado indiferenciado y serán utilizadas en la recuperación y regeneración del tejido normal dañado cuando sea necesario, representan un fondo de reposición para afrontar el daño o enfermedad tisular (figura 1.1). Consecuencia de la capacidad regenerativa es el uso frecuente de las mismas como terapia celular clínica.

Células Estromales/Madre Mesenquimales: prueba de concepto

Las Células Estromales/Madre Mesenquimales (MSCs, del inglés *Mesenchymal Stromal/Stem Cells*) representan una población celular heterogénea con capacidades multipotentes y de auto-renovación, y que se encuentran en la parte estromal de los tejidos adultos. Las MSCs son reconocidas por su crecimiento en cultivos adherentes al plástico, con morfología similar al fibroblasto, capacidad de diferenciación *in vitro* hacia tres linajes mesenquimales (osteogénico, adipogénico y condrogénico) y un fenotipo delimitado por moléculas marcadoras CD positivas y negativas. Sin embargo, el concepto ha sufrido muchas variaciones desde su primer descubrimiento hace 50 años por [Friedenstein *et al.* 1966]. La figura 1.2 muestra una línea temporal de los principales descubrimientos conceptuales del origen de las MSCs.

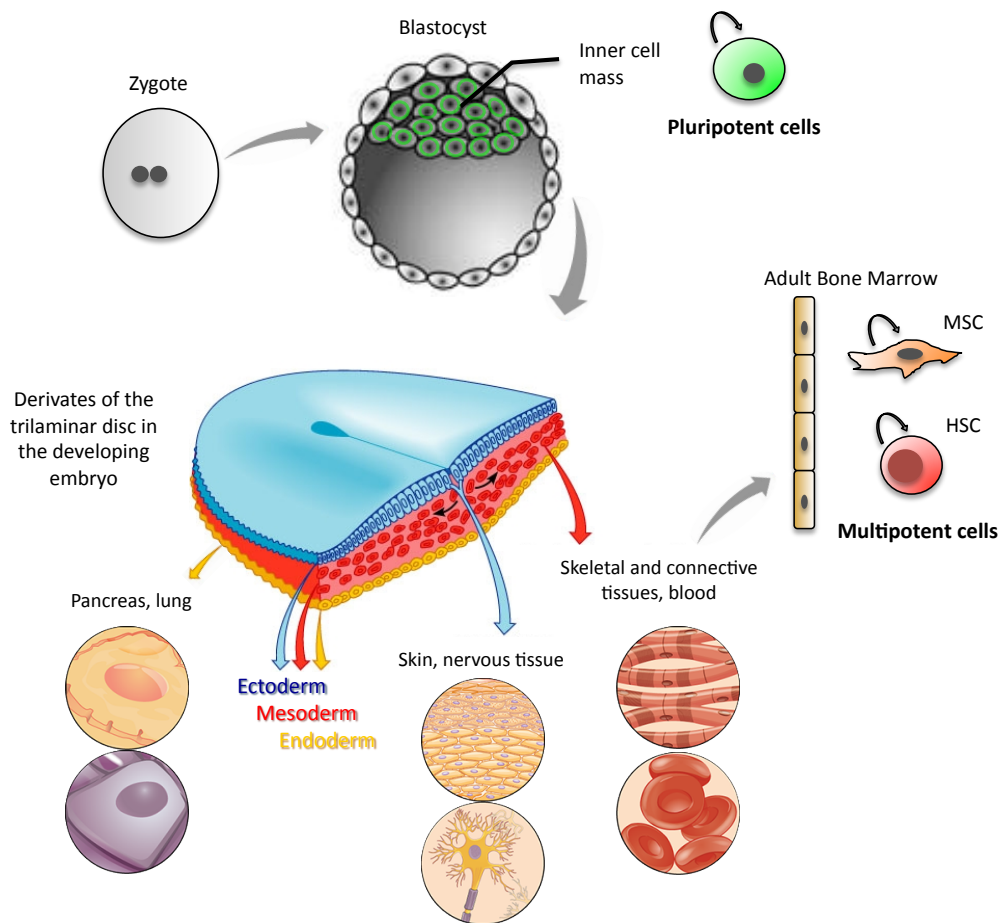


Figure 1.1 Deriva de células madre en el embrión (como células madre pluripotentes) a tejidos adultos (como células madre multipotentes).

Friedenstein y colaboradores demostraron a través del trasplante heterotópico la presencia de una población precursora no hematopoyética pero sí osteogénica en el estroma de la médula ósea (BM, del inglés *Bone Marrow*). Trascendentalmente, podrían asignar la capacidad autorrenovadora gracias al ensayo de establecimiento de una célula iniciadora de una única colonia fibroblástica, conocida como unidad formadora de colonias de fibroblastos (CFU-Fs, *Colony Forming Unit of Fibroblasts*). Otros experimentos *in vivo* produjeron la generación de múltiples tejidos esqueléticos (hueso, cartilago, tejido adiposo, tejido fibroso) y llevaron a asignar el concepto “célula madre” a estos habitantes del estroma medular. Caplan y colaboradores [Caplan et al. 1991], justo precediendo al estallido de las células madre embrionarias y en vista de los tipos mesodérmicos a los que dan lugar más allá de los tejidos óseos, acuñó el término de “células madre mesenquimales”, el cual fue posteriormente popularizado. Estudios ulteriores

contribuyeron a reforzar esta naturaleza *stem* [Prockop et al. 1997, Pittenger et al. 1999], aunque sólo se ha alcanzado veracidad completa sobre multipotencia para aquellas células del estroma de tejidos óseos que forman huesecillos heterotópicos. La figura 1.2 presenta los cambios conceptuales más relevantes sobre MSCs a lo largo de la historia.

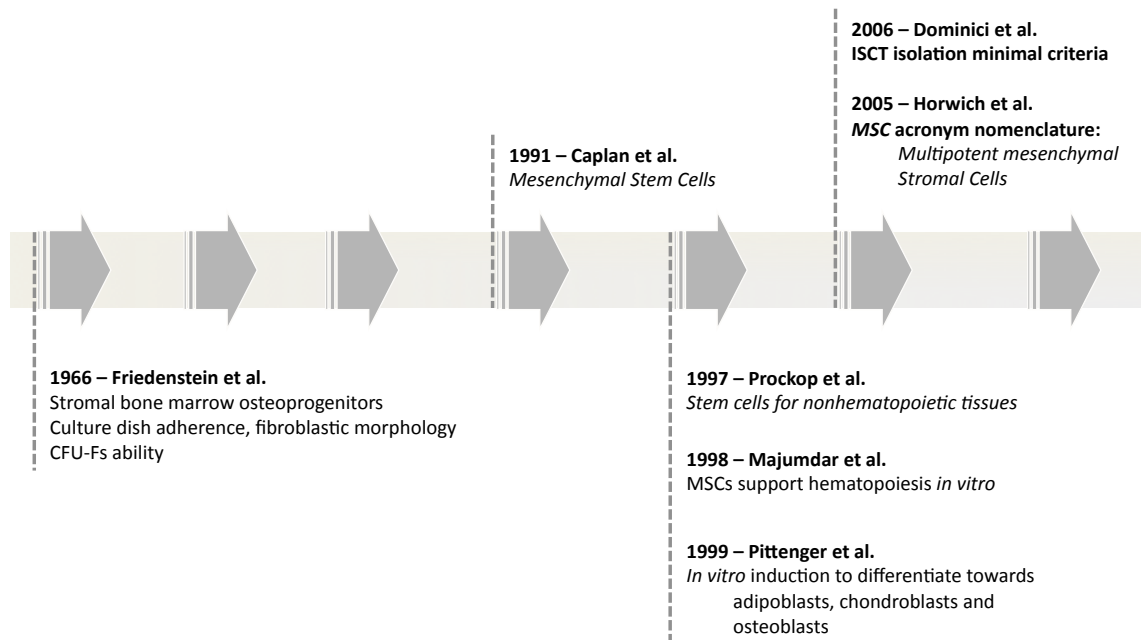


Figure 1.2 Línea histórica de los descubrimientos conceptuales sobre las MSCs.

El papel del estroma como parte fundamental en la hematopoyesis también se evidenció en co-cultivos a largo plazo de MSCs con progenitores CD34+, las MSCs mantenían la diferenciación progresiva de las células sanguíneas [Majumdar et al. 1998]. Este último hallazgo redirigió la atención mundial hacia su uso aplicativo en trasplantes de médula ósea, y generó la explosión de nuevos protocolos y fuentes para el aislamiento de MSCs. Sin embargo tal crecimiento se produjo sin un criterio común de identificación, haciendo imposible la comparación entre estudios de poblaciones independientes de MSCs. Para hacer frente al desorden general, la Sociedad Internacional de Terapia Celular (ISCT, <http://www.Celltherapysociety.org/>) propuso unas normas de nomenclatura y de protocolo de aislamiento para la identificación común de las MSCs [Horwitz et al. 2005, Dominici et al. 2006]. Estas poblaciones heterogéneas de MSCs así aisladas, que contienen un subconjunto de células madre mesenquimales (el que realmente cumpliría las cualidades de célula madre), deben ser mejor denominadas como *células estromales mesenquimales pluripotentes*. El consenso acordó tres criterios mínimos de identificación:

- Crecimiento selectivo en adherencia a una superficie de plástico bajo condiciones de cultivo estándar (descrito al detalle en la metodología de ésta tesis).
- Diferenciación *in vitro* a, al menos, tres de los linajes establecidos: adipoblasto, condroblasto y osteoblasto.
- Un inmunofenotipo positivo para CD105, CD73 y CD90. Este fenotipo debe ir acompañado de la falta de marcadores específicos de linaje hematopoyético, como son: CD45, CD34, CD14 or CD11b, CD19 or CD79 α , y HLA-DR.

Estos criterios siguen vigentes hoy en día, y también, muchas de las cuestiones de identidad subyacentes al tipo MSC. La mayor parte del material publicado sobre los beneficios clínicos de las MSC se ha realizado con poblaciones bajo los criterios de la ISCT. A lo largo de esta tesis, nos hemos centrado en la población así definida, y en todo momento me referiré a las MSC bajo estos criterios ISCT. Sea el lector siempre consciente de la naturaleza heterogénea de la población, idea que mejor corresponde a la complejidad biológica del sistema de estudio.

Potencial de diferenciación

La denominación “mesenquimal” de Caplan se basó inicialmente en el modelo de diferenciación hacia los tejidos considerados derivados del mesoderma embrionario (figura 1.3 muestra el esquema original de Caplan de lo que él llama “El proceso mesengénico”, reelustrado por Megan Kern en 2012: http://megankerndesign.com/artwork/1100817_Mesengenic_Process.html)

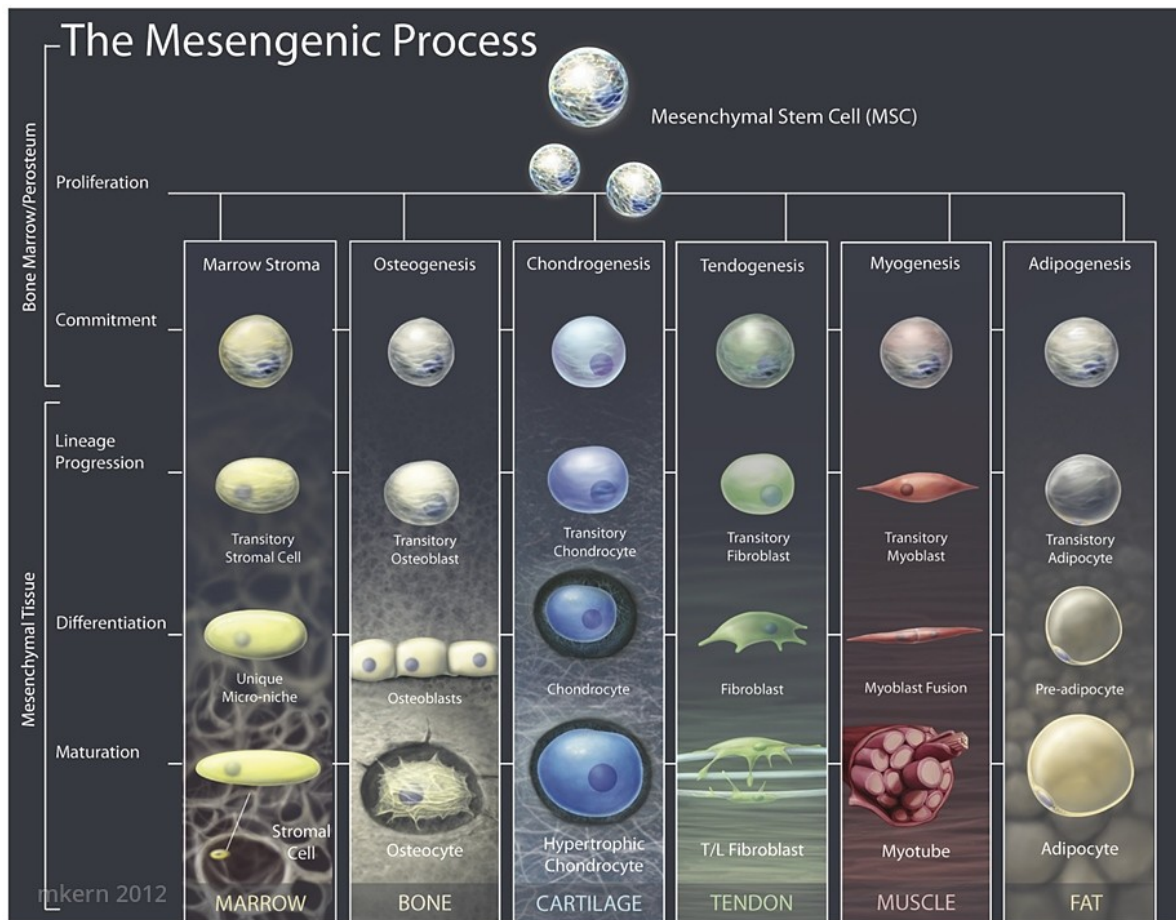


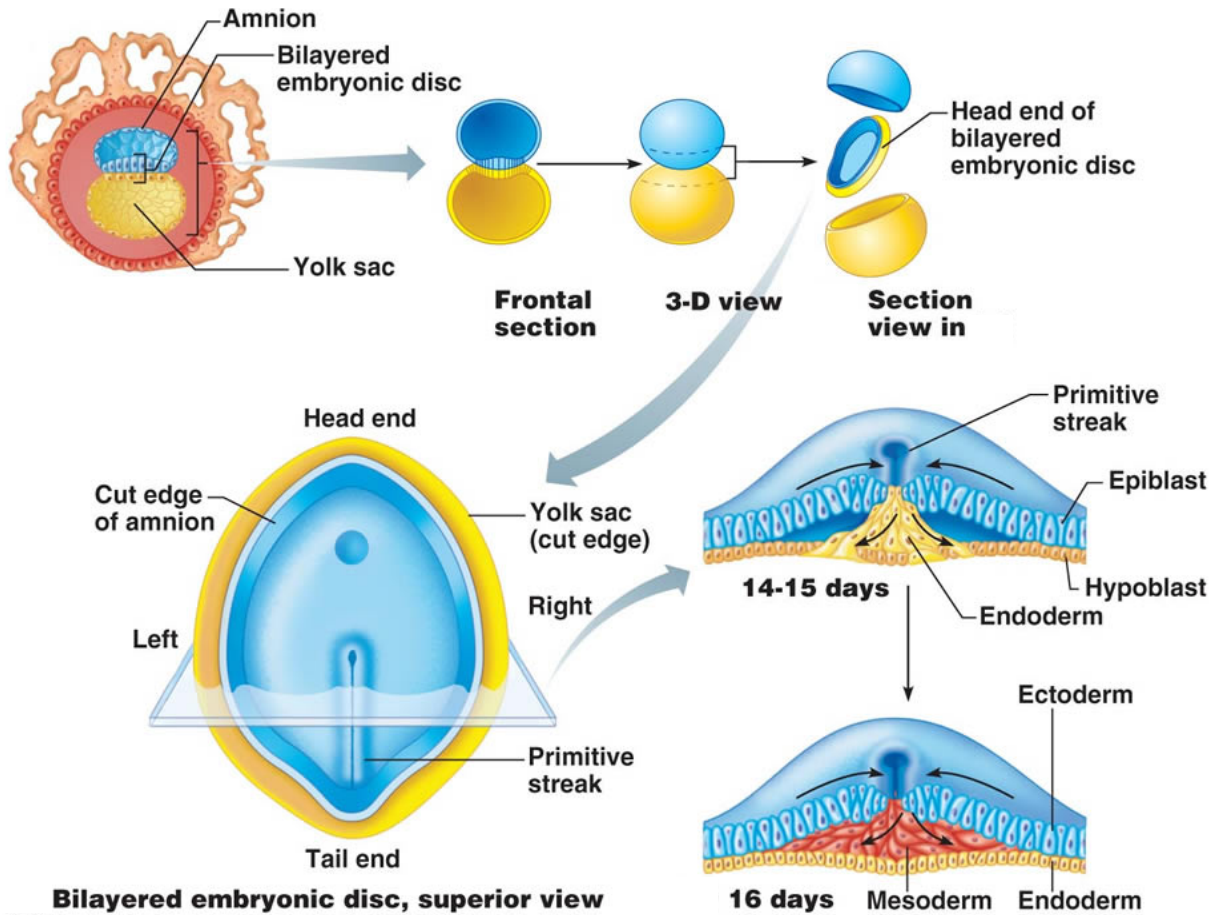
Figure 1.3 El Proceso Mesengénico de Caplan.

Aparte de los ya citados tejidos esqueléticos (osteoblastos, condrocitos, adipocitos, fibroblastos y células reticulares o adventicias perivasculares), otros tejidos mesodérmicos postnatales incluirían: la dermis, músculos (músculos liso, esquelético y cardíaco), tendones, ligamentos... y, potencialmente, todas las células soporte que componen el estroma de diferentes órganos. En la literatura podemos encontrar muchos ensayos *in vitro* para inducir la diferenciación de MSCs hacia los diversos tipos celulares mesodérmicos, a parte de los ya citados podemos incluir: miocitos [Zuk et al. 2001, Dezawa et al. 2005], cardiomiocitos [Makino et al. 1999, Choi et al. 2011]; e incluso de inducción no mesodérmica hacia: hepatocitos [Lee et al. 2004], o linajes neurales [Dezawa et al. 2004]. Sin embargo, la realidad es que las evidencias *in vivo* de la capacidad multipotente de diferenciación son aún inexistentes y el asunto sigue siendo bastante controvertido. Los ensayos de diferenciación *in vitro* hacia los tres linajes esqueléticos estándar (osteoblastos, condroblastos y adipoblastos) son la forma normalizada básica para probar la identidad multipotente de las MSC.

Ontogénesis

Entre el tercer y cuarto día después de la fecundación, la mórula crece, se cavita y se convierte en blastocisto (figura 1.4). La masa celular interna del blastocisto se convertirá en el embrión en sí, situándose en uno de los polos del blastocisto, mientras que en la capa exterior se ha formado el trofoblasto que desarrollará las membranas extraembrionarias. En la tercera semana comienza el proceso de gastrulación con la aparición de la línea primitiva y un nódulo en su cabecera. El disco bicapa embrionario formado por el epiblasto y el hipoblasto, se convertirá en tricapa. Las células del epiblasto migran a través de la línea primitiva, deslizándose hacia abajo a través de un proceso principal de invaginación. Parte de ellas se mueven hacia el hipoblasto, integrándose en lo que será el endodermo definitivo. Las otras se establecen en el espacio intermedio, fundando la capa del mesodermo. Las células que se mantienen en el epiblasto constituirán el ectodermo. Estas tres capas germinales dan lugar al embrión completo. El mesodermo evoluciona hacia cuatro estructuras principales: notocorda, somitas, mesodermo intermedio y mesodermo lateral. La figura 1.5 resume las partes del cuerpo desarrolladas por cada estructura mesodérmica. Sin embargo, durante la organogénesis prenatal en los mamíferos, la serie de tejidos considerados como mesodérmicos podrían ser originados por progenitores distintos de capas embrionarias distintas. Véase el caso de las estructuras óseas, con los huesos craneofaciales generados por el neuro-ectodermo, mientras que los huesos axiales y apendiculares son respectivamente derivados del mesodermo somítico y lateral. En este sentido, se ha sugerido un origen múltiple de las MSC esqueléticas por la diferente derivación embrionaria de las distintas estructuras óseas. De este modo, algunos autores han propuesto un desarrollo embrionario paralelo y coordinado de las MSCs con células progenitoras hematopoyéticas [Mendes et al. 2005]. Podrían existir orígenes mesenquimales en sitios anatómicos distintos durante la ontogenia, asociados a distintos microambientes hematopoyéticos tempranos: la región aorta-gónada-mesonefros (AGM), el hígado fetal y, finalmente, la médula ósea. En su trabajo, se detectaron las MSC en embriones de ratón de 11 días en la región AGM, en el momento en el que emerge la hematopoyesis; y aún más, las MSC parecen presentar preferencias de diferenciación asociadas a la ubicación dentro de la región AGM. Estudios de vanguardia que revelan un origen de la cresta neural para algunas MSC de ratón especialmente vinculadas a nichos hematopoyéticos, también conducen a la idea de que existen poblaciones de MSC fenotípicamente diferentes en la médula de los huesos largos que también pueden presentar destinos biológicos distintos [Isern et al. 2014].

Figure 1.4 Procesos de blastulación y gastrulación durante la embriogénesis para formar las tres capas embrionarias.



Modified from Pearson Education 2013

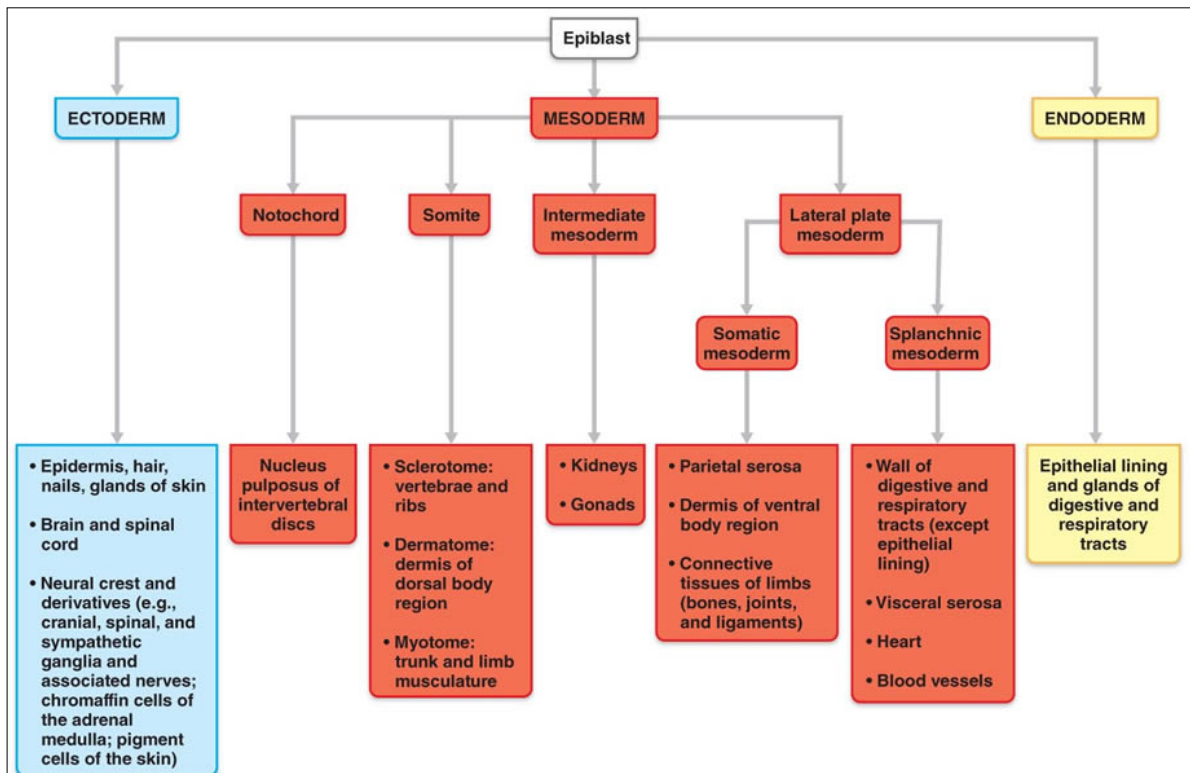


Figure 1.5 Derivados mesodérmicos en mamíferos.

Inmunofenotipo

Con la esperanza de encontrar un biomarcador específico que permita el aislamiento directo de una población más homogénea de MSCs y evitar los pasos de selección en cultivos adherentes, varias moléculas de superficie celular se han propuesto como antígenos para su reconocimiento a través de anticuerpos monoclonales (AcMo). En 1991, Simmons y Torok-Storb, [Simmons et al. 1991] desarrollaron el primer hibridoma murino productor de un AcMo que reconocía células precursoras estromales, no hematopoyéticas, de BM humanas. Dicho AcMo fue llamado Stro-1. Hoy en día todavía es el producto de un gen desconocido, de 75 kDa, y que no resulta específico a MSCs puesto que también reconoce células endoteliales [Ning et al. 2011].

En 1992, Haynesworth et al. [Haynesworth et al. 1992] produjeron tres líneas más específicas de MoAb contra BM-MSCs humanas, que no reaccionaban con células hematopoyéticas u osteocitos: SH-2, SH-3 y SH-4. SH-2 reconoce un epítipo del complejo receptor del TGF- β conocido como Endogлина (o CD105). SH-3 y SH-4 reconocen cada una un epítipo de la 5'-Nucleotidasa (o CD73). En 1998, Bruder et al. utilizaron el anticuerpo SB-10 para caracterizar las poblaciones de BM-MSCs indiferenciadas. El antígeno SB-10 se identificó como CD166, que desempeña un papel en la vía de diferenciación osteogénica [Bruder et al. 1998]. Más tarde, Boiret et al. [Boiret et al. 2005] utilizaron con éxito un marcador de superficie ya conocido de las poblaciones hematopoyéticas del timo (la glicoproteína Thy-1, o CD90) para refinar la caracterización de dicha población mesenquimal.

En el consenso de 2006, la ISCT dictó que por citometría de flujo debía medirse al menos un 95% de la población celular de MSCs con expresión para CD73 (ecto-5'-nucleotidasa), CD90 (Thy-1) y CD105 (endogлина), y no más de un 2% con expresión para CD34 (marcador de progenitores hematopoyético y células endoteliales), CD45 (marcador pan-leucocitario), CD11b o CD14 (marcadores de la línea monocito-macrófago), CD19 o CD79 α (marcadores de células B), y HLA-DR (marcador de células immunoestimuladas). Además de la lista de marcadores normalizados por el ISCT, otros antígenos de superficie se han descrito para definir las poblaciones de MSCs. En la década del 2000 varios experimentos de citometría caracterizaron el CD271 (también denominado receptor con baja afinidad del factor de crecimiento nervioso, LNGFR) como un marcador de clasificación para enriquecer en MSCs muestras crudas de la médula ósea [Quirici et al. 2002, Jones et al. 2002, Jones et al. 2006]. La población extraída con un enriquecimiento al 95% en CD271 resultó también expresar la molécula de adhesión celular vascular CD106 (VCAM-1), un marcador de las células del estroma reticular adyacentes al endotelio de los vasos. Más recientemente, la molécula de adhesión de células de melanoma CD146 (MCAM) fue identificada en MSCs de las zonas sub-endoteliales. Estas MSCs han sido

asemejadas a un tipo celular, también habitante medular, conocido como pericito, el cual comparten muchas propiedades con las MSCs [Sacchetti *et al.* 2007, Tormin *et al.* 2011]. Desafortunadamente, los antígenos identificados hasta la fecha se expresan también en otros linajes celulares y ninguno por se ha resultado lo suficientemente específico para permitir la separación de una población más pura de MSCs. Esta es la razón fundamental por la que se emplea un compendio de varios marcadores positivos y negativos para la identificación fenotípica básica de las MSCs.

Distribución corporal

Como se apuntado anteriormente, la distribución de las MSC a lo largo del organismo post-natal se ha correlacionado con ubicaciones perivasculares [Crisan *et al.* 2008]. En apariencia, las células mesenquimales pueden ser observadas en los espacios circundantes de casi todos los vasos sanguíneos en el cuerpo (arteriales y venosos). Aún más, todas las células del estroma del tejido conectivo comparten expresión de moléculas de adhesión que median la interacción con las células endoteliales a través de la interacción homofílica de MCAM (CD146) con receptores de superficie o antígenos implicados en la regulación de las células perivasculares: PDGFR- β (CD140b) o CD105 (Endogлина). Bajo estos términos, se pensó que las MSCs podían residir en "prácticamente todos" los órganos del cuerpo. Además de la médula ósea, otros órganos propuestos son: músculo, pulmón, riñón, hígado, páncreas, bazo, timo, cerebro, placenta o del cordón umbilical [Erices *et al.* 2000, Da Silva Meirelles *et al.* 2006, Parolini *et al.* 2008]. La mayor disponibilidad y facilidad de acceso de tejidos como los anexos fetales, impulsaron fuertemente la búsqueda de la aplicabilidad de sus derivados mesenquimales a modo de células madre en terapias celulares.

A pesar del cumplimiento general del fenotipo de caracterización propuesto por la ISCT, las MSC de diferentes tejidos muestran diferencias variables en su morfología, su cinética de crecimiento, sus niveles de expresión de moléculas marcadoras, y su capacidad de diferenciación. De cualquier modo, hasta la fecha no existe aún un marcador que resulte específico de alguno de los citados tejidos.

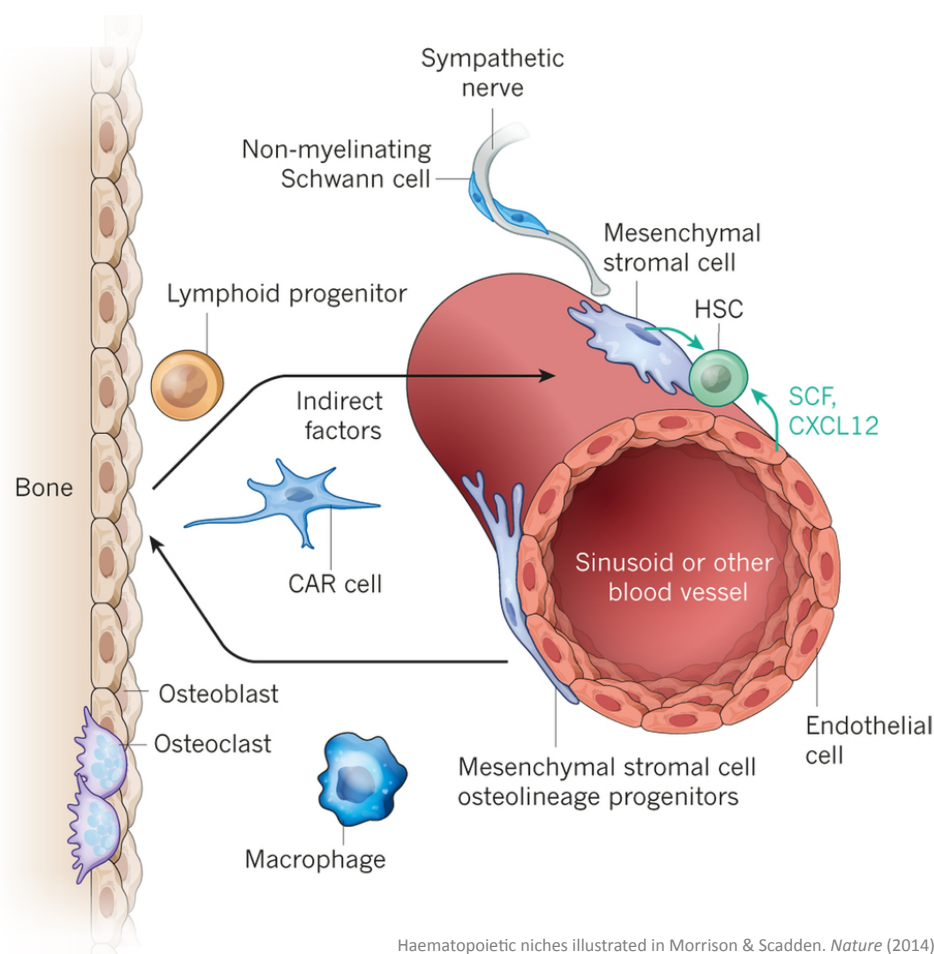
Además de la aparente presencia ubicua de poblaciones MSC, otra célula derivada del linaje mesenquimal, ampliamente conocida en la experimentación biológica, y presente en la mayoría de los tejidos conectivos sólidos, puede ser confundida con las MSCs: el fibroblasto. Las características que definen morfológicamente a las MSCs proceden en origen de los fibroblastos de la piel. Estos expresan todos los marcadores MSC aprobados por la ISCT, y muestran un

porcentaje variable de la capacidad de multipotencialidad e incluso inmunoreguladora [Haniffa *et al.* 2007, Haniffa *et al.* 2009, Halfon *et al.* 2011, Vaculik *et al.* 2012]. La frontera entre los fibroblastos y las MSC ha sido siempre más bien oscura y sigue siendo muy controvertida. Es interesante el hecho de que los fibroblastos presentan "una memoria de posición", es decir, presentan diferentes perfiles de transcripción y de marcadores de superficie en función de la parte del cuerpo de la que proceden [Chang *et al.* 2002]. Esto es una reminiscencia de los perfiles específicos de tejido observados para las MSC residentes en tejidos anatómicamente diferentes [Wagner *et al.* 2005]. Aunque algunos estudios han establecido las bases de este marco conceptual, la realidad es que la caracterización de perfiles relacionados con el tejido de las MSC sigue siendo una tarea permanente. Por otra parte, la detección de marcadores positivos para diferentes MSC de tejido, sigue siendo bastante empírica, y no hay consenso todavía existente en el fenotipo de MSCs de tejidos conectivos distintos de la médula ósea [Calloni *et al.* 2013].

El nicho ortodoxo de las MSCs en la médula ósea

El concepto de nicho celular se utilizó por primera vez a finales de los años 70. Estudios de Schofield en 1978 [Schofield *et al.* 1978] establecieron cómo distintos microambientes del bazo y la médula ósea podían inducir la hematopoyesis. El microambiente controlado de manera óptima promueve la ruptura de estados quiescentes con la activación proliferativa de las células madre. Es el propio nicho el que define con precisión cómo las células madre deben dividirse, cuál es el destino que las células hijas deben seguir, y cuándo empiezan a avanzar en el camino de la diferenciación. Por consiguiente, el nicho es una unidad estructural donde las interacciones a través de contactos celulares, factores solubles, y la matriz extracelular, poseen control de la biología de las células madre para que se sincronicen con las necesidades particulares del organismo. Otra dimensión del nicho significaría la pérdida de control sobre las células madre que lleva a mayores tasas de proliferación y baja diferenciación. Un desequilibrio en la tasa de proliferación-diferenciación daría lugar a la formación de una masa tumoral. Este aspecto adquiere relevancia esencial cuando se utilizan cultivos adherentes para investigar las MSCs, teniendo en cuenta que el sistema *in vitro* implica un cambio dramático, e impredecible, respecto al complejo estado *in vivo* en el que se encuentran las MSCs. El científicamente atractivo entorno de la médula ósea que suministra nichos hematopoyéticos específicos es compartido conjuntamente por las MSCs y las HSCs (células madre hematopoyéticas que producen toda la descendencia de sangre). Una ilustración de los nichos hematopoyéticos de médula ósea se adjunta en la figura 1.6 [Morrison *et al.* 2014]. Las MSCs representan originalmente una pequeña población en la rica superestructura celular de la médula ósea. Su frecuencia, medida por

Pittenger y colaboradores en 1999, oscila entre 0,001% a 0,01%. Las últimas investigaciones se han enfocado a seccionar diferentes poblaciones de células del estroma de la médula ósea y su papel en la promoción y mantenimiento de las HSCs. Las células mesenquimales perivasculares en los seres humanos, capaces de generar células osteoblásticas, expresan CD146 y los factores que promueven el mantenimiento de las HSCs, es decir, el factor de células madre (SCF, *stem cell factor* o KITL) y CXCL12 [Sacchetti *et al.* 2007]. Una caracterización más extensa de la expresión ha sido desarrollada en ratones. Factores como: Cxcl12 [Sugiyama *et al.* 2006], Nestine [Mendez-Ferrer *et al.* 2011], Lepr [Ding *et al.* 2012], Prx-1, Sp7 [Greenbaum *et al.* 2013] y Mx-1 [Park *et al.* 2012], definen poblaciones del estroma con capacidades similares a las CD146+. Células reticulares abundantes en CXCL12 (de nombre, células CAR), de nuevo junto a sinusoides, promovían una reducción de las HSCs cuando el gen CXCL12 era depleccionado, así como desacoplaban gravemente la capacidad adipogénica y osteogénica de estas células estromales [Sugiyama *et al.* 2006]. La posibilidad de que las MSC son parte del nicho hematopoyético fue apoyado además por la co-localización de células MSC Nestin-positivas y HSCs en el entorno de los vasos sanguíneos [Mendez-Ferrer *et al.* 2011]. Estos estudios proporcionan una fuerte evidencia de que BM-MSCs son un componente del nicho perivascular de las HSCs.



Haematopoietic niches illustrated in Morrison & Scadden. *Nature* (2014)

Figure 1.6 Vista esquemática de las MSCs en el nicho hematopoyético dentro de la médula ósea. *CAR cell: CXCL12-abundant reticular cell.*

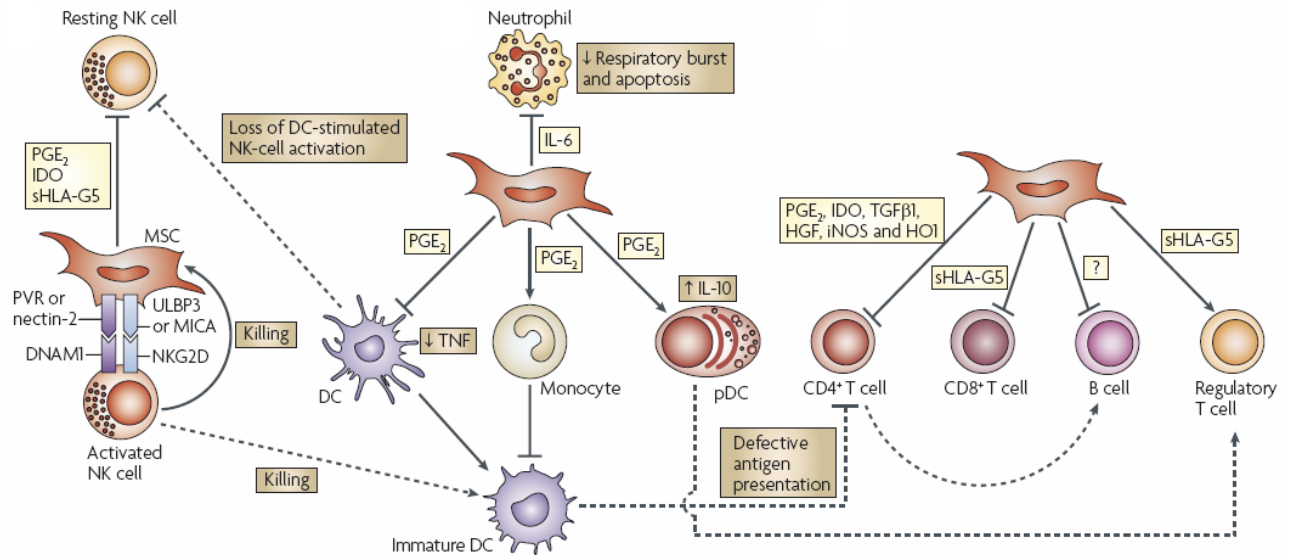
***Bona fide* MSCs de la médula ósea**

Multipotencia y auto-renovación fueron los caracteres pioneros inicialmente reconocidos por Owen y Friedenstein en las MSCs derivadas de médula ósea [Owen *et al.* 1988]. Ambos caracteres han sido progresivamente testados hasta la extensa colección de pruebas *in vivo*. Trasplantadas en vivo, estas células generan un hueso en miniatura, con su arquitectura histológica y su desarrollo completo de la hematopoyesis a partir de células huésped reclutadas [Sacchetti *et al.* 2007]. Como recientemente revisado por Bianco *et al.* [Bianco *et al.* 2013] “estas células son, o pueden ser útilmente conceptualizadas como, células madre de *bona fide*”. Se pueden formar todos los tejidos esqueléticos dentro de un segmento de hueso: tejido óseo, cartílago, adipocitos, fibroblastos y estroma de soporte de la hematopoyesis. En resumen, una sola célula de la médula ósea demostradamente productora de una progenie clonal, también ha demostrado ser la única responsable de la producción de todos estos otros tipos celulares. El potencial no esquelético de las MSCs individuales no se ha demostrado formalmente *in vivo*, ni la capacidad postulada de *trans-diferenciación* hacia linajes no mesenquimales. La identificación de poblaciones celulares que presentan estas habilidades de células madre se ha centrado en el estroma de las zonas sub-endosteales, y como se ha dicho antes, en componentes perivasculares dentro de la médula ósea, ambos sitios de residencia de los nichos hematopoyéticos [Morrison *et al.* 2014].

Comportamiento inmunomodulador y otros efectos tróficos con beneficios terapéuticos

Un gran hallazgo revelado por los trasplantes de médula ósea fue la baja inmunogenicidad que las MSCs parecían ejercer en los pacientes hospedadores. Progresivamente, los efectos inmunomoduladores y anti-inflamatorios de MSCs han sido estudiados tanto *in vitro* como *in vivo* [Aggarwal *et al.* 2009]. Las MSCs actúan tanto sobre el sistema inmune innato como sobre el adquirido, a través de la supresión de las células T, la supresión de maduración de células dendríticas, la reducción de la activación de células B y la proliferación, inhibición de la proliferación y la citotoxicidad de las células NK, o alterando a monocitos / macrófagos o neutrófilos (ver la revisión de [LeBlanc *et al.* 2012] para completar información). Algunos factores solubles secretados por las MSCs que juegan un papel clave en la regulación inmune e

inflamatoria incluirían: el factor- β 1 transformante (TGF- β 1), factor de crecimiento de hepatocitos (HGF), la prostaglandina E2 (PGE2), el interferón- γ (IFN- γ), la interleucina-10 (IL-10) y la 2,3-indolamina dioxigenasa (IDO). La figura 1.7 resume algunas interconexiones descritas entre MSCs y las células de la respuesta inmune [Uccelli *et al.* 2008].



Uccelli, Moretta & Pistola. *Nature reviews. Immunology.* (2012)

Figure 1.7 Acción inmunomoduladora de las MSCs.

Más recientemente, se descubrió que las MSC también pueden mostrar un efecto estimulante sobre la producción de linfocitos T reguladores CD8+ [Liotta *et al.* 2008]. Se ha propuesto un modelo análogo al sistema de activación de los macrófagos / monocitos [Waterman *et al.* 2010], donde dos estados de MSC pueden ser inducidos, uno inmunosupresor, el cual podría llegar a ser polarizado hacia el otro pro-inflamatorio. Los receptores tipo Toll TLR3 y TLR4 formarían el eje central que media la respuesta pro-inflamatoria a los estímulos de daño tisular.

Si el potencial como célula madre o las funciones de apoyo de la hematopoyesis ya eran suficientes para generar cientos de ensayos clínicos dirigidos hacia las MSCs, el citado privilegio de control sobre las reacciones inmunes cautivó por completo la atención de la comunidad biomédica hacia el tratamiento de enfermedades autoinmunes. Uno de los primeros casos tratados con éxito por MSCs fue la enfermedad post-trasplante de injerto contra huésped (EICH) [LeBlanc *et al.* 2004]. Aunque los efectos de las MSCs no persisten por mucho tiempo después de la administración [VonBahr *et al.* 2012], se cree que la acción trófica inmuno-regulatoria inmediata viene de la mano de los factores anteriormente citados.

Otra escena patológica, beneficiada en gran medida por la acción trófica de las MSCs, es el infarto de miocardio. Aunque se ha propuesto la diferenciación mioblástica de las MSC, los datos *in situ* no demuestran lo mismo [Iso *et al.* 2007]; es la liberación de factores lo que realmente favorece la supervivencia y mejora la funcionalidad de las células progenitoras cardíacas residentes, que finalmente lideran el progreso de regeneración del miocardio [Ranganath *et al.* 2012]. Beneficios equivalentes de estos factores de crecimiento se han demostrado también sobre progenitores de hepatocitos.

Del mismo modo, se descubrieron otros factores de crecimiento que facultan a las MSC con funciones neuroprotectoras. A mencionar: factor neurotrófico derivado del cerebro (BDNF), factor de crecimiento nervioso (NGF), neurotrofina-3 (NT-3), factor neurotrófico derivado de células gliales (GDNF), y algunos miembros de la familia del factor de crecimiento de fibroblastos (FGF) [Rosner *et al.* 2012]. Además, el daño mediado por respuesta inmune y apoptosis de las neuronas en diferentes enfermedades podrían ser rescatados por las MSCs, lo cual implica la inhibición de moléculas de estrés oxidativo liberadas por macrófagos activados y microglia en el sistema nervioso.

La hematopoyesis en cultivos a largo plazo de las MSC de médula ósea se basa en la expresión constitutiva de las interleucinas IL-6, IL-11, factor inhibidor de leucemia (LIF), macrófagos factor estimulante de colonias (M-CSF), y factor de células madre (SCF) [Majumdar *et al.* 1998]. La quimioquina altamente expresada CXCL12 es un potente atrayente y de retención para HSCs y MSCs, crítica en el mantenimiento de HSPCs en estado de reposo. La circulación de una proporción de HSPCs fuera de la médula ósea se controla sistemáticamente por inervación simpática gracias a los ritmos circadianos [Rankin *et al.* 2012].

Ante esta perspectiva, la potencial utilidad de las MSCs sigue siendo fuertemente investigada en el campo de la medicina regenerativa y terapia celular. En mayo de 2015, había 29,971 referencias en PubMed a “mesenchymal stem cell” o “mesenchymal stem cells”, y 19.596 a “mesenchymal stromal cell” or “mesenchymal stromal cells”. Respectivamente, para las mismas búsquedas en ClinicalTrials.gov, se registraron 498 y 98 estudios.

A pesar del gran conocimiento generado en las últimas décadas sobre las MSCs, todavía las vemos como piezas de un rompecabezas que no encajan bien unas con otras. Muchas preguntas sobre su identidad molecular siguen siendo inciertas. Preguntas que nacen de la heterogeneidad y versatilidad que demuestran, y que es inapropiadamente asignada a una única descripción fenotípica. El campo está todavía en espera de una mejor caracterización de los fenotipos de las MSCs con el fin de desentrañar su heterogeneidad y capacidades. La identificación de nuevos biomarcadores se convierte en una necesidad para llevar a cabo esta tarea, y por tanto, la

determinación de las moléculas clave que regulan la supervivencia celular, el crecimiento, el desarrollo y las funciones atribuidas a los patrones de secreción específicas desencadenadas y dirigidas al contexto biológico vecino.

Expresión génica

La expresión génica es el proceso por el cual la información contenida en la secuencia de un gen se utiliza en la síntesis de un producto génico funcional. Y la transcripción es el proceso molecular que lleva a cabo la expresión de un gen; a través de él, un segmento particular de ADN se copia en ARN mensajero (ARNm) por la enzima ARN polimerasa. La figura 1.8 ilustra cómo fluye la información de una región del gen del ADN, a una molécula de ARNm maduro, que sale del núcleo y se transcribe por los ribosomas en el producto funcional típico: una proteína.

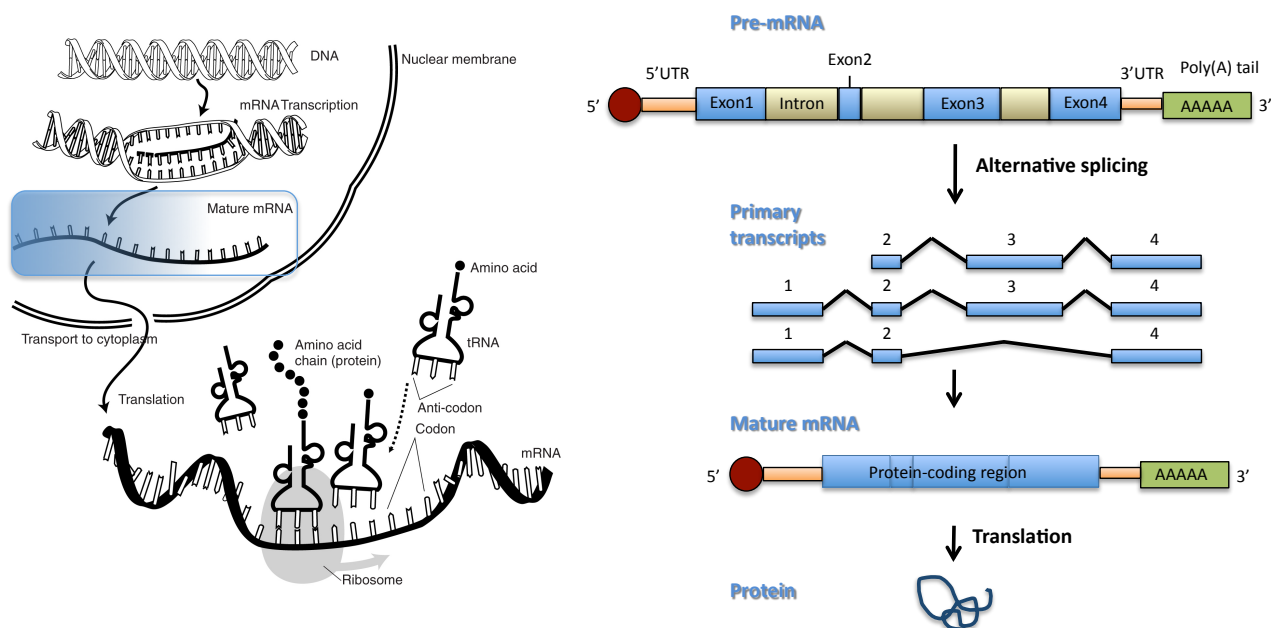


Figure 1.8 El proceso de expresión génica. *UTRs (*UnTranslated Regions*) son regiones no codificantes de los exones situadas en los extremos del ARNm.

También existen productos funcionales de genes que no son proteínas, son los considerados como genes no codificantes de proteínas: ARN ribosómico, ARN transferente, ARNs largos no codificantes, micro ARNs, etc. Estas versiones de ARN son biológicamente funcionales per se, y llevar a cabo diversas funciones: operativas, estructurales o reguladoras de la expresión génica. En este sentido, el ARNm representa la unidad intermedia de información transferida desde el genoma al proteoma. La cantidad de ARNm transcrito a partir de un solo gen es una medida de la

expresión de ese gen en el sistema celular. Aunque la producción de ARNm no se refleje exactamente en el repertorio proteico debido a los mecanismos adicionales de regulación postranscripcional, la biblioteca de todo el ARNm celular sigue siendo un buen recurso como punto de partida para buscar genes utilizados activamente por la célula. En una vista muy simplista, la expresión génica es la forma y la intensidad por la cual la información genética heredada se utiliza inicialmente por la célula, y por lo tanto, la primera puerta de regulación del fenotipo saliente.

Procesos de transcripción

La transcripción de genes eucariotas deja un transcrito primario de ARN (pre-ARNm), que primero tiene que someterse a una serie de modificaciones para convertirse en un ARNm maduro. Estos incluyen el encapuchado en 5', que es un conjunto de reacciones enzimáticas que añaden 7-metilguanosina (m7G) al extremo 5' del pre-ARNm, esta molécula aporta una protección al ARNm de la degradación por exonucleasas. La capucha de m7G se une a continuación al complejo heterodímero (CBC20 / CBC80), que ayuda en la exportación del ARNm al citoplasma y evita que pierda su capucha.

Otra modificación es la escisión y poliadenilación en 3'. Se producen si la secuencia de señal de poliadenilación (5'- AAUAAA-3') está presente en el pre-ARNm, al final de la secuencia de codificación de proteína. El pre-ARNm unido a una secuencia de ~200 adeninas (A) que forman la cola de poli-(A), la cual protege el ARN de la degradación. La cola de poli-(A) está unida a múltiples proteínas de unión a poli-(A) necesarias para la exportación del ARNm y la posterior iniciación de la traducción.

Probablemente la modificación más compleja del pre-ARNm eucariota para formar el ARNm maduro es el proceso de *splicing* (corte y empalme; ver el lado derecho de la figura 1.8). La secuencia de pre-ARNm contiene segmentos alternantes de exones e intrones. Durante el empalme, regiones intrónicas pueden ser quitadas. Para este objetivo, se pliegan a modo de estructuras en bucle, y el complejo del spliceosoma cataliza dos reacciones de transesterificación que liberan el intrón y empalman los dos exones vecinos juntos. Por otro lado, los exones pueden ser eliminados o retenidos en el ARNm maduro. Este proceso llamado "splicing alternativo" crea series de transcritos diferentes pero originados de un mismo gen. Estos transcritos con diferente carga de exones pueden ser traducidos a proteínas diferentes, cada una de ellas será reconocida como una isoforma de la proteína. Se ha demostrado que los tejidos de mamíferos hacen un uso discreto de isoformas; más aún, algunas isoformas podría llegar a ser muy específicas de un tejido determinado.

Tecnologías para cuantificar la expresión génica

Una de las técnicas más extendidas de biología molecular, para estimar la cantidad de transcritos por gen que cada célula está produciendo, es la reacción en cadena cuantitativa, o a tiempo real, de la polimerasa (qPCR, o RT-PCR). La síntesis se basa en las propiedades de hibridación y amplificación de ácidos nucleicos, y puede ser cuantificada simultáneamente durante los ciclos sucesivos de la reacción. Después de la reacción en cadena de polimerasa, cada vez que el ARN (o más exactamente, su copia complementaria de ADN -ADNc) se amplifica, la intensidad de la luz de las entidades recién sintetizadas puede ser detectada. Se utiliza normalmente para la cuantificación relativa, que mide el factor de cambio de los niveles de abundancia del ADNc respecto a una referencia control predefinido. A partir de la qPCR se obtiene una medida de la expresión del gen estudiado comparándola con la cantidad del gen control (por lo general, un gen de basal para la célula (*housekeeping*) altamente expresado).

LA PCR a tiempo real se ha convertido en una de las técnicas más fiables y rentables para cuantificar ácidos nucleicos; es por eso que hoy en día se considera el experimento estándar para la validación de resultados obtenidos por técnicas de expresión génica de alto rendimiento [VanGuilder *et al.* 2008].

Los *chips* génicos (o *microarrays*) y otras ingenierías génicas de alto rendimiento, desarrolladas principalmente en las dos últimas décadas, han automatizado y miniaturizado técnicas de genética molecular como la hibridación de ácidos nucleicos y la PCR. La producción a gran escala de datos, en tiempos cada vez más cortos, ha permitido el estudio de los sistemas celulares: desde la visión simplista de un gen, al nivel global de todo un genoma (el nivel "ómico"). Desde los años 90, cuando el primer secuenciador masivo de ADN fue creado (*Massively Parallel Signature Sequencing of Lynx Therapeutic*), estas tecnologías han ido en aumento en capacidad y reducción de costes, convirtiéndose en herramientas principales para la investigación biológica y biomédica actual (figura 1.9).

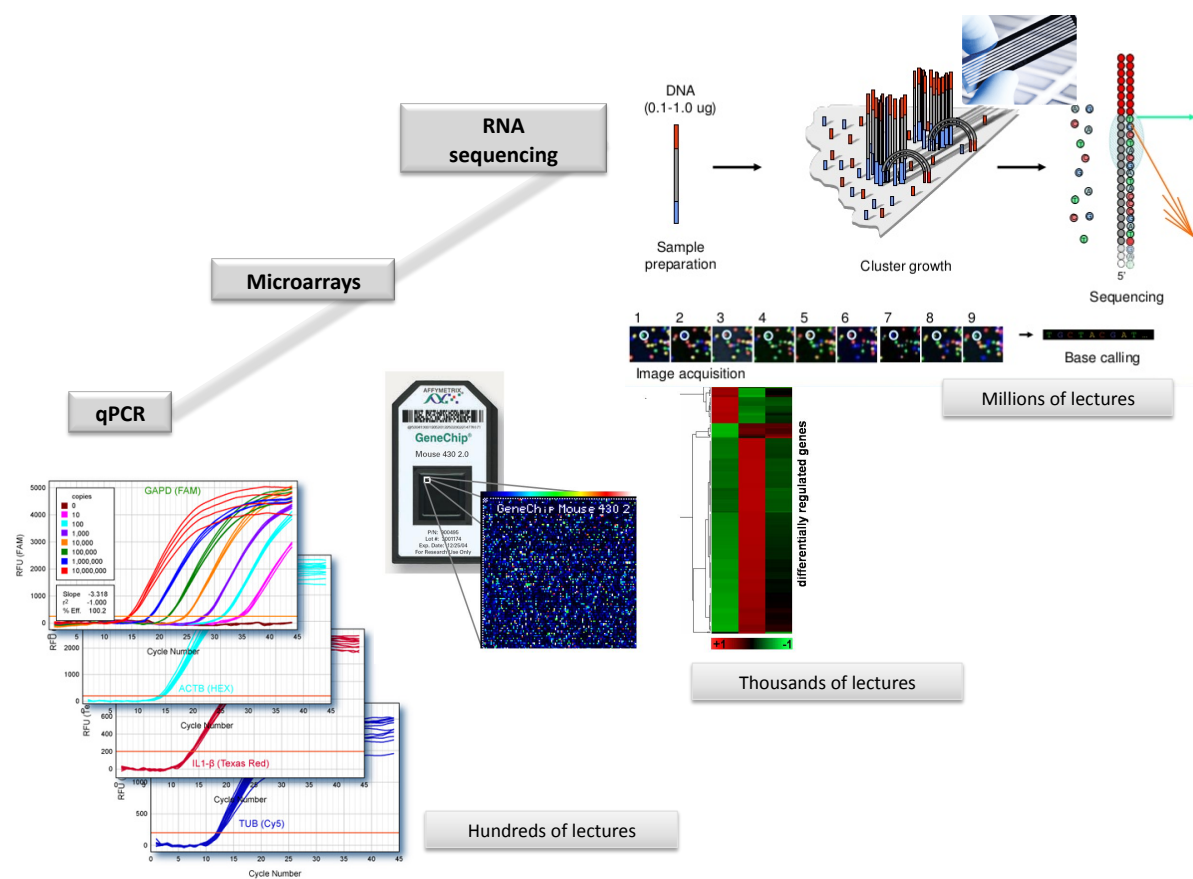


Figure 1.9 Gene expression technologies.

El término “-ómico” se determina al enfoque global que estas técnicas proporcionan. Se puede aplicar a diferentes compartimentos biológicos de estudio, tales como: transcriptoma, proteoma, metiloma, kinome, metaboloma ... y muchos otros términos que han aparecido a medida que han ido creciendo las nuevas plataformas.

En el campo de la transcriptómica, los *microarrays* de ADN de alta densidad protagonizaron una revolución, en particular en la investigación biomédica. Sin embargo, hoy en día, cuando la secuenciación de ARN se ha convertido en una realidad palpable, los *microarrays* siguen siendo útiles y utilizados en multitud de estudios. Sus costes asequibles y su gran reproducibilidad la han convertido en la plataforma de elección para la mayoría de los estudios de expresión génica. Cientos de publicaciones y cientos de muestras disponibles en diferentes repositorios públicos mantienen el interés de los investigadores en estos dispositivos [Rung *et al.* 2013].

Los *microarrays* son dispositivos físicos con miles de microceldas que contienen sondas de oligonucleótidos para hibridar contra el ARN (ADNc) de la muestra. Las señales de fluorescencia de pares hibridados sonda-tránsito se detectan y analizan, produciendo una imagen que puede ser leída como una matriz o array de valores de intensidad. Cada intensidad representa una

medida de la concentración de ARN unido. Una consideración esencial de esta tecnología es que esas sondas de oligonucleótidos tienen que ser definidos de antemano en base al conocimiento existente del genoma de referencia.

La importancia que ha adquirido la tecnología de *microarrays* y su entrada en la comunidad investigadora se puede observar navegando en uno de los mayores repositorios de datos de expresión: *The Gene Expression Omnibus, GEO* (<http://www.ncbi.nlm.nih.gov/geo/>) [Barrett et al. 2013]. A finales de 2014, el modelo de "Genoma Humano 133 Plus 2.0" de Affymetrix (entrada: GPL570) contenía 82,695 muestras hibridadas, distribuidas en 3.029 series de datos. Más aún, Affymetrix representaba 63 entradas de plataformas diferentes de hibridación de oligonucleótidos.

La restricción de los diseños basados en conocimiento previo de los *microarrays*, capaces de cuantificar genes sólo anotados, ha impulsado el desarrollo de tecnologías diseñadas sin requerimiento de información *a priori*, como son las plataformas de *Next Generation Sequencing* (NGS).

La secuenciación masiva en paralelo de ARN (ARN-seq) nos da la posibilidad real de abarcar todo el transcriptoma, proporcionando la herramienta perfecta para el descubrimiento de nuevas entidades de transcripción. Para hacer RNA-Seq, el ARN (total, o seleccionado por las colas de polyA) se convierte en una biblioteca de fragmentos de ADNc con adaptadores unidos a uno o ambos extremos, y cada molécula es entonces secuenciada por NGS. Las lecturas resultantes se alinean sobre un genoma, o transcriptoma, de referencia, para producir un mapa de transcripción a gran escala, que consiste tanto en la estructura, como en el nivel de expresión, del gen y sus transcritos. La tecnología ARN-seq ofrece una escala de medida ilimitada de abundancia no relativa, y a muy alta reproducibilidad [Marioni et al. 2008]. Estas ventajas, y las futuras reducciones de precios, hacen del ARN-seq una estrategia cada vez más atractiva, sobre todo, para organismos de genomas desconocidos.

Estudio de la expresión génica en MSCs humanas

Una población de células madre y estromales, susceptibles de ser aisladas a partir de diversos tejidos humanos, es lo que conocemos como la población de Células Estromales/Madre Mesenquimales (MSCs). Los rasgos específicos que conducen a la separación de estadios no comprometidos de los estadios diferenciados aún no han sido descifrados. La extensa ocupación, en localizaciones anatómicas del cuerpo, es también una fuente de información poco explorada que nos acerca a la variabilidad entre MSCs. Esta información podría estar asociada a diferentes presiones de linaje, determinadas por el tejido, dicho de otro modo, la variabilidad de

predeterminación del compromiso celular, así como el papel específico de señalización de tejido en la renovación natural del tejido, y la adaptable especialización funcional de las MSCs en el tejido de residencia. Tanto por la variabilidad determinada por el tejido, como por los aspectos comunes a las MSCs, establecer un fenotipo definido para las MSC está aún en trámites de caracterización.

Como las correas del fenotipo se vienen tiradas por la expresión de los genes, muchos estudios de perfiles genómicos se han abordado con el fin de responder a las preguntas no resueltas sobre las MSC. Beneficiados por la gran explosión de las herramientas de análisis de la genómica, los genes expresados diferencialmente entre las MSC de distintos tejidos fueron inicialmente explorados (como por ejemplo, en estudios pioneros de Wagner et al. [Wagner et al. 2005]). La mayoría de los estudios desde entonces, se han centrado en genes regulados durante procesos de diferenciación *in vitro* (largamente revisado en [Menicanin et al. 2009]).

A medida que la tecnología ha ido avanzando, las firmas han conseguido mejorar en sensibilidad y en especificidad. Sin embargo, los diseños de sondas distintas o protocolos químicos interfieren en la comparación de resultados entre análisis independientes. Las principales firmas publicadas que se han centrado en tecnologías de alto rendimiento se han revisado a lo largo de este trabajo y muestran correspondencias importantes en algunos grupos de genes.

Con el fin de caracterizar mejor los rasgos que identifican a las MSC humanas, hemos aprovechado los datos transcriptómicos accesibles recopilados de diferentes tipos de células con diferente grado de compromiso. También hemos completado el estudio con datos de producción propia para explotar la variabilidad inherente presente en MSC de diferentes tejidos. Se ha demostrado que esta variabilidad se mantiene a lo largo del periodo inicial de expansión en cultivo [Chang et al. 2002, Covas et al. 2008]. Creando un marco transcriptómico amplio con las células estromales hemos sido capaces de delimitar las distancias relaciones entre ellas, el posicionamiento de unas respecto a las otras en contextos funcionalmente significativos: etapas de compromiso celular o tejido de origen. Una lote de 15 microarrays de expresión de producción propia, el RNA de 6 cultivos de MSC secuenciados, y 264 muestras de *microarrays* de datos públicos, completan el fondo de información que alimenta el estudio realizado.

Escarbando un poco más en los perfiles de expresión génica mediante una aproximación más sofisticada hemos sido capaces de identificar una firma génica pulida que recoge las características del linaje mesenquimales en 489 genes. La vinculación funcional entre estos genes permitió establecer las rutas basales que estas células necesitan poner en marcha para su supervivencia. También se examinaron los genes específicos asociados a cada tejido, y en particular sus patrones asociados de citoquinas. Hemos ido más allá a través del transcriptoma

de la población de médula ósea (BM-MSK) y hemos investigado las potenciales interacciones dentro del nicho hematopoyético, con células madre y progenitoras hematopoyéticas (HSPCs). Las señales intercambiadas y el diálogo cruzado entre estas dos, determinan el establecimiento del microambiente de una médula ósea funcional.

Por último, la superposición de los resultados de nuestras diferentes aproximaciones analíticas, con otras firmas ya publicadas, ha permitido rescatar genes que aparecen repetidamente en resultados no relacionados, y que pueden demarcar nuevos rasgos característicos de las MSC.

Objetivos

Enfoque principal

Esta tesis ha sido formulada con un enfoque decriptivo y dirigido por los propios datos, en lugar del pensamiento tradicional basado en hipótesis de trabajo. Se consideró que las poblaciones de células madre mesenquimales estromales humanas, a pesar de los criterios existentes que comúnmente las definen, siguen siendo poblaciones fenotípicamente mal definidas, dentro, y entre, diferentes orígenes tisulares. Por lo tanto, el alcance general de esta tesis es el de mejorar la caracterización del fenotipo de células mesenquimales estromales/madre humanas, hacia el análisis de la totalidad de su paisaje de expresión genómica. Este enfoque convergerá en la generación de nuevas hipótesis acerca de la biología de éstas células.

Objetivos centrales

1. Aislamiento y cultivo de MSCs humanas a partir de diferentes tejidos: tejido adiposo, médula ósea y placenta coriónica; siguiendo los criterios de caracterización de la ISCT (Sociedad Internacional de Terapia Celular) de inmunofenotipo y capacidad de diferenciación tri-línea *in vitro*.
2. Definir, aplicando tecnologías de expresión de alto rendimiento y métodos bioinformáticos, los perfiles de expresión de genes específicos asociados a cada subtipo de MSC de tejido, y compararlos con otros tipos de células relacionados, como las células madre y progenitoras hematopoyéticas (HSPCs), células diferenciadas derivadas de las MSC y otras células del estroma, como fibroblastos.
3. Obtener el perfil transcriptómico común a las MSC, identificar los genes que pueden ser considerados en forma activa, y definir una firma común de expresión génica de MSC humanas.
4. Construir redes relacionales de genes derivados de las firmas de expresión para las MSC, y detectar los nodos clave, como factores de transcripción (TFs), citoquinas específicas o nuevos marcadores de CD para una mejor caracterización del retrato transcriptómico de las MSC.

Metodología

Brevemente, y de forma esquemática:

En el ámbito experimental - molecular y celular:

- i. Aislamiento y crecimiento de células madre mesenquimales (hMSCs) humanas derivadas de placenta, médula ósea, sangre de cordón umbilical y tejido adiposo. Todos los procedimientos realizados están en conformidad con la Declaración de Helsinki y todas las muestras humanas fueron recogidas tras consentimiento informado de los donantes, aprobado formalmente por el Comité de Ética del Área de Salud y el Hospital Universitario de Salamanca.
- ii. Aislamiento de células madre hematopoyéticas humanas (hHSCs) derivadas de médula ósea.
- iii. Realización de los ensayos de diferenciación celular, en los que conseguimos verificar la multipotencialidad propia de las hMSCs expandidas in vitro.
- iv. Verificación de los patrones citométricos definidos por la ISCT en los sucesivos pases en cultivo.
- v. Extracciones y purificaciones del mRNA celular para los estudios transcriptómicos.
- vi. Hibridación del mRNA en un set de 15 arrays de exones (Affymetrix), en el que se incluyeron 3 arrays por cada tipo celular de interés.
- vii. Secuenciación masiva de RNA (RNA-Seq, utilizando la plataforma Illumina) de un set de 6 muestras de hMSCs (procedentes de médula ósea y placenta).

En el ámbito computacional - bioinformático:

- i. Minería de bases de datos genómicas públicas (p. ej. GEO) para la recolección y análisis integrado de series de datos de expresión genómica (realizados con plataformas de microarrays Affymetrix) de hMSCs, hHSCs, y células derivadas de ambas con distintos grados de diferenciación.
- ii. Diseño y creación de un repositorio con todos los datos recopilados, que nos permita establecer un panel transcriptómico en el que poder analizar la dinámica del desarrollo de nuestras células.
- iii. Tratamiento estadístico y computacional de los datos del repositorio: pre-procesamiento, integración y normalización de las series de datos de distinto origen.

- iv. Realización de un meta-análisis bioinformático dirigido a la búsqueda patrones transcriptómicos responsables de las decisiones biológicas propias de las MSCs humanas.
- v. Perfilado del patrón de expresión de los marcadores biomoleculares propios de hMSCs y hHSCs a partir de los datos obtenidos en nuestro set experimental de arrays de exones.
- vi. Análisis de los datos extraídos por secuenciación masiva del mRNA celular (RNA-Seq). Caracterización de la huella transcriptómica común a las hMSCs de médula ósea y de placenta.
- vii. Análisis de predicción de los factores de transcripción (TFs) reguladores de las hMSCs.
- viii. Generación de redes de regulación génica para las hMSCs.

Resultados

Resumen de Resultados

1. Variaciones citológicas entre células estromales: MSCs de médula ósea, e placenta, de tejido adiposo, y fibroblastos de piel.

- La morfología de las diferentes células en cultivo resultó similar, con la característica forma fibroblástica en cultivos adherentes al plástico.
- La cinética de crecimiento presentó ligeras variaciones. El tiempo de duplicación celular resultó menor en fibroblastos y MSCs de placenta. Sus tasas de crecimiento fueron mayores respecto a MSCs de médula y tejido adiposo.
- Los ensayos de citometría verificaron el inmunofenotipo esperado para todas las poblaciones de MSCs. Así como resultó muy similar para fibroblastos de piel.
- Los ensayos de diferenciación *in vitro* verificaron la capacidad de las MSCs de generar adipoblastos, condroblastos y osteoblastos.

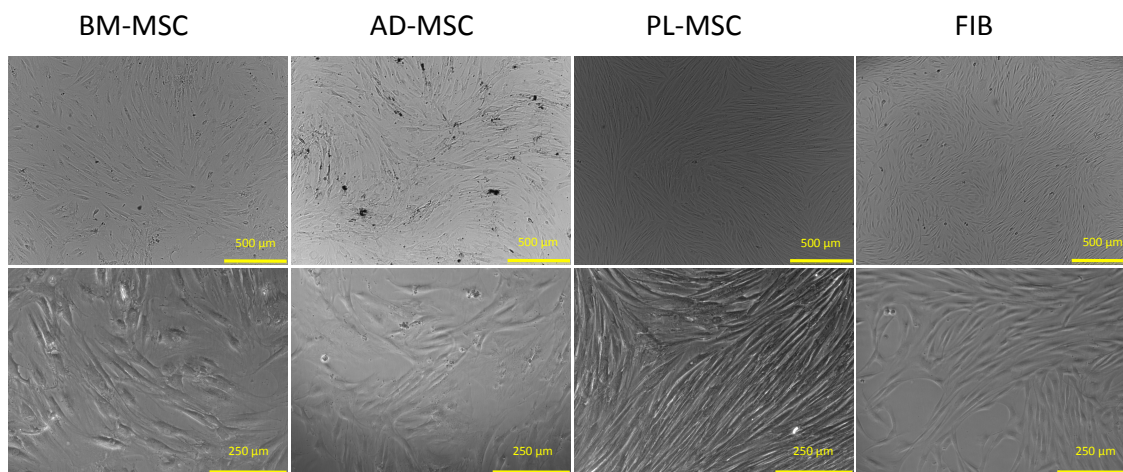


Figura 4.1: Fotografías del microscopio de células estromales humanas en cultivo: micrografías de contraste de fase tomadas en el pase tres, visto a ampliaciones 4x y 10x.

DT Contrasts	2.5% CI of difference	97.5% CI of difference	p-value
** BM vs PL	3.58	12.66	2.35E-04
BM vs AD	-7.74	9.33	6.49E-01
** PL vs AD	-11.68	-3.12	3.23E-03
** FIB vs AD	-14.57	-4.47	4.04E-03
FIB vs PL	-2.37	0.01	5.39E-02
** FIB vs BM	-15.03	-4.31	3.11E-04

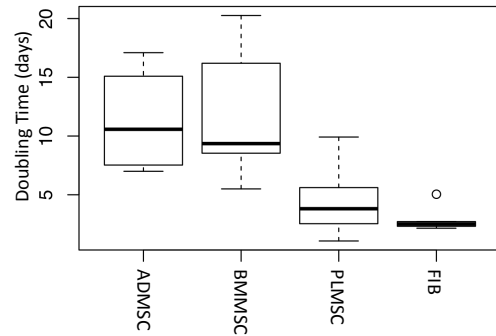


Figura 4.2: Análisis de tiempos de duplicación poblacional: (izquierda) una tabla con los resultados de los test de Wilcoxon de los tiempos de duplicación contrastadas entre los tipos celulares estromales; (derecha) un diagrama de cajas de las distribuciones de los tiempos de duplicación por cada cultivo celular. DT: Tiempo de duplicación. IC: Intervalo de confianza.

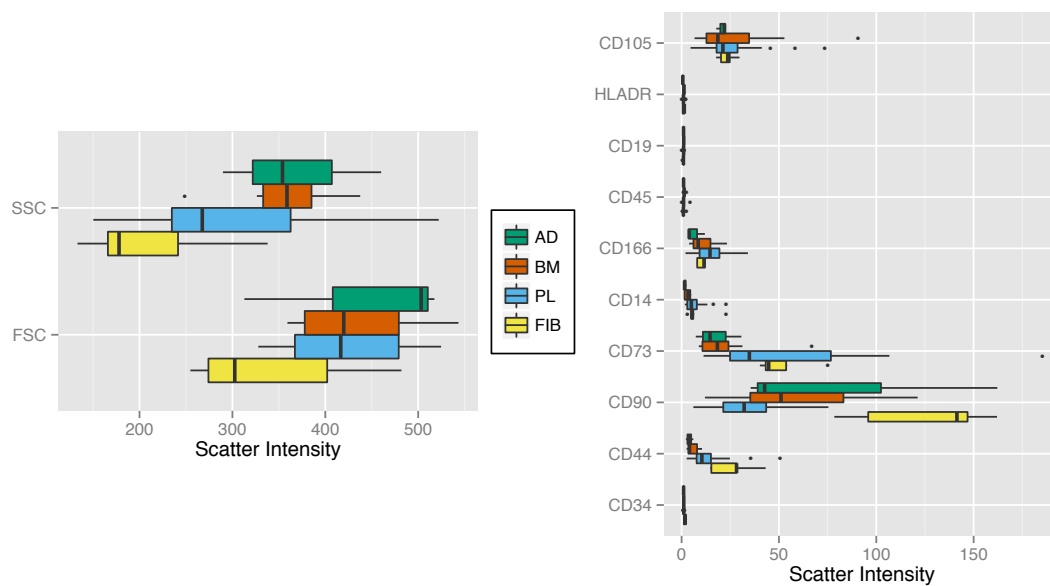


Figura 4.3: Diagramas de caja de intensidades de dispersión del citómetro de flujo, normalizados, recogidos de ensayos de inmunofenotipado. SSC: dispersión lateral. FSC: dispersión frontal. Las etiquetas de los tipos de células del estroma: (AD) MSCs de tejido adiposo; (BM) MSCs de médula ósea; (PL) MSCs placentarias; Fibroblastos (FIB) dérmicos.

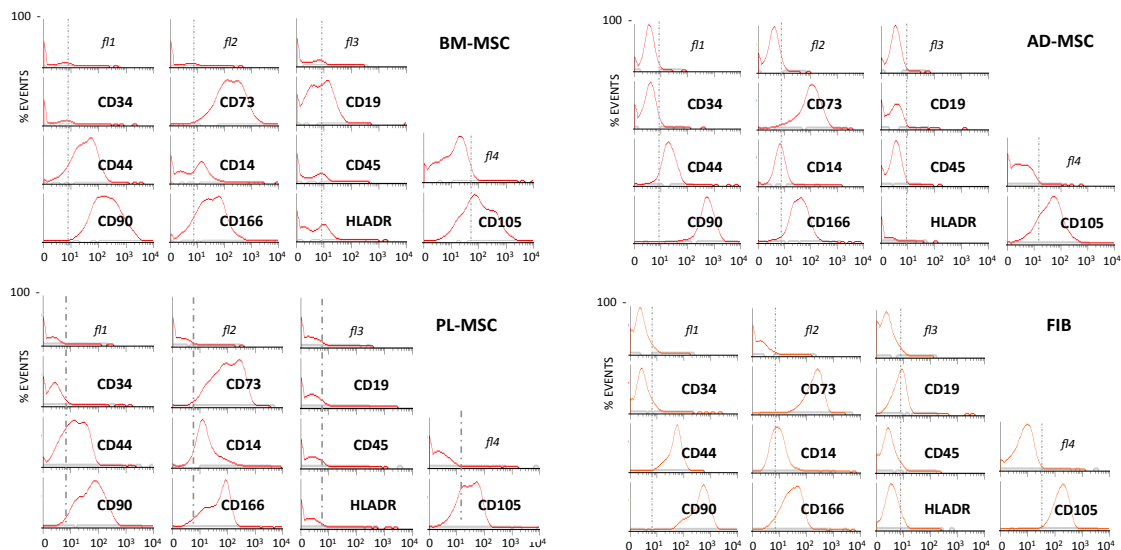


Figura 4.4: Histogramas de citometría de flujo que muestran el inmunofenotipado estándar esperado para las poblaciones de células del estromales estudiadas.

2. Huella de expresión génica de las MSCs obtenida a partir de secuenciación masiva de ARN (ARNseq)

- Se midieron los niveles absolutos de ARNm producidos por MSCs de placenta y médula ósea (BM-MSC y PL-MSC). Los ensayos de secuenciación dieron un promedio de 42,14 millones de lecturas por muestra.
- Se valoró la bondad de la técnica de secuenciación a través de la expresión obtenida para los marcadores CD estándar de citometría de flujo. Como cabía esperar, CD90, CD73, y CD105, presentaron niveles de expresión elevados, acorde con los resultados de citometría. Los marcadores de línea hematopoyética, CD34 y CD45, no produjeron lecturas de secuenciación por encima de valores detectables.

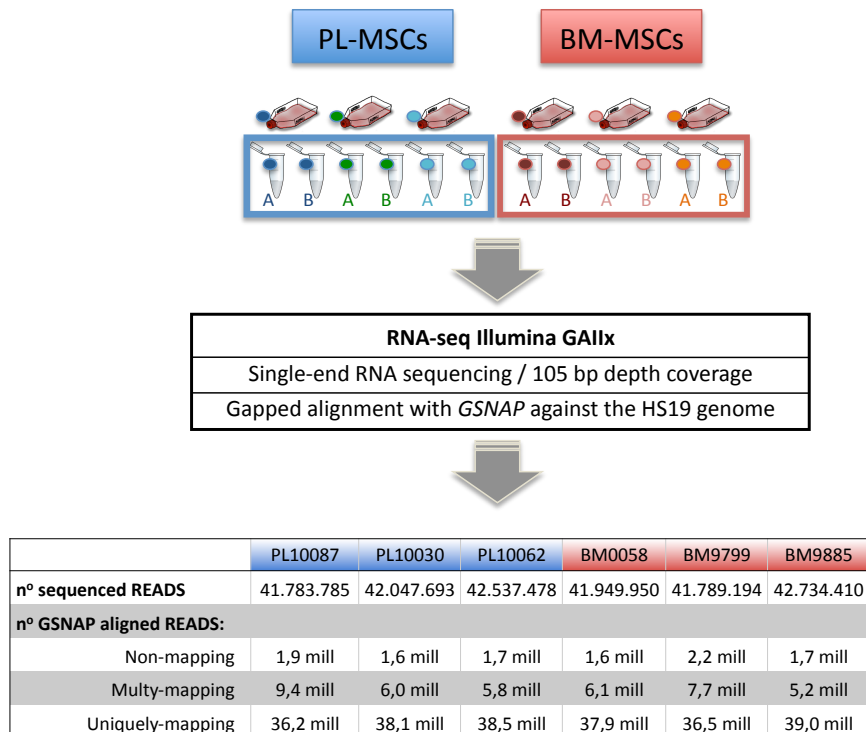


Figura 4.5: Esquema del diseño de los ensayos de secuenciación y recuento de las lecturas de salida. De arriba a abajo: Introducción de las muestras (2 réplicas técnicas de cada librería de ADNc fueron generadas). Detalles de la secuenciación y el alineamiento de las lecturas. Tabla final del las lecturas obtenidas y resultado del alineamiento.

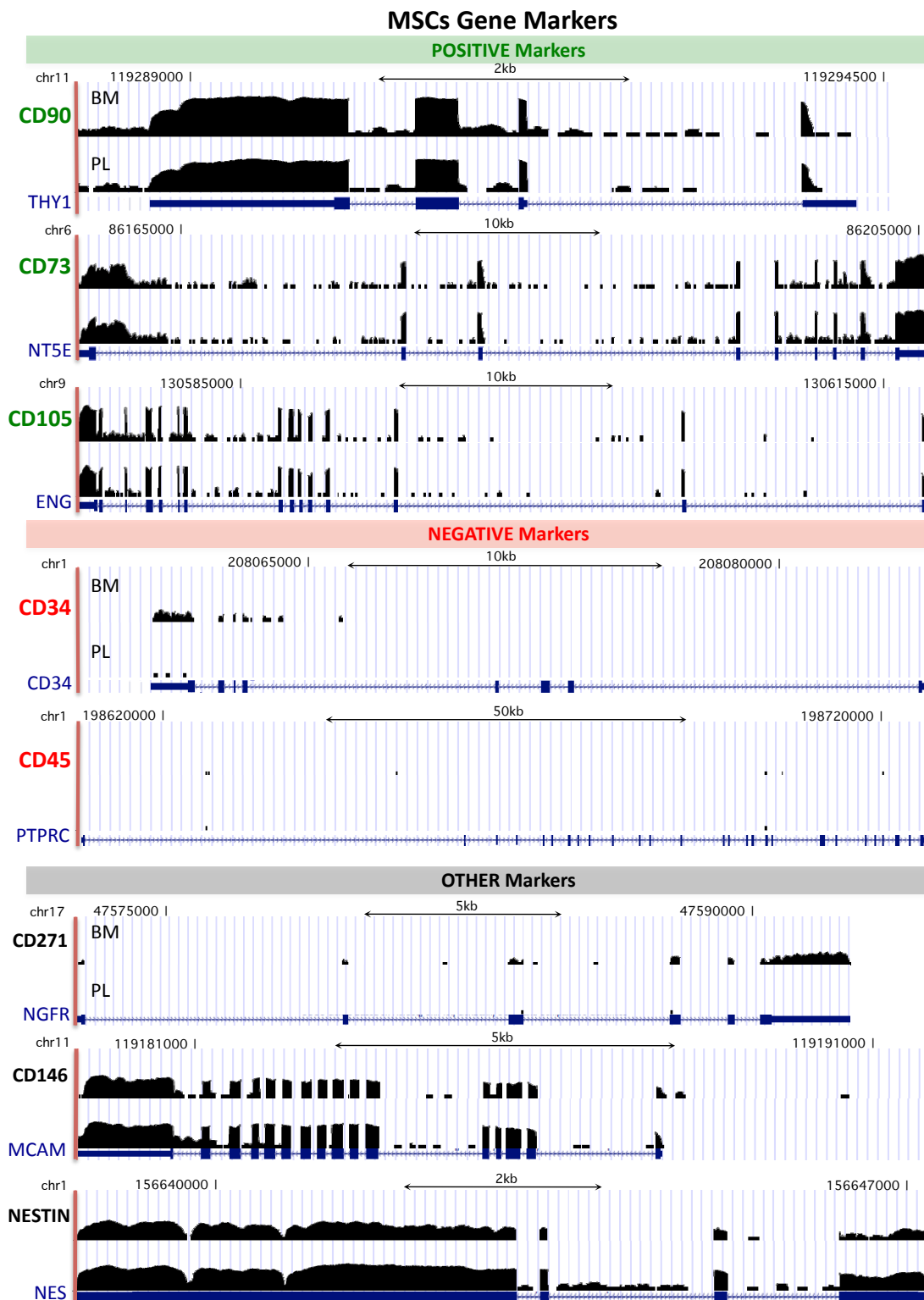


Figura 4.6: Lecturas de ARNseq mapeadas sobre los loci de 8 genes marcadores específicos.

- A partir de los datos de secuenciación masiva se generó una huella de expresión génica de MSCs que comprende 5.271 genes codificantes de proteínas.

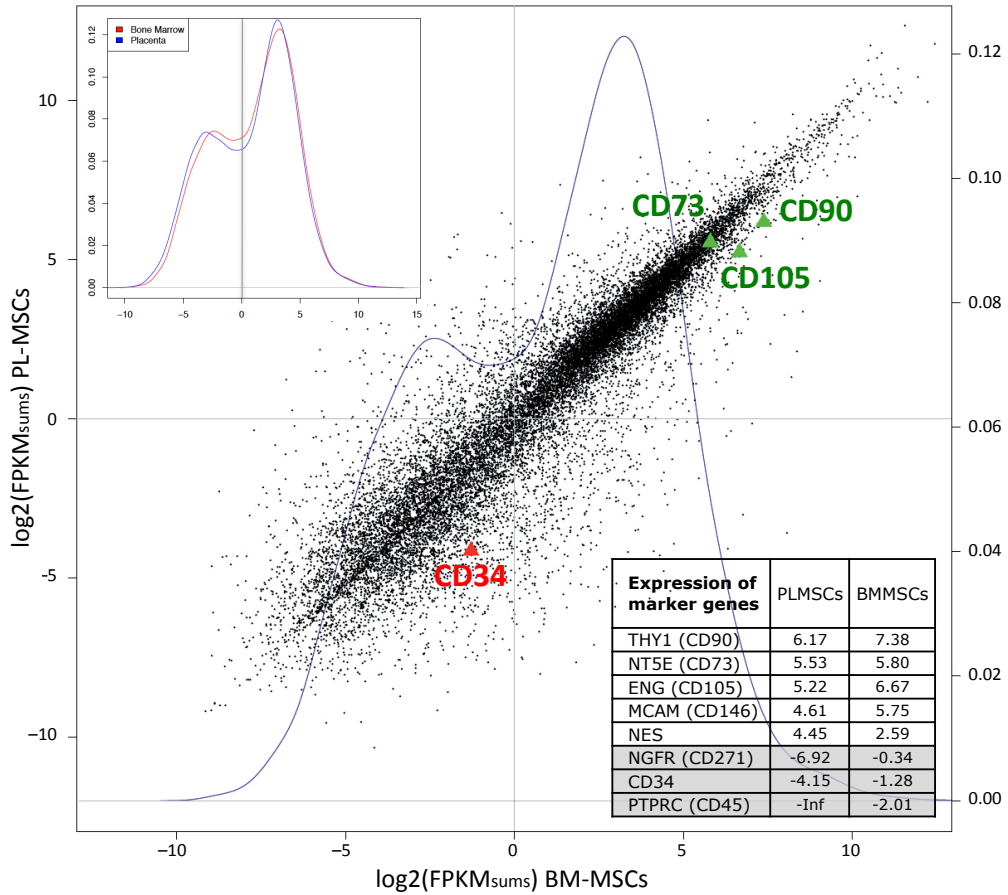


Figura 4.7: Expresión global: los datos de expresión de ARNseq de MSC humanas. **Centro:** Gráfico de dispersión de los valores log2 (FPKMsum) para cada gen en las muestras de BM-MSC (eje X) frente a las muestras PL-MSC (eje Y). Se ha destacado en verde y rojo la posición de los genes marcadores estándar positivos y negativos. **Inserto arriba-izquierda:** Curvas de densidad de los FPKMsum de cada tipo MSC. **Inserto abajo-derecha:** Tabla que indica los valores log2 (FPKMsum) correspondientes a los 8 genes marcadores. FPKM > 0 en blanco; FPKM < 0 en gris.

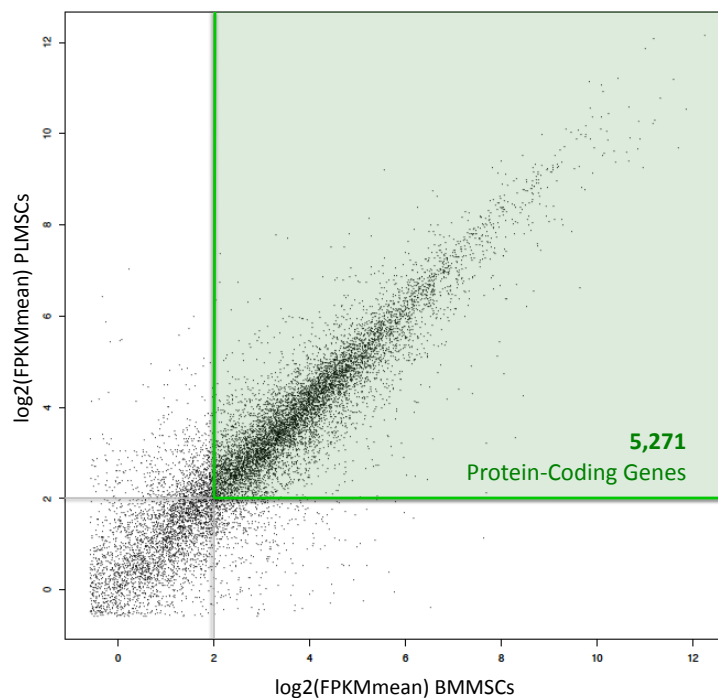


Figura 4.8: Gráfico de dispersión de los \log_2 (FPKMmeans) que incluye la huella de expresión de genes codificantes de BM- y PL-MSc (región verde) y comprende el 95% de la distribución.

- En un análisis más funcional de la huella delineada, se observaron bloques funcionales de genes ocupan áreas discretas a lo largo del espectro de la huella de expresión. Se mapearon genes *housekeeping*, factores de transcripción, y genes propios de la capacidad *stem*.
- Seguidamente se realizó un análisis de enriquecimiento sobre "vías de señalización" de KEGG dentro de la huella expresión de 5,271 genes. En particular, se observó un alto enriquecimiento en la vía de señalización mTOR, así como en ERBB, TGF β y vías de señalización Notch. También se detectó enriquecimiento en la vía de diferenciación de los osteoclastos.

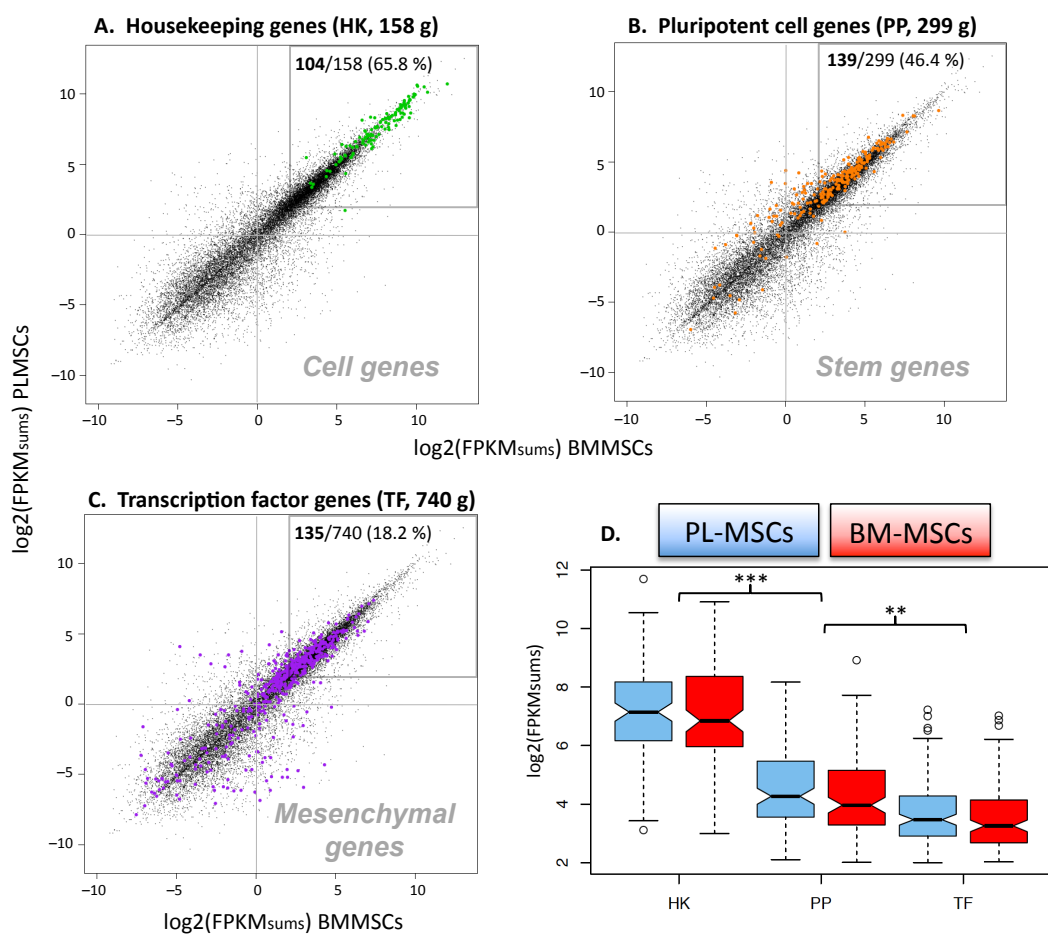


Figura 4.9: Mapeo de varios conjuntos de genes en los datos de expresión de ARNseq de MSC. Gráficos de dispersión de log₂ (FPKMsum), en BM- contra muestras PL-MSC, marcados en colores diferentes los distintos grupos de genes.

- Un análisis en mayor profundidad de los factores de transcripción hallados en la huella de expresión permitió definir una serie de factores meta-reguladores, o controladores maestros del transcriptoma de MSCs.

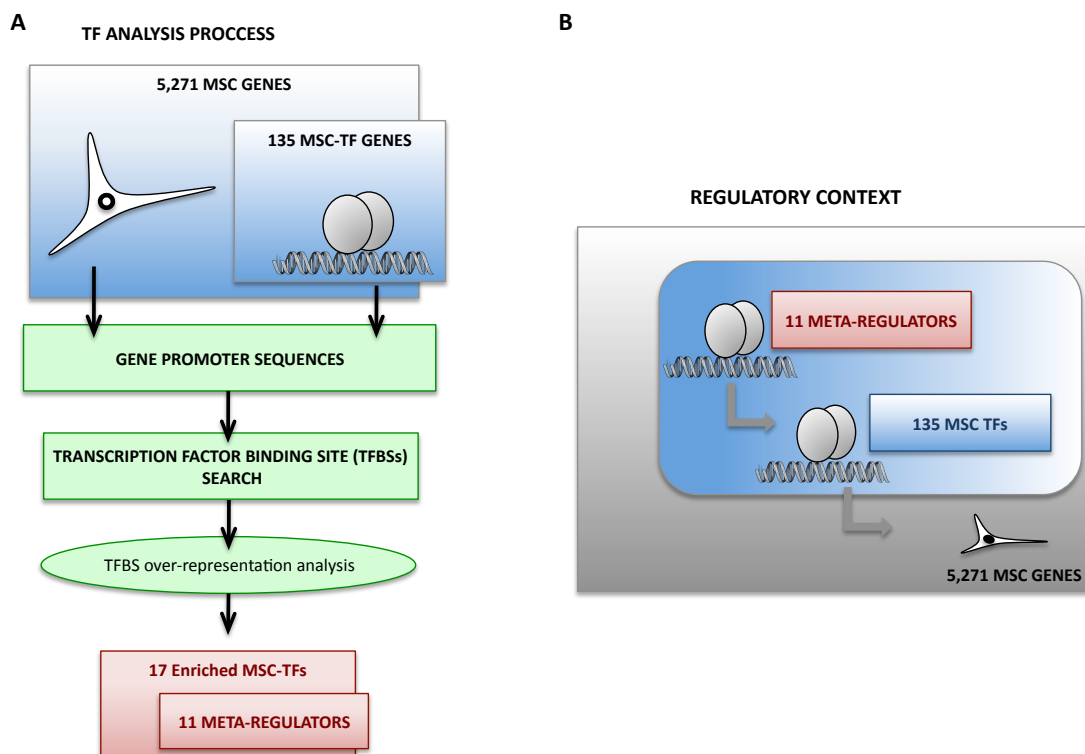


Figura 4.10: Vista esquemática de (A) el análisis de enriquecimiento de sitios de unión a factores de transcripción (TFBS), y (B) el modelo regulador sugerido.

Tabla de 11 factores meta-reguladores, cuyos TFBSs resultaron enriquecidos a lo largo de las regiones cis-reguladoras de 135 genes de factores de transcripción detectados en la huella de la expresión de MSC.

Meta-Regulator TFs (expressed & enriched)	TFBSs (MATRIX_ID)	RSAT p-value (FDR) <0.05	RSAT (RRS)	OPOSSUM Z-score >5	TRANSFIND p-value (FDR) <0.05	Sum
ARNT	MA0004.1	1	0	1	NA	2
ARNT	MA0006.1	1	0	1	NA	2
ARNT	V\$ARNT_01	1	0	NA	1	2
ATF4(CREB2)	V\$ATF4_Q2	1	0	NA	1	2
ATF4(CREB2), CREB3	V\$CREB_Q4	1	0	NA	1	2
ATF4(CREB2), CREB3	V\$CREBATF_Q6	1	0	NA	1	2
HES1	V\$HES1_Q2	1	0	NA	1	2
KLF4	MA0039.2	1	0	1	NA	2
MYC	V\$MYC_Q2	1	0	NA	1	2
MYC, MAX	V\$MYCMAX_03	1	0	NA	1	2
NFYA	MA0060.1	1	0	1	NA	2
USF1, USF2	V\$USF_01	1	0	NA	1	2
USF1, USF2	V\$USF_02	1	0	NA	1	2
USF1, USF2	V\$USF_C	1	0	NA	1	2
VDR	V\$VDR_Q3	1	0	NA	1	2

- Se estudiaron los estados de metilación de las islas CpG de los factores de transcripción definidos para MSCs.

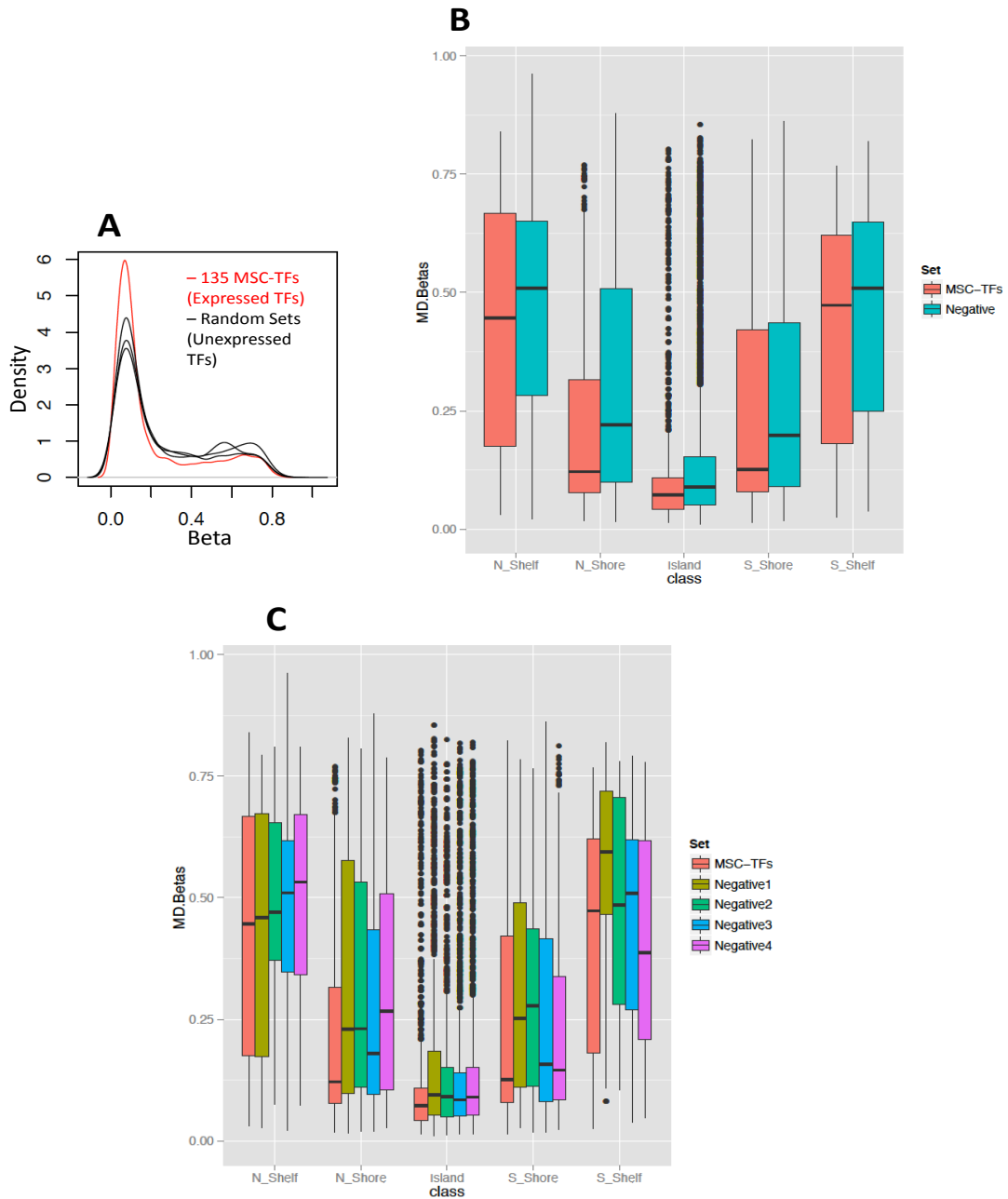


Figure 4.11 (A) Distribuciones de metilación del ADN -desidades de valores Beta de metilación- que corresponden a 135 factores expresados (rojo) y de otros conjuntos seleccionados al azar como control negativo (negro). El gráfico muestra mayor acumulación de mediciones de metilación en torno a valores más bajos de Beta para MSC-TFS. (B) Diagrama de cajas de las distribuciones de metilación de los 135 factores de transcripción de MSCs (en rojo) en comparación con un conjunto negativo de otros 135 TFS

(en azul) que no están presente en la huella de MSC. Diferentes regiones de las islas CpG también se dividieron para su análisis. (C) De igual modo, este diagrama de caja compara las distribuciones de metilación de los 135 MSC-TFs contra 4 series negativas aleatorias. **MD.Betas:** La mediana de los valores de beta a lo largo de las muestras.

- A partir de los 135 factores de transcripción definidos en estas MSCs se generó una red de de interacción entre ellos, que permite intuir de qué modo interactúan para promover la regulación de la expresión celular.

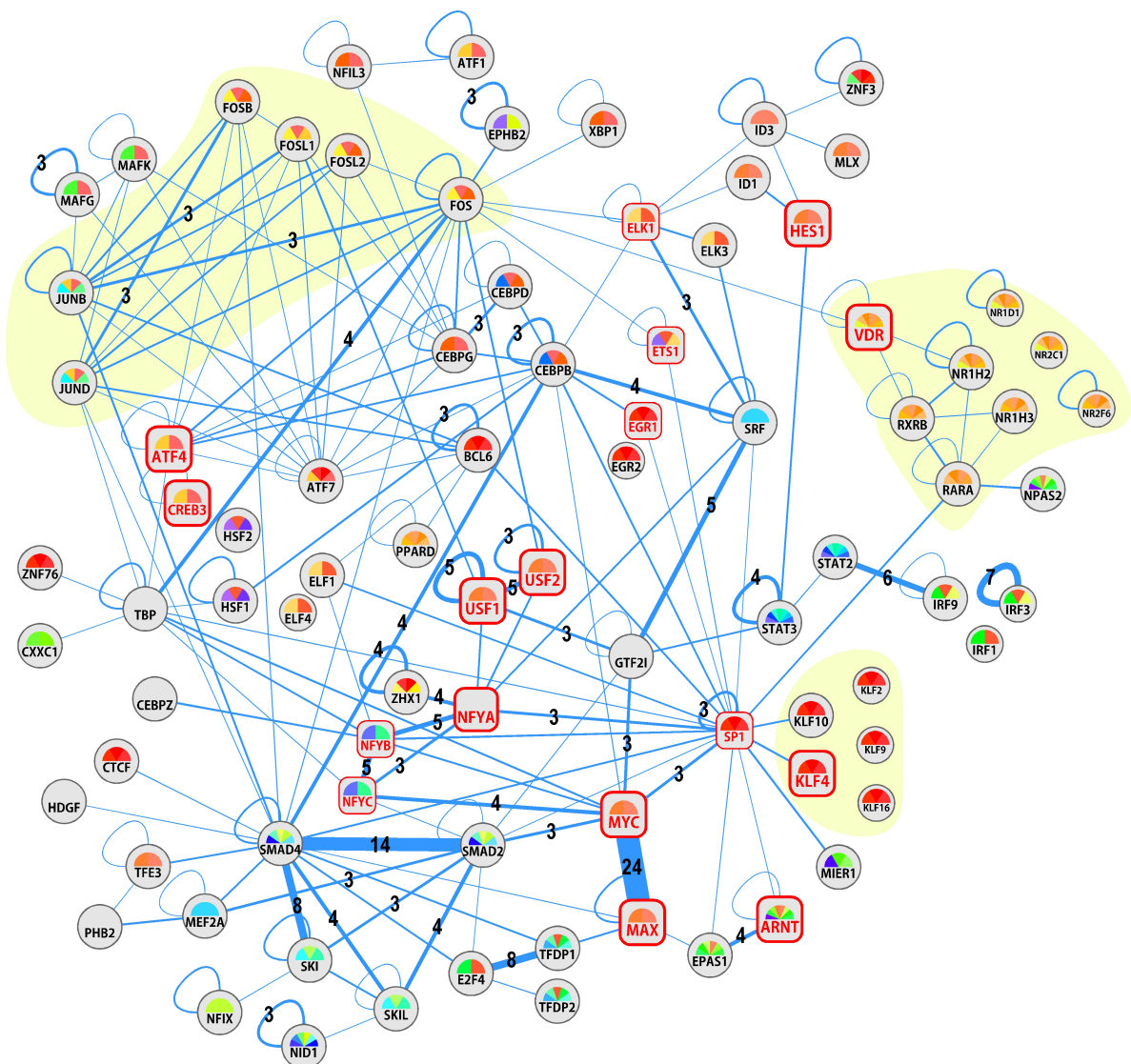


Figure 4.12 Red de interacción de los 135 TFs que se encuentran en las MSC.

- Un análisis de expresión diferencial de los datos ARNseq permitió definir 203 genes cambiantes entre PL- y BM-MSCs: 125 resultaron sobre-expresados en BM-MSC y 78 en PL-MSC. Los genes de BM-MSC fueron enriquecidos en funciones tales como: la biogénesis de hueso, la formación de hueso, morfogénesis de vasos sanguíneos, organización de la matriz extracelular y respuesta inflamatoria. Por otro lado, los genes PL-MSC apuntaban a términos más vinculados a procesos de células madre, tales como: la morfogénesis embrionaria, la activación del ciclo celular y la regulación negativa de la muerte celular;

3. Relaciones entre las diferentes células estromales estudiadas a través de *microarrays* de expresión

- Un meta-análisis de varias colecciones de datos públicos nos permitió mostrar la distancia transcriptómica que separa las células estromales de tipos celulares relacionadas por linaje y de otros tipos no relacionados. 264 *microarrays* de expresión de 18 series de datos fueron seleccionados a través de un exigente proceso de filtro de más de 500 muestras. Varios pasos de normalización de los datos incluyeron las siguientes pruebas con combinaciones alternativas de algoritmos de normalización: (A) RMA; (B) RMA + LOESS; (C) RMA + non-parametric COMBAT; (D) RMA + parametric COMBAT; (E) RMA + LOESS + non-parametric COMBAT; (F) RMA + LOESS + parametric COMBAT; (G) frozenRMA.

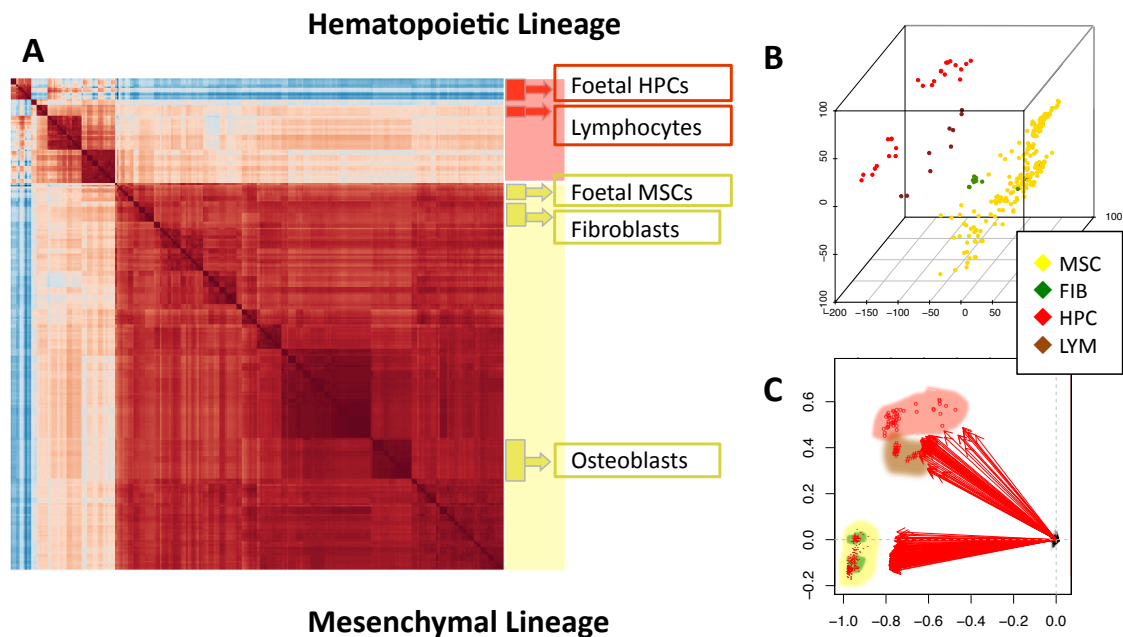


Figure 4.13 (A) Un mapa de calor de 264 *microarrays* que incluye mesenquimales y células de linaje hematopoyético. La agrupación de las muestras se basa en distancias euclidianas. **En el margen derecho:** las muestras de linaje hematopoyético se han sombreado en rojo; y muestras de linajes mesenquimales en amarillo. La escala de colores representa la distancia más corta o más lejana, del rojo oscuro al azul oscuro. (B) Gráfico 3D de PCA realizado para las muestras de células del estroma y células hematopoyéticas. Cada punto representa un *microarray* del repertorio de 264. (C) Biplot de PCA realizado sobre los genes de los mismos datos.

- Paralelamente, se produjo también un conjunto de datos propios a partir de microrarrays *Affymetrix Human Exon 1.0*. Las poblaciones cultivadas de MSC de tejido adiposo (AD), médula ósea (BM) y placenta (PL) han sido sometidas a ensayos de *microarrays* de expresión genómica, junto con los fibroblastos de la piel (FIB) y HSPCs como tipos de células contraste. Sobre estos datos se realizó un análisis de agrupamiento semi-supervisado considerando sólo 358 genes marcadores (CDs de citometría canónicos), que permitieron separar perfectamente los tipos celulares entre sí.

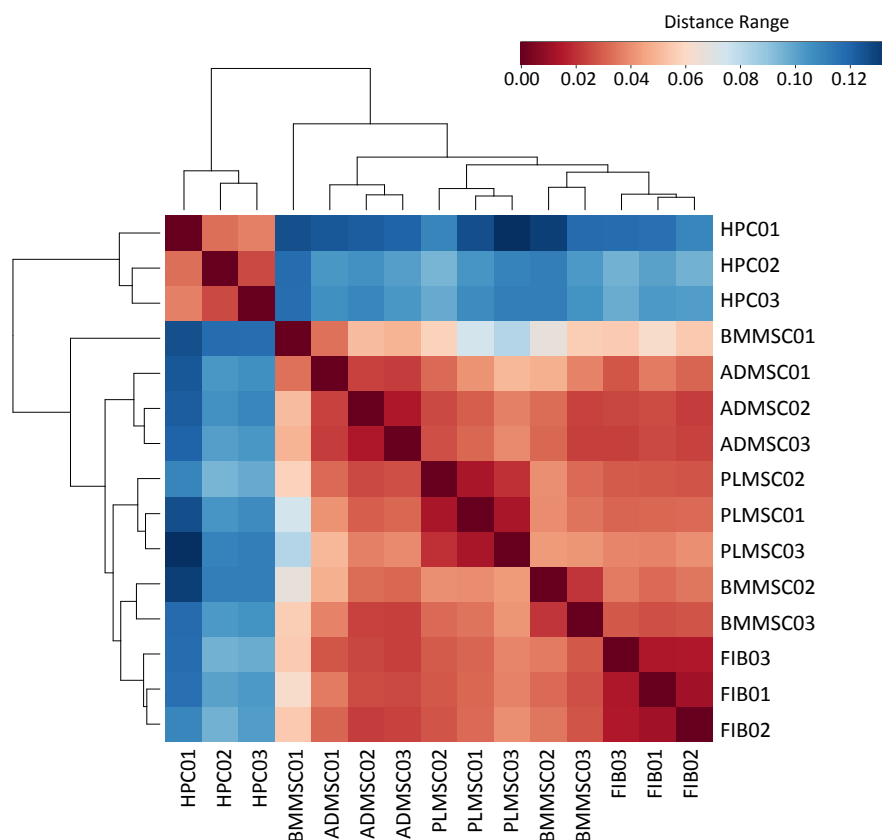


Figure 4.14 Mapa de color de las distancias de correlación de Pearson, muestra a muestra, de la expresión génica de los 15 *microarrays* producidos *de novo*.

- Análisis de expresión diferencial y genes específicos de tejido obtenidos a partir de los dos contextos de datos genómicos.
- En los resultados del meta-análisis, cuando cruzamos las listas de genes sobre-expresados en BM-MS de ambos contrastes (frente a Los FIBs y frente HSPCs), un gen marcador CD y dos factores de transcripción se mantuvieron en la intersección: CD146, ID3 y PAWR. Seguidamente se realizó un análisis dirigido al significado

funcional de los genes expresados diferencialmente. Cuando se compara con HSPCs, funciones como el desarrollo de los vasos sanguíneos, la localización de microvesículas, la migración celular, o la regulación de la muerte celular, se enriquecen en los genes de BM-MSC. Lateralmente, los procesos de hematopoyesis se enriquecen en los genes HSPC. Significativamente, los genes más altos en BM-MSC frente a fibroblastos aparecen anotados a las tareas de desarrollo de hueso (por ejemplo osificación), así como la organización de fibras esqueléticas (por ejemplo, disco Z, miofibrilla, o citoesqueleto de actina). Entre los genes de fibroblastos predomina la organización de la matriz extracelular en general.

Tabla: Meta-análisis de expresión diferencial después del protocolo de re-muestreo. Esta tabla resume los genes DE obtenidos de cada contraste siguiendo una estrategia de re-muestreo

Contrast	CELL TYPE I	CELL TYPE II	Median of nDEgenes (in 100 runs)	nIter per run	Selected DE-genes	Up-Reg in CELL TYPE I	Up-Reg in CELL TYPE II
1	BMMSC	HPC	4878	150	4748	2124	2624
2	BMMSC	FIB	426	50	401	207	194
3	BMMSC	OSTB	3	20	--	--	--
4A	BMMSC	dOST	38	100	--	--	--
4B	dOST	OSTB	1238	100	--	--	--
4C	OSTB	stOST	0	100	--	--	--

- Análisis de expresión diferencial han sido también implementados sobre nuestro propio conjunto de datos. Gracias a ellos, se han podido definir y caracterizar varias firmas génicas específicas de cada MSC de tejido estudiada.
- En relación a lo anterior, también se han definido los patrones de expresión del citoquinoma, asociados a cada uno de estos tejidos de origen de las MSC.

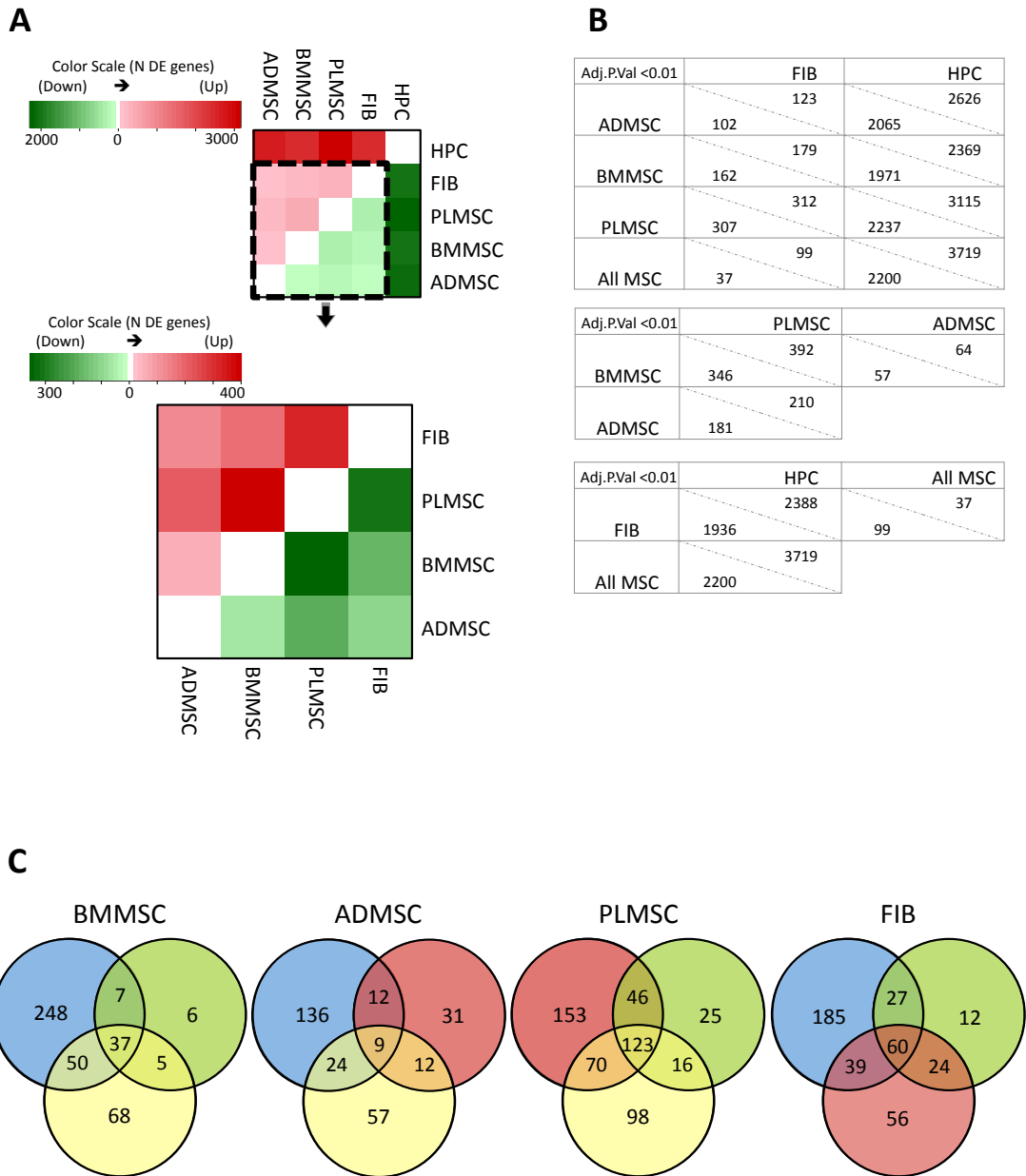


Figure 4.15 Análisis de expresión diferencial sobre nuestro conjunto de datos de microarrays. **(A)** Mapa de calor de distancias, medida como el número de genes expresados diferencialmente entre cada tipo de célula. Escala roja representa los genes sobre-expresados en el tipo celular marcado en el panel de la derecha; simétricamente, las intensidades verdes representan el número de genes infra-expresados. Cuanto más intenso es el color, mayor es la distancia entre los tipos de células implicados. **(B)** Recuentos de números de genes expresados diferencialmente por contraste. **(C)** Genes sobre-expresados en cada contraste: BM- (rojo), PL- (azul), AD- (verde) y FIB (amarillo)

Tabla: Firmas génicas específicas de tejido .

Cell Type	Specific genes
BMMSC	KLHL3, MAOB, TM4SF20, AP001422.1, C20orf197, ANK3, RP11-145A3.3, NOTCH3 , FBXO16, EPHX1, HEY2, DLX5, JAG1, ZNF423, IGF2, C20orf103, EYA2 , KCNMB1, ENTPD1, ITGA7, LBP, AGT , KIAA1217, CHRDL1, IGFBP2, SLPI, CXCL16 , EYA4 , NPR3, SFTA1P, LEPR , EYA1 , C5orf23, FMO3, ACAN, VCAM1 (CD106)
ADMSC	AP000843.1, PDLIM3 , PPAPDC1A, Z69713.1, MFAP3L, AC104654.2, PCDH9, MATN3 , AP000695.2 FABP4, HTR2B, TFPI2, SYTL5 , TNFSF4 , ALDH1A1, F2RL1, TNFSF18 , HOXD10, TRHDE, PCDH10, PLAT, CAPN6, SYT14 , AC090625.2, NCAM1 , SCN9A, AL121578.3, ANO4, PMAIP1, GBP4, MCTP1, PAMR1, ST6GALNAC5, SERPINB2, RARB, C9orf167, GPR126, ADAM23 , SGIP1, LPCAT2, AC026250.2, ZNF804A, REN, HTR1D, AC003092.2, LAMA1, C3orf72, C12orf59, GAS2L3, INA, TMEM154, PLBD1, FJX1, PLCXD3, AMIGO2 , IFIT3, IFIT2, LCP1, KRT19, RP4-706A16.2, PDCD1LG2, CARD16, SNCA, NR2F1, NEFM, RASGRP1, SLITRK4, TRPC4, NETO2, RP11-251J8.1, HMGA2, GPR37, FAM105A, RPS6KA6, SOX5, SSTR1, SYT16 , LIMS1, RP11-87H9.2, BRIP1, CAMK1G, SOCS1, E2F7, DGKH, HSD17B2, C10orf57, AC069155.1, CKS2, FAM60A, MUM1L1, S1PR3, SLC16A4, TNFSF15 , STRA6, SNCAIP, CADPS2, PDGFC , NID1, RP11-332P22.1, MAPK8, SORBS1, ALDH1A2, SYPL2, CYTSB, ATRNL1, SLC1A2, CACNA1H, CTSH, CNTN4, RP11-548O1.1, KCNA3, ETV4, MAP3K5, GCLC, RP11-117P22.1, NNAT, AC012409.1, ADRA1D, C16orf52, FAM155A, SLC4A8, MSI2, CD226 , AR, SNORD27, TMEM51, GLRX3, LRRC1, ZNF516, DTNA
PLMSC	KIAA0802, MASP1, MBNL3, AK5, AIM1, ADAM33 , CCDC102B, C9orf21, PRLR, GTSF1, AC112217.2, PAX3, OSR2, PHACTR3, PLEKHG1, TWIST2 , CLIC2, OMD, SVIL, OR1Q1, THBS4, DNM1, PDGFRL, AC008440.2, GLI3, GLDN, APOD, TWIST1 , GRPR, TFAP2C, SECTM1, SIPA1L2, RCAN2, ROBO2, RXFP1, RP11-392O17.1, IGSF10, THRB, KIT , SEMA3D, MMP27 , SLC9A9, BMPER, STK32B, CH25H, OR1H1P, TGFB3, CCBE1, AQP9, ADH1B, CDON, PLXNC1, CYP7B1, TDO2, LRRC15, CCRL1, APCDD1, DKK2, COLEC12
FIB	

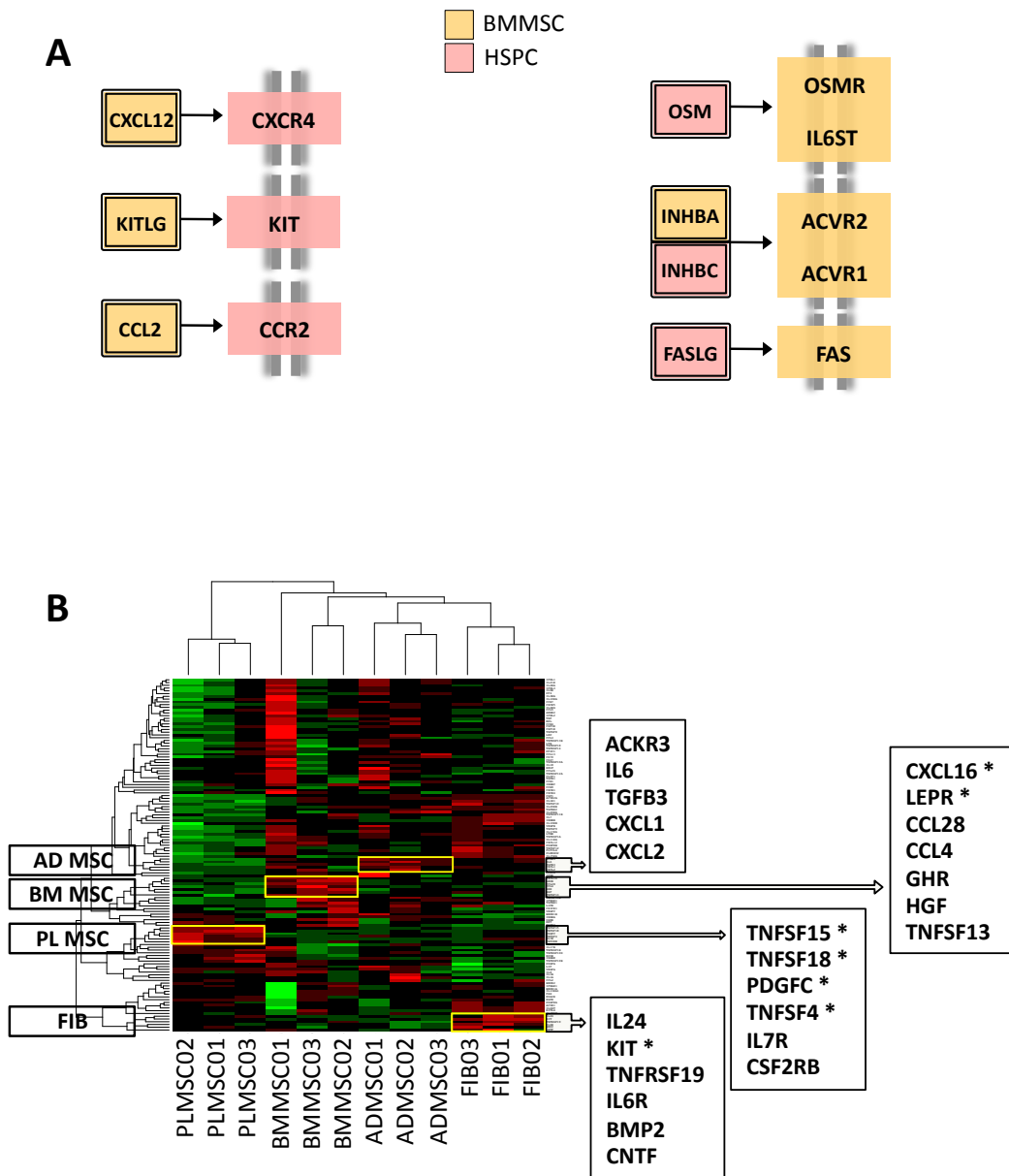


Figure 4.16 (A) Interacciones de citoquinas. **(B)** Patrones de expresión de citoquinas según los tejidos de origen de las MSC.

4. Cruce de resultados de las tres aproximaciones genómicas desarrolladas.

- A partir de los datos de secuenciación, del meta-análisis y de nuestro propio conjunto de *microarrays*, hemos definido una firma que denominamos: el núcleo de expresión de linaje mesenquimal. Está compuesto por 489 genes.

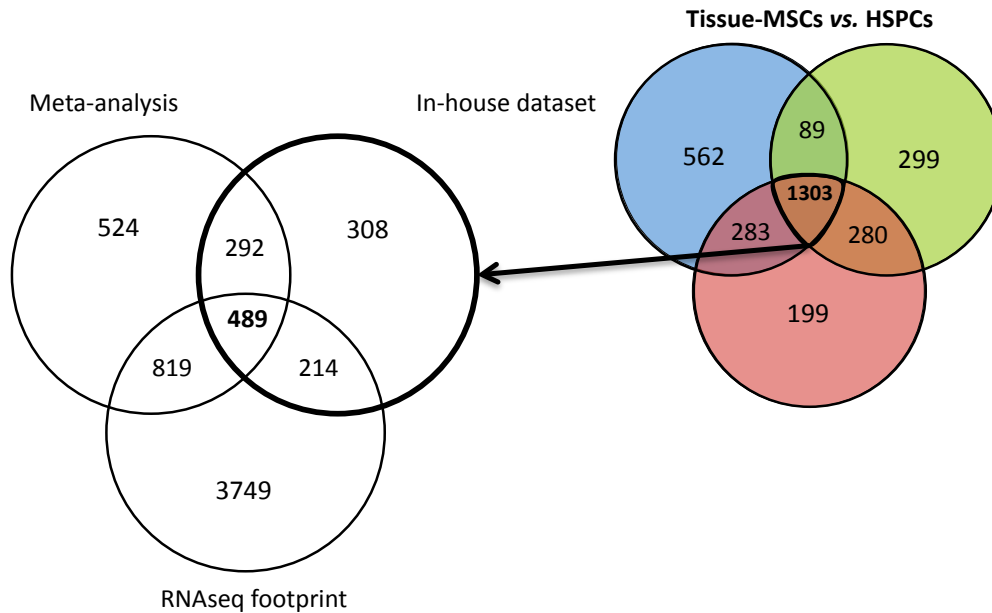


Figure 4.17: La firma de linaje mesenquimal de 489 genes. Diagramas de Venn que muestran el flujo transversal de los genes expresados diferencialmente de los tres contextos genómicos.

- El análisis funcional de ésta firma ha permitido generar una red basada en las relaciones funcionales entre los 489 genes.

Tabla: Enriquecimiento y análisis de agrupamiento de anotaciones funcionales de la firma de linaje mesenquimal de 489 genes (realizado con la herramienta web *GeneTerm Linker*)

MG	Silhouette Width	Genes	pValue	Gene symbols	Terms
1	0,477	20	4,51E-11	DDR2,ILK,ACVR1,BMPR1A,BMPR2,PDLIM7,CAV1,MYOF,CDH13,CTSB,PTRF,S MURF2,CDKN2B,PPP2CB,THBS2,PDGFRB,FSTL1,EXT2,COL5A2,RAB2A	BMP signaling pathway (BP), Caveola (CC), Mesoderm formation (BP), Positive regulation of osteoblast differentiation (BP), SMAD binding (MF)

2	0,443	24	6,22E-13	PGM3,SAR1B,SEC24A,GFPT1,UAP1,SEC23A,SEC31A,STT3A,GANAB,TUSC3,MCFD2,YKT6,TMED10,RAB2A,GJA1,MBTPS1,CKAP4,ARCN1,MGAT4B,GGCX,P4HB,SEL1L,PSMB5,CAPN2	Antigen processing and presentation of peptide antigen via MHC class I (BP), COPII vesicle coating (BP), ER to Golgi transport vesicle membrane (CC), ER to Golgi vesicle-mediated transport (BP), N-Glycan biosynthesis
3	0,319	29	5,63E-20	LAMB1,COL6A3,COL4A1,LAMB2,COL5A2,TNC,VEGFC,VCL,LAMC1,ACTN1,THBS2,COL4A2,LAMA4,COL5A1,COL3A1,ITGA3,ITGA5,SDC4,ANXA2,LOX,P4HA1,COL12A1,PLOD3,SERPINH1,PXDN,BGN,MAP2K1,CCND1,MMP2	Collagen fibril organization (BP), Extracellular matrix structural constituent (MF)
4	0,272	24	1,21E-13	ILK,MYH9,TPM4,VCL,TPM1,FERMT2,ACTN1,LIMA1,PDLIM7,TPM2,CALD1,DSN,MAP2K1,SGCB,TNFRSF12A,CERCAM,GJA1,PTPRM,DLG5,ASPH,MYOF,FZD2,GTF2IRD1,TNFAIP1	Cell-cell adherens junction (CC), Stress fiber (CC), Structural constituent of muscle (MF)
5	0,245	26	8,36E-15	FGFR1,ITGA3,MAP2K1,MYH9,ITGA5,RAS,VCL,RRAS2,ACTN1,PDGFRB,MRAS,FGF7,SAR1B,RHOC,RAB2A,RND3,RPS6KA2,ARHGEF12,RAB31,GNA11,CASK,ARF4,TUBB2A,RRAGC,TUBG1,EHD2	Ras protein signal transduction (BP)
6	0,230	27	1,44E-15	IGF2R,SORT1,SCARB2, CD63 ,NPC1,HEXA,CLTB,LAPTM4A,LAMP1,HEXB,MCOLN1,GJA1,CAV1,SEC31A,CLTC,CLTA,YKT6,STS,TMEM55A,ANKFY1,WIP1,APPL2,CTSA,MYOF,BNIP3,EHD2,PCYOX1	Late endosome membrane (CC), Nuclear envelope (CC)
7	0,214	26	3,62E-12	CLIC4,SORT1,NRP1,TNFRSF12A,ITGA3,APP,ADAM9,PTPRK,SRPX,LPAR1,STX2,ILK,ITGA5,PTPRM,VCL,HEG1,TMEM47,SDC4,CDC42BPB,IGF2R,BMPR2,LGALS1,BMPR1A,PTPN21,DUSP14,TNS1	Peptidyl-O-phospho-tyrosine phosphatase, Protein-tyrosine/Dual-specificity phosphatase, Protein tyrosine phosphatase activity (MF), Protein-tyrosine phosphatase , catalytic
8	0,202	28	6,06E-17	NRP1,FGFR1,PTPRK,BAMBI,PDGFRB,TNFAIP1,CDC42BPB,MAP2K1,MYH9,ACTN1,VCL,CCND1, CREB3 ,SORT1,MYO6,ARHGEF12,ITGA5,RRAS,ALDH7A1, SNAI2 ,WDR1,GPNMB,BECN1,RPS6KA2,OPTN,ITGA3,FGF7,MAP4K4	Prostate cancer
9	0,202	39	7,30E-21	MARCKS,CALD1,MYH9,MYO1E,CASK,CTTN,PDLIM7,MYO6,ITGA5,TNFRSF12A,S100A11,ACTN1,LIMA1,ANTXR1,LASP1,CORO1C,KIF1C,DYNLT3,DYNC1LI2,CNN3,MYO1D,CDC42BPB,FGF7,EPSS8,DSP,SRPX2,DLG5,PPP2CB,CDC42EP3,CCL2,FERMT2,BAMBI, THY1 ,APP,SYNC,PDZRN3,DSTN,BECN1,STX2	Actin cytoskeleton reorganization (BP), Actin filament binding (MF), Cytokinesis (BP), Motor activity (MF), Myosin complex (CC), Myosin head, motor domain, Neuromuscular junction (CC), Regulation of cell shape (BP), Ruffle (CC)

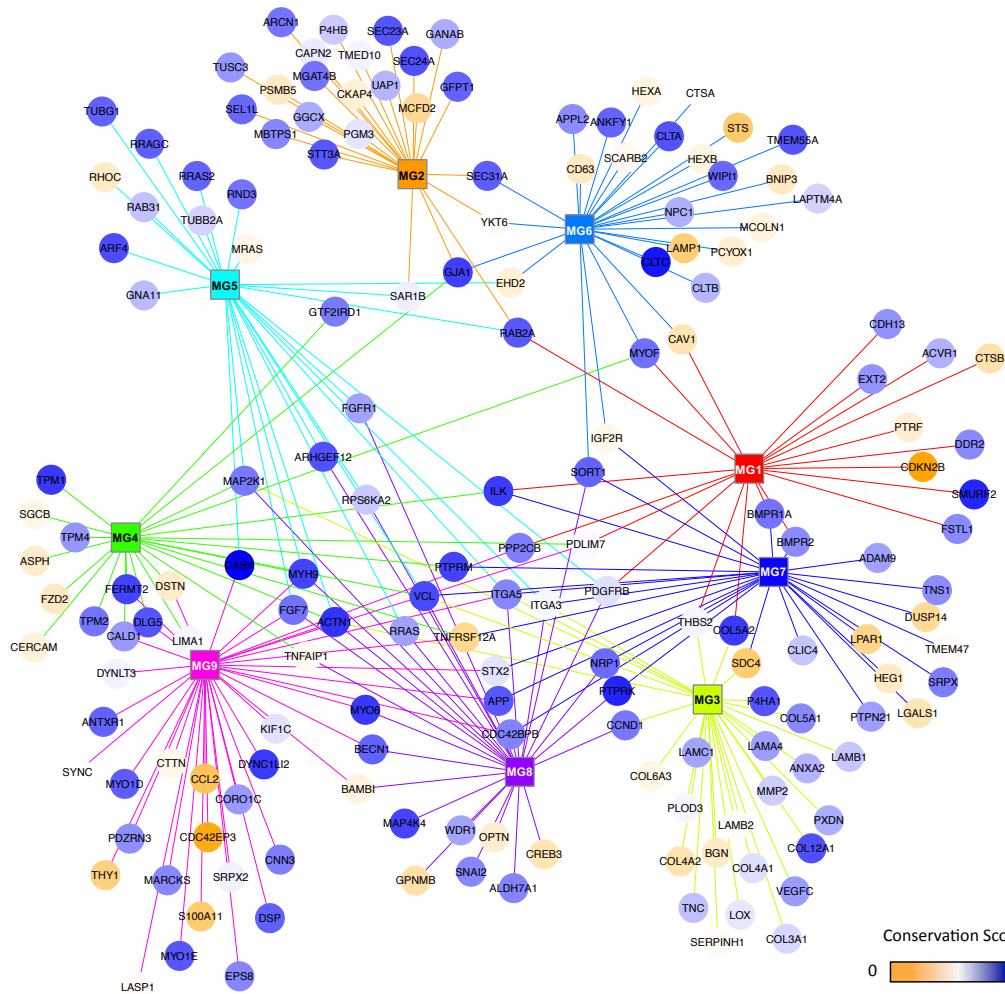


Figure 4.18 Red funcional basada en el análisis de enriquecimiento y agrupación funcional de la firma de 489 genes mesenquimales.

- La comparación de nuestros resultados con algunas otras listas ya caracterizadas previamente ha permitido determinar algunos genes repetitivos, que son de suma relevancia para la biología de las MSC.

Conclusiones

Esta sección resume las conclusiones generales de esta tesis:

1. Hemos validado la presencia de una población de MSCs dentro de las secciones de placenta coriónica humana (PL-MS), pudiendo establecer un protocolo para su aislamiento y selección en cultivos adherentes. Los cultivos primarios obtenidos superaron los criterios identificativos marcados por la ISCT, y mostraron la proximidad citológica y transcriptómica esperada respecto a las MSC mejor definidas, las de médula ósea (BM-MS).
2. Utilizando una técnica de secuenciación masiva de ARN (*RNA-Seq*) sobre poblaciones de BM-MS y PL-MS hemos obtenido un retrato completo del transcriptoma de estas células. La huella transcriptómica definida abarca 5.721 genes codificantes de proteínas activamente expresados o “encendidos” que engloban los todos los componentes biomoleculares, desde genes constitutivos genéricos, hasta factores de transcripción altamente específicos de las MSC humanas.
3. El planteamiento de un meta-análisis, capaz de integrar datos públicos de microarrays de expresión, con un diseño biológico coherente y una metodología bioinformática robusta, nos ha permitido establecer las distancias transcriptómicas entre las MSCs y otras poblaciones celulares relacionadas. Los resultados obtenidos, concuerdan con el conocimiento existente, y además proporcionan una amplia y precisa caracterización de las MSC de médula ósea basada en la expresión de sus genes.
4. El análisis y la integración bioinformática de los diferentes datos ómicos nos ha permitido definir una firma transcriptómica de 489 genes que recoge los genes centrales del linaje mesenquimal de poblaciones celulares con capacidad multipotente.

5. La producción propia y controlada de un conjunto de datos de microarrays de expresión de MSCs aisladas a partir de varios tejidos, nos permitió definir la variabilidad entre tejidos de origen. Así, hemos definido las firmas de expresión génica específicas para: médula ósea, tejido adiposo y placenta, así como para fibroblastos dérmicos.

6. Los fibroblastos de piel son células estromales consideradas una población de células diferenciadas útiles en estudios comparativos sobre el fenotipo asociado a la capacidad *stem* de las MSCs. Hemos observado que la definición de identidad actual de las MSC no permite una clara distinción de los fibroblastos. Sin embargo, hemos demostrado que existen diferencias de expresión, cuando se observan los transcriptomas completos.

7. Hemos podido encontrar evidencias significativas de la diferencia de expresión del gen CD146 (MCAM) entre BM-MSc y fibroblastos. Este gen ya ha sido postulado como marcador de multipotencia y auto-renovación de células progenitoras del estroma en la médula ósea humana. Este marcador también lo hemos encontrado expresado por secuenciación en MSCs de placenta. La expresión de otros marcadores propuestos, como Nestina, también se han probado en BM- y PL-MSc. Sin embargo no se detectó expresión para CD271.