

# Investigación española en células madre (1997-2010).

## Tema 2 Localización y evolución de las principales líneas de investigación a través de los KeyWords Plus

**Gisela Cantos-Mateos**

Consejo Superior de Investigaciones Científicas / Instituto de Políticas y Bienes Públicos (CSIC-IPP) / SCImago Research Group (España)  
gisela.cantos@csic.es

**Maria Angeles Zulueta**

Universidad de Alcalá, Facultad de Documentación / SCImago Research Group (España)  
ma.zulueta@uah.es

**Benjamin Vargas-Quesada**

Universidad de Granada, Facultad de Comunicación y Documentación / SCImago Research Group (España)  
benjamin@ugr.es

**Zaida Chinchilla-Rodríguez**

Consejo Superior de Investigaciones Científicas / Instituto de Políticas y Bienes Públicos (CSIC-IPP) / SCImago Research Group (España)  
zaida.chinchilla@csic.es

608

### RESUMEN

- **Objetivos:** analizar y representar la evolución y la dinámica de las principales líneas de investigación en la producción científica española sobre células madre entre los años 1997 y 2010.
- **Material y Métodos:** los documentos recuperados proceden de la base de datos *Science Citation Index* (SCI). Las unidades de análisis han sido los descriptores *KeyWords Plus* (KW+). Las herramientas utilizadas para la visualización han sido el software *Pajek*, en combinación con el algoritmo *PathfinderNetwork* (PfnET) y el software *VOSviewer*. Metodológicamente, el estudio analiza el periodo completo (1997-2007) y la evolución de las líneas de investigación en tres subperiodos: 1997-2001, 2002-2006, 2007-2010.
- **Resultados y Discusión:** Los resultados del periodo completo permiten localizar hasta seis líneas de investigación. En el primer periodo, se observa una fuerte presencia de los descriptores que representan documentos relacionados con la hematología y oncología, en cambio, el resto de líneas no comienzan a detectarse de manera más clara hasta el segundo periodo. Aquí, es posible localizar agrupaciones de descriptores relacionados con la investigación sobre células madre hematopoyéticas, otros referidos a los estudios sobre la generación, proliferación y diferenciación de células madre y otra agrupación emergente relacionada con las células progenitoras neurales. Finalmente, los resultados muestran que el peso de estas últimas agrupaciones tienen una presencia mucho más evidente en el último periodo.
- **Conclusión:** La visualización de las relaciones que se establecen entre los KW+ ha permitido obtener dos imágenes complementarias de la situación y evolución de la investigación española sobre células madre. La metodología ha permitido identificar tanto líneas de investigación claramente consolidadas como las emergentes y se intuye la capacidad de predicción del desarrollo de un dominio temático a lo largo del tiempo. Los programas de visualización se complementan muy bien y coinciden tanto en las líneas de investigación como en la localización de los descriptores más influyentes de la red.

**Palabras clave:** Células Madre, Análisis de Co-Words, Visualización de la Información, Análisis de Dominios Científicos

## ABSTRACT

- **Objective:** To render and analyze the evolution and dynamics of the main research lines in the Spanish scientific output on stem cells between 1997 and 2010.
- **Material and Methods:** We retrieved the documents from the Science Citation Index (SCI). The units of analysis were the descriptors KeyWords Plus (KW +). The tools used for visualization have been software Pajek combined with Pathfinder Network (PFnet), and VOSviewer software. This study analyzes the period includes between 1997-2007, and the evolution of the research into three sub-periods: 1997-2001, 2002-2006 and 2007-2010.
- **Results and Discussion:** The results of the full period locate up to six main research lines. In the first period, there is a strong presence of the descriptors that represent documents related to hematology and oncology. However, the other research lines do not begin to be clearly detected until the second period. Here, we can locate clusters of descriptors related to hematopoietic stem cell research, others related with the generation, proliferation and differentiation of stem cells and finally, alternate associated to an emerging cluster of neural progenitor cells. The results show that the weight of these latter groups make them more evident in the last period.
- **Conclusion:** The visualization of the relationships between the KW + has yielded two complementary images of the situation and evolution of Spanish research on stem cells. The methodology has identified areas of research both consolidated and emerging, intuiting the development of a thematic domain over time. The visualization software complement each other quite well, matching in the identification of the main research lines and in the location of the most influential descriptors.

**Keywords:** Stem Cells, Co-Words Analysis, Information Visualization, Scientific Field's Analysis

## 1. INTRODUCCIÓN

La investigación española con células madre es un tema biomédico de gran importancia y repercusión científica por las expectativas que genera. Desde que en 1949, el científico J. Hammond (1949) descubriese el método para mantener los embriones de ratón en cultivo *in vitro*, la investigación con células madre se ha desarrollado gradualmente ofreciendo en los últimos años perspectivas muy esperanzadoras para el tratamiento de enfermedades, hasta ahora, incurables. Actualmente, la dirección de la investigación está orientada, fundamentalmente, al desarrollo de nuevas terapias para enfermedades hematológicas, cardiovasculares, neurodegenerativas, genéticas, cáncer, diabetes, etc. (Martínez Serrano y Bjorklund, 1996; Bishop et al., 2002; Cao et al., 2002; Di Giorgio et al., 2007). Por su naturaleza, la investigación con células madre trasciende inevitablemente a otros campos de carácter tan diverso como el político, ético, cultural, jurídico, etc., situándose en el escenario de la controversia social.

Por otro lado, resulta cada vez más evidente, la utilidad que tienen los estudios bibliométricos para reflejar el estado en el que se encuentra la investigación en un momento determinado, predecir tendencias y facilitar la toma de decisiones en materia de política científica. Esto impulsa el desarrollo y perfeccionamiento de nuevas técnicas y herramientas bibliométricas que tratan de mejorar la comprensión y análisis de los dominios científicos. Una muestra es la incorporación de las redes sociales (Wasserman y Faust, 1998) como técnica de análisis y visualización de dominios científicos (Boyack et al., 2009; Leydesdorff y Rafols, 2009; Vargas-Quesada et al., 2010; Rafols et al., 2010).

Existe una gran diversidad de métodos y técnicas a la hora de delimitar y visualizar un campo científico con el fin de detectar líneas de investigación. Hay estudios que utilizan la

cocitación de autores (ACA) (White y Griffith, 1981; White y McCain, 1998; Small and Upham, 2009; Chen et al., 2010; Zhao y Strotmann, 2011), el análisis de emparejamiento bibliográfico (Kessler, 1963; Boyack y Klavans, 2010; Chen et al., 2011; Glänzel y Thijs, 2011), estudios de co-words (Callon et al., 1983; Rip y Courtial, 1984, Cambrosio et al., 1993; Van Raan y Tijssen, 1993; Ding et al., 2001; Lee, 2008; Leydesdorff y Welbers, 2011; Zulueta et al., 2011) o los estudios híbridos en los que se combinan tanto ACA como co-words (Braam et al., 1991a, b; Zitt y Bassecoulard, 1996; Zitt et al., 2011). Según Zitt y colaboradores (2011), las palabras, a diferencia de las citas, parecen adaptarse mejor a la «ciencia viva» porque son universales. Para estos autores, las unidades lingüísticas tienen la capacidad de reflejar mejor la incidencia de los contextos científicos, sociales y políticos, propios de los dominios más controvertidos y áreas emergentes. Además, el análisis de co-words es considerado como una técnica de análisis de contenido muy eficaz en el mapeo de la fuerza de asociación entre los ítems procedentes de datos textuales (Cobo et al., 2011).

En vista de las ventajas e inconvenientes que condicionan la elección de una metodología sobre otra, para el presente estudio se ha optado por el análisis de co-words (Zitt et al., 2011), continuando así estudios precedentes sobre la investigación española sobre células madres. En un estudio anterior (Cantos-Mateos et al., 2012), se analizó la producción científica desarrollada entre 1997 y 2007 sobre esta temática. Se analizó mediante indicadores bibliométricos la producción científica relativa al dominio y se generó una representación topográfica de los descriptores *KeyWords Plus* más frecuentes. Este trabajo puede considerarse una continuidad, ya que, se detectaron las principales líneas de investigación del período comprendido entre 1997 y 2007. En este caso particular, el objetivo fundamental se centra en la localización y representación de las principales líneas de investigación que tienen lugar en las distintas fases o periodos temporales en los que se ha dividido el estudio con el fin de apreciar la evolución y la dinámica, tanto de las líneas de investigación consolidadas como de los emergentes.

## 2. MATERIAL Y MÉTODOS

### 2.1. FUENTES, UNIDADES DE ANÁLISIS Y UNIDADES DE MEDIDA

La base de datos utilizada fue Science Citation Index (SCI). Entre las razones para su elección destaca que el campo *Address* está normalizado, característica fundamental para estudios que pretenden centrarse en un dominio geográfico determinado. Además, emplea los descriptores *KeyWords Plus* (KW+) para indizar los documentos. Estos descriptores son obtenidos mediante un algoritmo de búsqueda que extrae las palabras clave de los títulos de las referencias contenidas en los documentos (Garfield, 1990; Garfield y Sher, 1993).

La elección de los KW+ como unidades de análisis se debe a su elevada representatividad, ya que están presentes en el 78,25% de los documentos (Tabla 2) Los *Keywords Author* se descartaron por su baja representatividad, ya sólo están presentes en el 52,96% (Cantos-Mateos, et al, 2013). Además, al tratarse de términos procedentes de los títulos de las referencias, son los propios autores, los que seleccionan los términos esenciales para expresar el contenido más importante de sus trabajos de investigación. Esta característica favorece la aparición de una terminología más actualizada y específica, en comparación con otros tipos de descriptores como son los procedentes de un vocabulario controlado (Braam et al., 1991; Cantos-Mateos, et al. 2013). Estas cualidades, pueden contribuir mejor a reflejar los aspectos temáticos más dinámicos de un dominio. Entre los inconvenientes hay que destacar que presentan una amplia dispersión terminológica. La falta de normalización de los

términos y de una selección estandarizada de los mismos contribuye a que se generen muchos descriptores para expresar un mismo concepto o idea (Quin, 2000). Para paliar en cierto modo esta limitación, se ha realizado una normalización unificando criterios gramaticales (plurales y singulares) y ortográficos (guiones, mayúsculas y minúsculas).

La estrategia de búsqueda empleada para la recuperación de los documentos consistió en utilizar los términos *stem\* cell\** en el campo *Topic*, *Spain* en el campo *Address* y limitada al periodo temporal de 1997-2010. El criterio de selección de la ventana temporal está determinado por la disponibilidad de los *Journal Citation Report* a partir del año 1997, con el objetivo de analizar el impacto de las principales revistas del área. De cara al estudio evolutivo, además de analizar el período completo en su conjunto (1997-2010), éste se ha subdividido en otros tres: 1997-2001; 2002-2006 y 2007-2010. Esta división temporal estuvo motivada precisamente por el volumen documental, teniendo en cuenta el crecimiento exponencial de la producción científica (De Solla Price, 1963). Quizá, el haber considerado más subdivisiones hubiera fragmentado excesivamente la información impidiendo ver diferencias importantes entre un periodo y el siguiente. Y por el contrario, realizar menos divisiones hubiera agrupado demasiada información ofreciendo una imagen más estática y corriendo el riesgo de detectar sólo líneas muy consolidadas. El número de años por período es de 5, a excepción de último que contiene 4. Al tratarse de un período de 14 años, se ha preferido que los dos primeros tramos temporales sean más amplios ya que, el volumen de producción de los primeros años es mucho menor en comparación con los últimos.

Para la identificación y visualización de las principales líneas de investigación, se han calculado las relaciones que se producen entre los KW+ a partir de su aparición conjunta en los documentos, lo que se conoce como *Co-Words Analysis*. Dicho análisis consiste, en generar una matriz cuadrada de  $N \times N$  elementos, donde  $N$  es el descriptor KW+ a representar, a partir de las veces que ocurre en los documentos. El resultado es una matriz de co-ocurrencias que refleja el número de veces que un par de descriptores aparecen conjuntamente en dos documentos.

## 2.2. VISUALIZACIÓN DE LA INFORMACIÓN

Para la visualización de la estructura relacional entre los KW+, se han utilizado dos software libres especializados en la generación de redes. Por un lado, Pajek (Batagelj and Mrvar, 2010) y en concreto, el algoritmo *Kamada-Kawai*, (Kamada y Kawai, 1989). Aunque Pajek permite representar grandes redes, se decidió simplificarlas con el fin de que las visualizaciones fueran los más claras posibles. Por un lado, se decidió representar únicamente los descriptores más frecuentes. Tras varios ensayos con distintos umbrales de frecuencias, se observó que las visualizaciones más claras eran las compuestas en torno a 100 nodos. Para la reducción del espacio dimensional se aplicó el algoritmo de poda *Pathfinder* (Moya-Aneón et al., 2007a, b), basado en el principio de desigualdad del triángulo bajo los parámetros  $r = \infty$  y  $q = n - 1$ . Éste ayuda a preservar y poner de relieve las relaciones más destacadas entre los descriptores (White, 2003; Vargas-Quesada et al., 2008).

Para la representación del periodo completo, se utilizó el software VOSviewer v.1.5.4 (Van Eck y Waltman, 2010), que se presenta como una alternativa a las técnicas tradicionales de representación multidimensional y visualización de redes. VOSviewer combina técnicas de visualización y clustering, lo que favorece el análisis. En este estudio se han utilizado los mapas etiquetados con nombres, donde cada descriptor está representado por una etiqueta y cuyo tamaño es proporcional a su peso, junto con los denominados «mapas de calor» (o mapas de densidad). Estos últimos se caracterizan porque cada nodo se representa con un color que va del rojo al azul, reflejando la densidad de las relaciones entre los

descriptores. Cuanto mayor es la densidad (co-ocurrencia entre los descriptores), más se aproximarán a la tonalidad roja y cuando hay más dispersión por la menor co-ocurrencia, se aproximan más al color azul.

### 3. RESULTADOS Y DISCUSIÓN

Se recuperaron un total de 4.148 documentos para el período 1997-2010. La tabla 1 muestra el volumen de producción por tramos temporales. 724 documentos en el primero, 1.725 documentos en el segundo (2002-2006) por y el tercero (2007-2010) por 2.174 documentos.

**TABLA 1. Número de documentos por años.**

Período	Primer Período					Segundo Período					Tercer Período			
Años	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Nº docs.	94	142	176	149	163	167	165	296	271	351	395	494	608	677
Total docs.	724					1.250					2.174			

#### 3.1. PERÍODO TEMPORAL COMPLETO (1997-2010)

De los 4.148 documentos recuperados, el 78,25% están indizados con KW+, contabilizándose un total de 28.733 descriptores totales y 9.465 únicos (Tabla 2). En la visualización que ofrece Pajek (Figura 1), se pueden apreciar hasta 6 agrupaciones que, por su significado y relación con los demás, podemos considerar que definen las principales líneas de investigación. Como primera aproximación, es posible distinguir estructuralmente dos naturalezas temáticas claramente diferenciadas. Por un lado, la relacionada con la investigación clínica representada por los descriptores agrupados bajo *Bone-Marrow-Transplantation* y por otra parte, la investigación básica representada por el resto de agrupaciones. Ambas zonas quedan conectadas por dos KW+ que actúan como puentes: *T-Cells* y *Colony-Stimulating Factor*. Esta distinción resulta aún más evidente en la Fig. 3, donde se distinguen claramente estos dos focos de la investigación.

**TABLA 2. Datos relativos al Período Completo (1997-2010)**

Período Completo (1997-2010)			
Documentos	Totales		4.148
	Indizados con KW+ (%)		3.246 (78,25)
KW+	Totales		28.733
	Únicos	Sin Normalización	9.833
		Con Normalización	9.465
	Por documento		8,85
	Frec. Media de aparición		3,04
Visualización	Pajek	Nº Nodos	99
		Frec $\geq$	32
	Parejas		1.703
	VOSviewer	Nº Nodos	668
		Fuerza	5

En particular, los descriptores que conectan con *Bone-Marrow-Transplantation* (Fig. 1 y 3) están relacionados con los estudios ligados al uso terapéutico de las células madre en enfermedades hematológicas. Concretamente, muchos de los descriptores están relacionados con el tratamiento de los procesos oncológicos como la Leucemia (*Chronic Lymphocytic-Leukemia, Acute Myeloid-Leukemia; Leukemia, etc.*) a partir de trasplantes de progenitores hematopoyéticos (*Marrow-Transplantation; Stem-Cell Transplantation*).

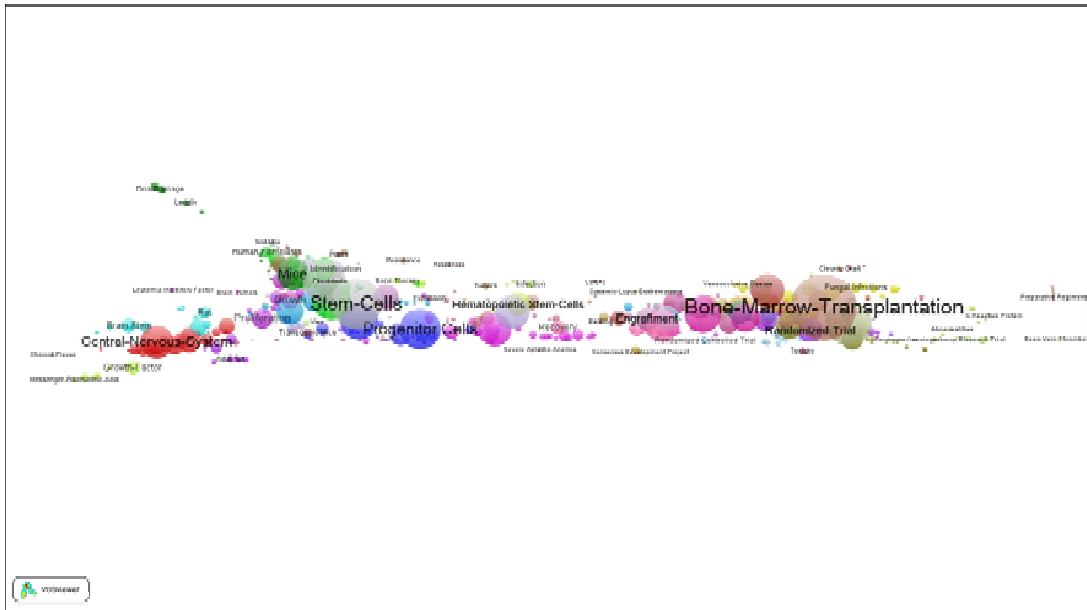
La investigación básica está representada por las cinco agrupaciones restantes de (Figura 1) Existen dos agrupaciones que conectan directamente con la investigación clínica: en la parte central derecha del mapa, los descriptores vinculados con *Stem Cells*, relacionados con los estudios encaminados a la investigación de células madre hematopoyéticas que son las células madre procedentes de la médula ósea. Por otro lado, con aquellos descriptores que forman parte de la línea de investigación relacionada con el estudio de las células madre embrionarias. Concretamente, con los descriptores relacionados con los procesos de expresión (*Expression*) y diferenciación celular (*Differentiation*). En realidad, esta línea se encuentra desagregada en tres agrupaciones: las dos que se acaban de describir (*Expression y Differentiation*) y la que se detecta en torno al descriptor *In-Vitro*. En esta última, se pueden localizar descriptores relacionados con los procesos y las biotecnologías necesarias para localizar, producir, crecer y sobre todo, analizar in-vitro las células madre embrionarias.

La última agrupación que se detecta y que conecta con la línea de investigación de las células madre embrionarias, está relacionada con la investigación con células madre neurales (*Central-Nervous-System*). Estos descriptores están muy relacionados con los estudios centrados en el desarrollo e identificación de las células madre neuronales que se diferencian activamente en el tejido cerebral.

**FIG. 1 Período Completo (1997-2010). Representación en Pajek de los 99 KW+ más frecuentes. (Frecuencia  $\geq 32$ ; coocurrencias = 1.703)**

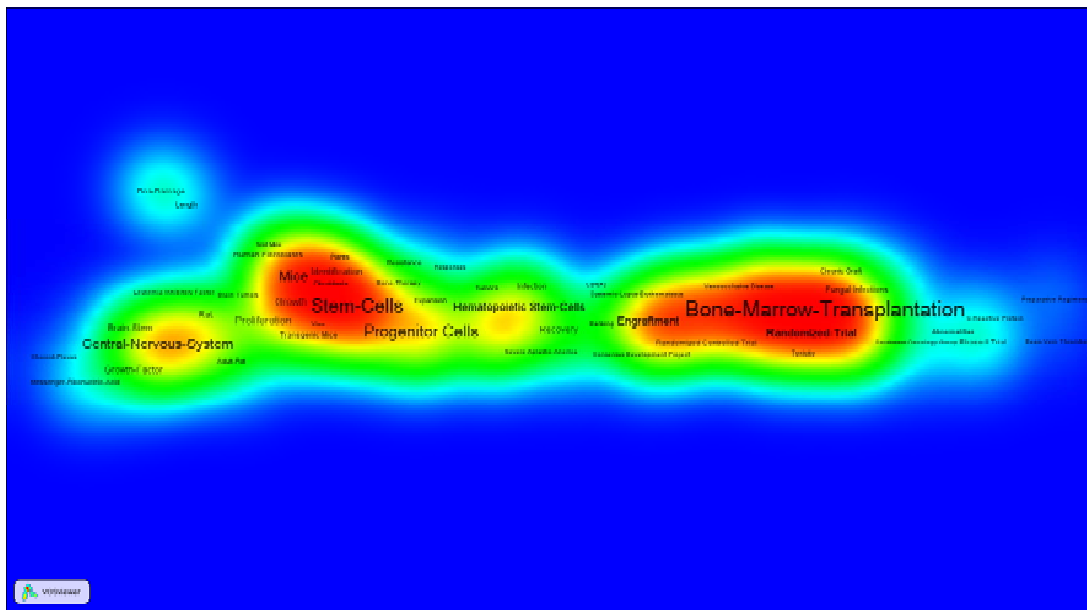


**FIG. 2 Período completo (1997-2010). Mapa de clúster representado con VOSviewer.**  
(Nº de nodos = 668; Fuerza de Asociación = 5)



En la representación de VOSviewer (Fig. 2), no se identifica de una manera tan nítida, el mismo número de agrupaciones que en Pajek. Sólo se identifican claramente 3 líneas de investigación. La representada por el clúster de color marrón encabezada por el descriptor *Bone-Marrow-Transplantation*, la situada en la parte central bajo el descriptor *Stem-Cells*, y la situada en la zona derecha del mapa, representada por el clúster de color rojo encabezada por el descriptor *Central-Nervous-System*. El resto de nodos, aparecen muy entremezclados y son difícilmente detectables.

**FIG. 3 Período Completo (1997-2010). Mapa de calor representado con VOSviewer.**  
(Nº de nodos = 668; Fuerza de Asociación = 5)



### 3.2. PRIMER PERÍODO (1997-2001)

El volumen de documentos para este periodo es de 724 y el número de descriptores sobre el que se realiza el análisis relacional, de 2.080 KW+ únicos (Tabla 3).

Tabla 3. Datos relativos al Primer Período (1997-2001)

Primer Período (1997-2001)			
<b>Documentos</b>	Totales		724
	Indizados con KW+ (%)		560 (77,35)
<b>KW+</b>	Totales		4.631
	Únicos	Sin Normalización	2.134
		Con Normalización	2.080
	Por documento		8,27
	Frec. Media de aparición		2,23
<b>Visualización</b>	Pajek	Nº Nodos	95
		Frec $\geq$	8
		Parejas	678
	VOSviewer	Nº Nodos	543
		Fuerza	15

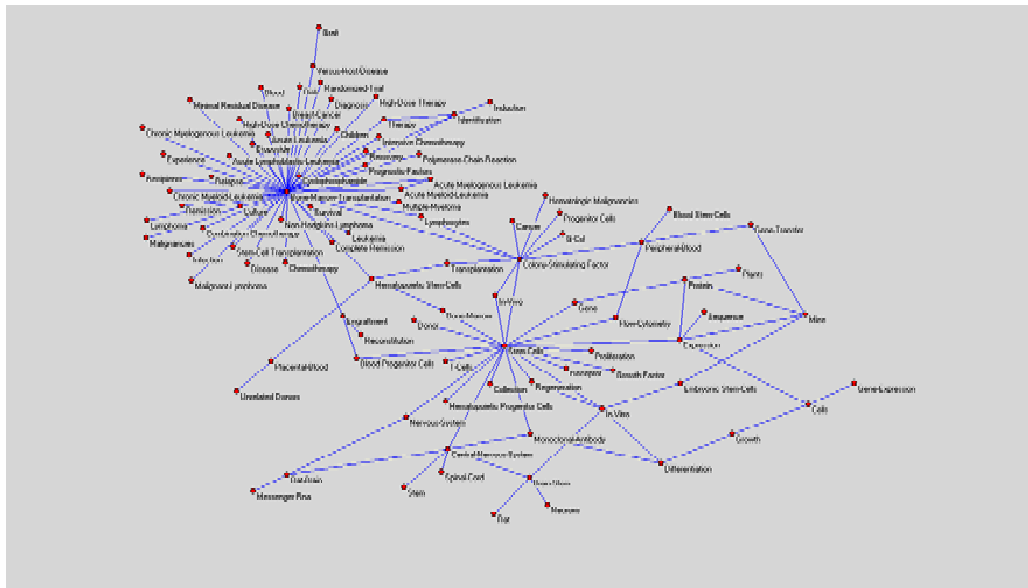
La visualización del dominio ha cambiado sustancialmente, respecto a la imagen del período completo. A pesar de apreciar importantes cambios estructurales en las visualizaciones, la distinción entre los estudios relacionados con la investigación clínica y la básica, es incluso más clara que en el período completo.

La investigación clínica se agrupa claramente en la zona superior izquierda, con un grupo muy numeroso de KW+ relacionados con el uso terapéutico de las células madre en enfermedades hematológicas. La composición de los descriptores de esta línea de investigación es muy similar a la descrita en el período completo. Debido a la clara definición que esta agrupación refleja, es posible considerar que está bien consolidada desde los primeros años del periodo.

El resto de líneas de carácter más básico, no parecen estar tan definidas. Aunque es posible detectar (Figura 4) la mayor parte de los descriptores que estaban presentes en el período anterior, incluso los que ejercían mayor influencia dentro de cada línea de investigación (*Stem Cells, Expression, Differentiation, In-Vitro, Central-Nervous-System*), el comportamiento relacional entre ellos es distinto. La mayoría de las relaciones parten del descriptor *Stem Cells*, nodo principal de la investigación de carácter básico.



**FIG. 4 Primer período (1997-2001). Representación en Pajek de los 95 KW+ más frecuentes. (Frecuencia  $\geq$  8; cocurrencias = 678)**



Las visualizaciones de este período sólo permiten identificar 4 agrupaciones. Por un lado, la investigación clínica, y por otro, la básica. En esta última es posible localizar otras dos agrupaciones más, de carácter más emergente y muy poco definidas pero que pueden tener cierta identidad respecto al resto de descriptores. Éstas muestran estudios relacionados con la investigación de células madre neurales y con algunos de los procesos que intervienen en la proliferación de líneas celulares. Es posible que estas débiles agrupaciones se desarrollen de manera más consistente en el futuro, pudiendo ser consideradas como líneas de investigación emergentes.

### 2.3. SEGUNDO PERÍODO (2002-2006)

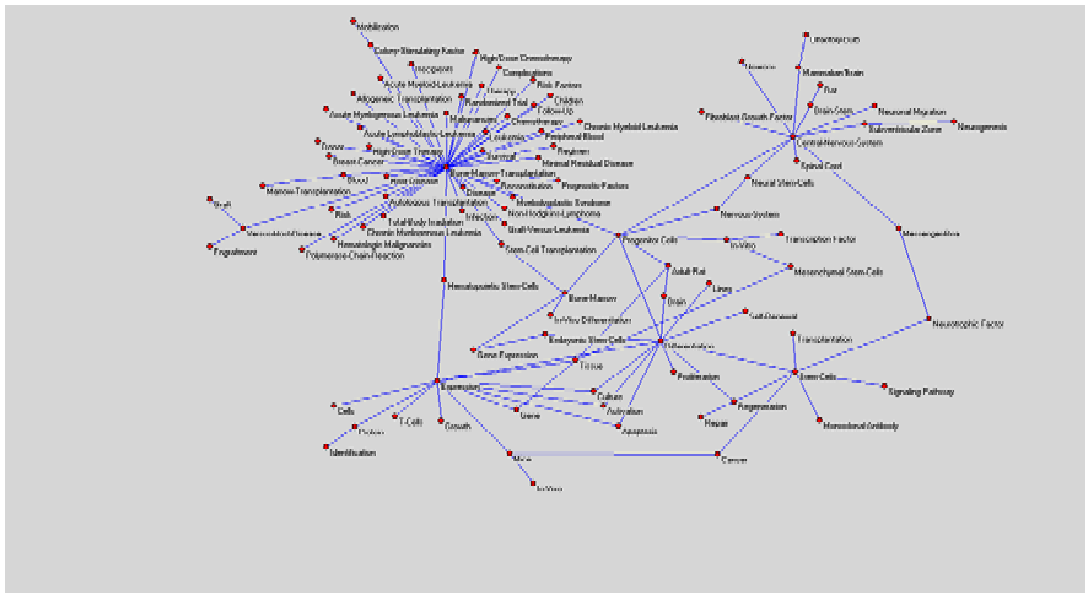
En este período el número de documentos ascendiendo a 1.250 documentos (Tabla 4) y también incrementa el número de KW+ únicos a 3.852 descriptores. En cuanto a las relaciones que se producen entre los KW+, se repite una clara distinción entre los descriptores relacionados con la investigación clínica y la básica con mucho mayor densidad en el caso de la clínica.

**TABLA 4. Datos relativos al Segundo Período (2002-2006)**

Segundo Período (2002-2006)			
<b>Documentos</b>	Totales		1.250
	Indizados con KW+ (%)		979 (78,32)
<b>KW+</b>	Totales		8.581
	Únicos	Sin Normalización	3.966
		Con Normalización	3.852
	Por documento		8,76
	Frec. Media de aparición		2,23
<b>Visualización</b>	Pajek	Nº Nodos	96
		Frec $\geq$	11
	Parejas		820
	VOSviewer	Nº Nodos	363
		Fuerza	30

La agrupación más numerosa y definida es la que se localiza en torno al descriptor *Bone-Marrow-Transplantation* (figura 5), que se identifica temáticamente con los estudios relacionados con el uso terapéutico de las células madre en enfermedades hematológicas. La presencia de esta línea de investigación, detectada en el primer período confirma la consolidación de estos estudios a lo largo del tiempo.

FIG. 5 Segundo período (2002-2006). Representación en Pajek de los 96 KW+ más frecuentes. (Frecuencia ? 11; coocurrencias = 820)



Respecto a la investigación de carácter más básico, estructuralmente se muestran mejor definidos. Los estudios relacionados con las células madre hematopoyéticas identificados mediante los descriptores conectados con *Stem-Cells*, ya no concentran la mayor parte de las relaciones como ocurría en el primer período, sino que han pasado a formar una agrupación con mayor independencia relacional y temática. Al definirse mejor esta línea de investigación también se han definido con mayor claridad los descriptores relacionados con los estudios sobre las células madre embrionarias. Si en el primer período estaban muy entremezclados con los estudios de la células madre hematológicas y tan sólo los procesos de proliferación de líneas celulares podían estar algo más definidos, en este período además de definirse mejor estos estudios (zona inferior del mapa), es posible detectar la presencia de aquellos que se relacionan con los procesos de generación de líneas celulares en torno al descriptor *Expression*. A pesar de la presencia del descriptor *In-Vitro*, aún no se pueden identificar de manera definida los estudios relacionados con los procesos y las biotecnologías necesarias para localizar, producir, crecer y analizar in-vitro las células madre embrionarias como ocurría en el período completo. Es muy posible que esta línea se desarrolle en el siguiente período.

Finalmente, los estudios relacionados con las células madre neurales que aparecían en el período anterior bajo el descriptor *Central-Nervous-System*, no sólo siguen estando presentes para estos años sino que, además, están mejor definidos temáticamente y estructuralmente. Por lo tanto, es posible considerar la consolidación de esta línea de investigación para este período de estudio.

#### 2.4. TERCER PERÍODO (2007-2010)

Este es el periodo con el mayor número de documentos (2.174 documentos) y el mayor número de descriptores (6.178 KW+ únicos). Esta elevada producción puede afectar a la visualización del período completo por la presencia de determinados KW+, y sus frecuencias de aparición que alcanzan valores mayores que otros periodos

TABLA 5. Datos relativos al Tercer Período (2007-2010)

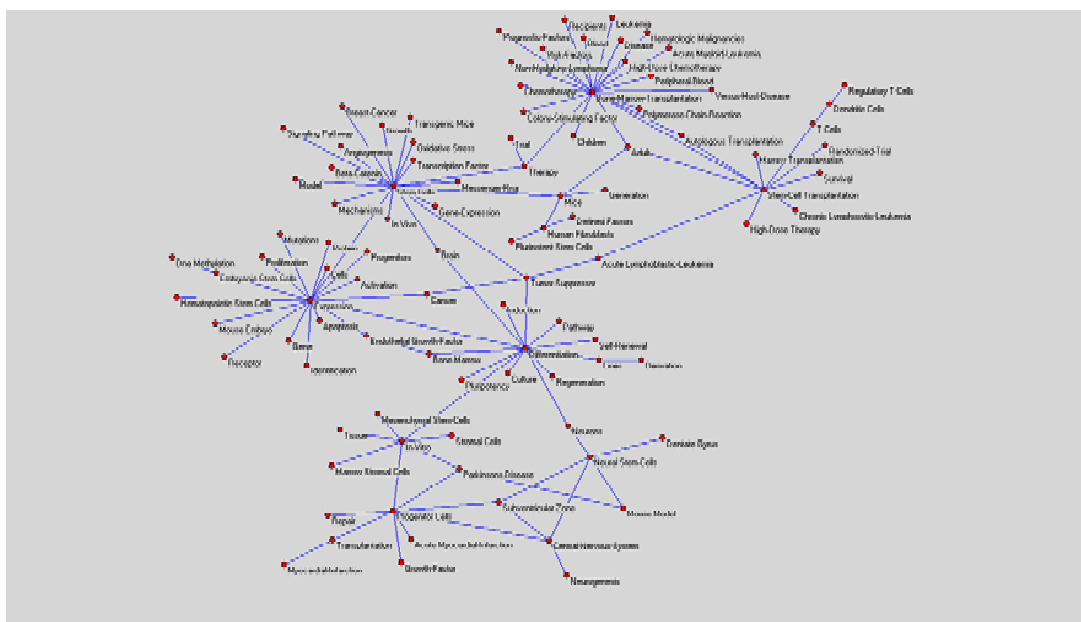
Tercer Período (2007-2010)			
<b>Documentos</b>	Totales		2.174
	Indizados con KW+ (%)		1.707 (78,52)
<b>KW+</b>	Totales		15.521
	Únicos	Sin Normalización	6.393
		Con Normalización	6.178
	Por documento		9,1
	Frec. Media de aparición		2,51
<b>Visualización</b>	Pajek	Nº Nodos	98
		Frec $\geq$	18
		Parejas	1.103
	VOSviewer	Nº Nodos	390
		Fuerza	50

Las visualizaciones permiten detectar cambios importantes en la evolución de la estructura y composición temática de las líneas de investigación localizadas en los periodos anteriores. La distinción entre los descriptores relacionados con la investigación clínica y la básica sigue siendo clara, pero éstos últimos aparecen mucho más desarrollados.

La presencia de descriptores de carácter clínico es más proporcionada al resto de las líneas de investigación, tanto en número de descriptores como en el peso de sus relaciones. Lo que llama la atención en estos años, es la distinción clara que se aprecia (fig. 6) entre los estudios relacionados con los procesos oncológicos, identificados por el descriptor *Bone-Marrow-Transplantation* y que conectan con la investigación básica a través de *Therapy*. Y por otro lado, los que están más relacionados con el trasplante de progenitores hematopoyéticos, identificados con el descriptor *Stem-Cel Transplantation* y que conectan con la investigación básica a través de *Acute-Lymphoblastic-Leukemia*.

Se identifican cinco agrupaciones de carácter más básico. La primera estaría constituida por los estudios relacionados con las células madre hematopoyéticas. Esta línea de investigación, identificada por los descriptores relacionados con *Stem-Cells*, aparece mucho mejor definida que en el período anterior (figura 6, zona superior izquierda del mapa). Esto significa, que estos estudios se han ido desarrollando hasta consolidarse a lo largo de tiempo.

**FIG. 6. Tercer período (2006-2010). Representación en Pajek de los 98 KW+ más frecuentes. (Frecuencia  $\geq$  18; cocurrencias = 1.103)**



Esta línea de investigación está conectada con los estudios relacionados con el uso terapéutico de las células madre en enfermedades oncológicas y con la investigación relacionada con las células madre embrionarias. Si en el período anterior era posible localizar estudios relacionados con los procesos de proliferación y diferenciación celular, en este período aparecen mejor definidos. Por lo que se puede decir que esta línea de investigación se consolida. Pero lo hace incorporando una novedad con respecto a periodos anteriores. Se trata de la localización e identificación de manera clara de los estudios relacionados con los procesos y las biotecnologías necesarias para localizar, producir, crecer y sobre todo, analizar *in vitro* las células madre embrionarias (figura 6, zona inferior). Si bien es cierto, que en el período anterior no era posible identificar de manera definida estos estudios, en este período si es posible.

La última línea de investigación la conforman los estudios relacionados con las células madre neurales. A diferencia del período anterior, los KW+ que identifican esta línea, han perdido definición y ya no constituyen un grupo tan homogéneo. Ahora, es posible localizar dos aspectos de la investigación en este ámbito: la consolidación de los estudios relacionados con el desarrollo e identificación de las células madre neurales mediante los descriptores relacionados con *Neural-Stem-Cell*, y por otro, los relacionados con el infarto de miocardio mediante el empleo de progenitores celulares. Es posible que la pérdida de definición de esta agrupación en este período se deba a la gestación de una nueva línea de investigación que es posible que se consolide en los próximos años como ha ocurrido en el caso de los estudios relacionados con los procesos de proliferación y diferenciación celular.

#### 4. CONCLUSIONES

La visualización de las relaciones que se establecen entre los KW+ ha permitido obtener dos imágenes complementarias de la investigación española sobre células madre, proyec-

tando tanto líneas de investigación claramente consolidadas como las emergentes, lo que nos permite intuir metodológicamente la capacidad de predicción de este tipo de estudios en el desarrollo de un dominio temático a lo largo del tiempo.

El estudio del período completo ha mostrado la estructura estática del dominio con la identificación de las 4 líneas de investigación más consolidadas. Por un lado la investigación relacionada con el uso terapéutico de las células madre en enfermedades hematológicas, por otro, la investigación relacionada con las células madre hematopoyéticas, también la investigación de las células madre embrionarias desagregada en tres aspectos fundamentales como son la proliferación, generación y diferenciación de las líneas celulares, y finalmente, la investigación relacionada con las células madre neuronales.

En cambio, a través del análisis y comparación de la evolución de los resultados por subperíodos, se ha podido reflejar la consolidación de las líneas de investigación emergentes en unos casos y la pérdida de protagonismo en otros. En el primer periodo (1997-2001), se observó una fuerte presencia de los descriptores que representaban los documentos relacionados con la hematología y oncología. En cambio, el resto de líneas no comenzaron a detectarse de manera más clara hasta el segundo período (2002-2006). Fue aquí, donde fue posible localizar claramente agrupaciones de descriptores relacionados con la investigación sobre células madre hematopoyéticas, los descriptores referidos a los estudios sobre la generación, proliferación y diferenciación de células madre y los descriptores relacionados con las células progenitoras neurales. Finalmente, los resultados han mostrado que el peso de estas últimas agrupaciones tuvo una presencia mucho más evidente en el último periodo (2007-2010), en detrimento de las líneas de investigación relacionadas con la aplicación e investigación clínica sobre el trasplante de progenitores hematopoyéticos, a diferencia de lo que ocurría en el primer y segundo período.

En cuanto a los programas de visualización utilizados, el estudio ha mostrado que tanto las visualizaciones en Pajek como en VOSviewer, coinciden tanto en las líneas de investigación como en la localización de los descriptores más influyentes de la red. Pero además de coincidir, se complementan muy bien, ya que Pajek permite detectar de manera más clara la estructura relacional, mientras que VOSviewer posibilita la detección de los grupos y los focos de investigación de una manera más plástica. Sin embargo, el empleo de este tipo de software presenta la limitación de generar sólo visualizaciones correspondientes a un momento determinado de la investigación, sin poder reflejar la evolución en sí misma del dominio científico.

## REFERENCIAS

- Batagelj, V., y Mrvar, A. (2010). Pajek 2.0: package for large network analysis, disponible en: <http://vlado.fmf.uni-lj.si/pub/networks/pajek/> (accedido 19/01/2012).
- Bishop, A.E., BATTERY, L.D.K., y Polak, J.M. (2002). Embryonic stem cells. *Journal of Pathology*, 197 (4), 424-429.
- Boyack, K.W., Börner, K., y Klavans, R. (2009). Mapping the structure and evolution of chemistry research. *Scientometrics*, 79 (1), 45-60.
- Boyack, K.W., y Klavans, R. (2010). Co-Citation Analysis, Bibliographic Coupling, and Direct Citation: Which Citation Approach Represents the Research Front Most Accurately? *Journal of the American Society for Information Science and Technology*, 61 (12), 2389-2404.
- Braam, R.R., Moed, H.F., y Van Raan, F.J. (1991a). Mapping of science by combined co-citation and word analysis. I: Structural Aspects. *Journal of the American Society for Information Science*, 42 (4), 233-251.

- Braam, R.R., Moed, H.F., y Van Raan, F.J. (1991b). Mapping of science by combined co-citation and word analysis. II: Dynamical Aspects. *Journal of the American Society for Information Science*, 42 (4), 252-266.
- Callon, M., Courtail, J.P., Turner, W.A., y Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information Sur Les Sciences Sociales*, 22, 191-235.
- Cambrosio, A., Limoges, C., Courtial, J.P., y Laville, F. (1993). Historical scientometrics? Mapping over 70 years of biological safety research with co-word analysis. *Scientometrics*, 27 (2), 119-143.
- Cantos-Mateos, G., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., y Zulueta, M.A. (2012). Stem cell research: bibliometric analysis of main research areas through KeyWords Plus. *Aslib Proceedings*, 64 (6), 561-590.
- Cantos-Mateos, G., Zulueta, M.A., Vargas-Quesada, B., y Chinchilla-Rodríguez, Z. (2013). *Estudio comparativo sobre la visualización de redes de co-words a través de los descriptores del Science Citation Index y Medline*. I Congreso ISKO España y Portugal / XI Congreso ISKO España. Fundamentos, métodos y modelos para la organización y representación de la información. Oporto, noviembre (paper).
- Cao, Q.L., Benton, R.L., y Whittemore, S.R. (2002). Stem cell repair of central nervous system injury. *Journal of Neuroscience Research*, 68 (5), 501-510.
- Chen, C., Ibekwe-SanJuan, F., y Hou, J. (2010). The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for Information Science and Technology*. 61 (7), 1386-1409.
- Chen, D.Z., Huang, M.H., Hsieh, H.C., y Lin, C.P., (2011). Identifying missing relevant patent citation links by using bibliographic coupling in LED illuminating technology. *Journal of Informetrics*, 5 (3), 400-412.
- Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., y Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. *Journal of Infometrics*, 5 (1), 146-166.
- De Solla Price, D.J. (1963) *Little Science, Big Science*, New York, USA: Columbia University Press.
- Di Giorgio, F.P., Carrasco, M.A., Siao, M.C., Maniatis T., y Eggan, K. (2007). Non-cell autonomous effect of glia on motor neurons in an embryonic stem cell-based ALS model. *Nature Neuroscience*, 10 (5), 608-614.
- Ding, &, Chowdhury, G., y Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis. *Information Processing & Management*, 37 (6), 67-78.
- Garfield, E. (1990). Keywords plus-ISI's breakthrough retrieval method. Part 1. Expanding your searching power on Current Contents on Diskette. *Current Contents*, 1, (32), 5-9.
- Garfield, E., y Sher, I.H. (1993). Keywords Plus™ – Algorithmic derivative indexing. *Journal of the American Society for Information Science*, 44 (5), 298-299.
- Glänzel, W., y Thijs, B. (2011). Using 'core documents' for the representation of clusters and topics. *Scientometrics*, 88 (1), 297-309.
- Hammond, J. (1949). Recovery and culture of tubal mouse ova. *Nature*, 163, (4131), 28-29.
- Kessler, M.M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14 (1), 10-25.
- Lee, W.H. (2008). How to identify emerging research fields using scientometrics: An example in the field of Information Security. *Scientometrics*, 76 (3), 503-525.
- Leydesdorff, L., y Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60 (2), 348-362.
- Leydesdorff, L., y Welbers, K. (2011). The semantic mapping of words and co-words in contexts. *Journal of Informetrics*, 5 (3), 469-475.

- Martínez Serrano, A., y Bjorklund, A. (1996). Protection of the neostriatum against excitotoxic damage by neurotrophin-producing, genetically modified neural stem cells, *Journal of Neuroscience*, 16 (15), 4604-4616.
- Moya-Anegón, F., Chinchilla Rodríguez, Z., Corera-Álvarez, E., Gómez-Crisóstomo, R., González-Molina, A., Muñoz-Fernández, F., y Vargas-Quesada, B. (2007a). *Indicadores bibliométricos de la actividad científica española: 1990-2004*. Madrid: Fundación Española para la Ciencia y la Tecnología (Fecyt).
- Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Muñoz-Fernández, F.J., y Herrero-Solana, V. (2007b). Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology*, 58 (14), 2167-79.
- Quin, J. (2000). Semantic similarities between a keyword database and a controlled Vocabulary Database: an investigation in the antibiotic resistance literature. *Journal of the American for Information Science (JASIS)*, 51 (3), 166-180.
- Rafols, I., Porter, A., y Leydesdorff, L. (2010). Science overlay maps: a new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61 (9), 1871-1887.
- Rip, A., y Courtial, P. (1984) Co-word maps of biotechnology: An example of cognitive scientometrics. *Scientometrics*, 6 (6), 381-400.
- Small, H., y Upham, S. P. (2009). Citation structure of an emerging research area on the verge of application. *Scientometrics*, 79 (2), 365-375.
- Van Eck, N.J., y Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84 (2), 523-38.
- Van Raan, A.F.J., y Tijssen, R.J.W. (1993). The neural net of neural network research. An exercise in bibliometric mapping. *Scientometrics*, 26 (1), 169-192.
- Vargas-Quesada, B., Moya-Anegón, F., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., y Guerrero-Bote, V. (2008). Development of the Spanish scientific landscape: ISI web of science 1990-2005. *Profesional de la Información*, 17 (1), 22-37.
- Vargas-Quesada, B., Chinchilla-Rodríguez, Z., González-Molina, A., y Moya-Anegón, F. (2010). Showing the Essential Science Structure of a Scientific Domain and its Evolution. *Information Visualization*, 9, 288-300.
- Wasserman, S., y Faust, K. (1998). *Social network analysis: methods and applications*. Cambridge: Cambridge University Press.
- White, H.D., y Griffith, B.C. (1981). Author co-citation: a literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32, 163-171.
- White, H., y McCain, K. (1998). Visualizing a Discipline: an Author Co-Citation Analysis of Information Science, 1972-1995. *Journal of the American Society for Information Science*, 49 (4), 327-375.
- White, H.D. (2003). Pathfinder networks and author co-citation analysis: a remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology*, 54 (5), 423-34.
- Zhao, D., y Strotmann, A. (2011). Intellectual structure of stem cell research: a comprehensive author co-citation analysis of a highly collaborative and multidisciplinary field. *Scientometrics*, 87, 115-131.
- Zitt, M., y Bassecouard, E. (1996). Reassessment of co-citation methods for science indicators: Effect of methods improving recall rates. *Scientometrics*, 37 (2), 223-244.
- Zitt, M., Lelu A., y Bassecouard, E. (2011). Hybrid citation-word representations in Science mapping: portolan charts or research fields? *Journal of the American society for information Science and Technology*, 62 (1), 19-39.
- Zulueta, M.A., Cantos-Mateos, G., Sánchez, C., y Vargas-Quesada, B. (2011). Research involving women and health in the Medline database, 1965-2005, Co-term analysis and visualization of main lines of research. *Scientometrics*, 88 (3), 679-706.