

ANÁLISIS DE LA PRODUCCIÓN CIENTÍFICA ESPAÑOLA CITADA EN PATENTES BIOTECNOLÓGICAS EN EE.UU.

Luis M. Plaza, Armando Albert*

Resumen: A partir de las experiencias obtenidas en dos recientes estudios cuantitativos sobre transferencia de tecnología, se dan a conocer algunas de las aplicaciones que ofrece el análisis de la literatura científica citada en documentos de patente, así como las limitaciones y dificultades metodológicas que presenta este tipo de estudios. Este artículo describe el procedimiento metodológico utilizado para analizar el flujo de información científica generada y publicada por los investigadores de nuestro sistema público de I+D que ha servido de base para el desarrollo de patentes nacionales y extranjeras. La metodología aquí descrita permite obtener una serie de indicadores para determinar las verdaderas capacidades de nuestras universidades y centros de investigación para producir conocimientos potencialmente utilizables por el sector productivo, para incrementar la generación de patentes, impulsar la transferencia de tecnología y la cooperación con la industria y promover la creación de nuevas empresas de base tecnológica.

Palabras clave: Transferencia de tecnología; Patentes; Análisis de citas

Abstract: Based on two recent studies about technology transfer, this contribution deals with both usefulness and difficulties for the analysis of scientific literature cited in patent documents. This article describes a methodology used to assess the flow of scientific information produced and published by Spanish scientists of the public R&D system that constitutes the background for generation of in-house and foreign patented technology. The described methodology allows us to obtain a set of indicators and reliable information to determine the real capabilities of our Universities and Research centres to produce knowledge potentially useful for the productive sector, increase patent generation, improve technology transfer and co-operation with Spanish industries and to promote generation of spin off companies.

Keywords: Technology transfer; Patent citation analysis

1 Antecedentes

Durante la década de los 80, diversos investigadores iniciaron lo que luego sería una larga serie de estudios basados en el análisis bibliométrico de la literatura científica citada en documentos de patentes USA, como los de Broad en 1997 (1) y Narin et al. en 1985, 1988 y 1992 (2, 3, 4). El origen de este tipo de estudios obedecía a la necesidad de justificar, ante las autoridades de los EE.UU., el fuerte ritmo de crecimiento en inversión dedicada a investigación científica básica, ya que desde distintas esferas se alegaban diversos argumentos sobre la escasa rentabilidad de la misma. Ante la polémica suscita-

* Centro de Información y Documentación Científica (CINDOC-CSIC). España. Correo-e: lplaza@cindoc.csic.es; aalberto@cindoc.csic.es.

Recibido: 10-3-03; 2.^a versión: 14-5-04.

da, se procedió a analizar la literatura científica que aparecía citada en las primeras páginas de las patentes USA (estado de la técnica en cada patente), entendiendo que los indicadores derivados de este tipo de estudios constituyen una manera eficaz de valorar la transferencia de conocimientos entre el ámbito científico y la industria o, dicho de otra forma, el impacto de la investigación científica sobre el desarrollo tecnológico. Como resultado de dicho estudio se pudo comprobar que una gran parte de la literatura científica citada correspondía a trabajos de investigación científica básica y que, además, un elevado porcentaje de estos trabajos habían sido financiados con fondos públicos.

Independientemente de las repercusiones que este trabajo y otros, en la misma línea, tuvieron en la política científica estadounidense, este tipo de estudios constituyeron una llamada de atención sobre las posibilidades que ofrece el análisis de los documentos de patente en el ámbito de la cienciometría. De hecho, a partir de aquel momento se produjo una fuerte corriente de interés sobre este ámbito de actuación, cuya evolución puede ser fácilmente observable a través de los trabajos presentados en los congresos y conferencias internacionales de cienciometría celebrados durante los últimos años, así como en las revistas especializadas.

Análisis de literatura citada en patentes, realizados en España

En España el número de estudios de esta naturaleza ha sido particularmente escaso. Este hecho puede considerarse como especialmente negativo por cuanto, para un país con una balanza tecnológica de pagos fuertemente deficitaria, es esencial conocer con rigor y en profundidad los mecanismos y los factores que condicionan la generación de patentes y la transferencia de conocimientos entre el sector público de investigación y el sector productivo, máxime teniendo en cuenta que las interacciones ciencia-industria en nuestro país están aún lejos de alcanzar la intensidad y complejidad que deberían tener. Como experiencias españolas en este tipo de estudios cabe mencionar un proyecto de investigación sobre transferencia de conocimientos en distintos sectores tecnológicos realizado en 2000 para el ámbito de la Comunidad de Madrid (5). Este estudio, realizado por el CINDOC, permitió obtener una serie de resultados muy significativos sobre el escaso impacto de la investigación científica española en el desarrollo de nuevas tecnologías y, por tanto, de patentes de titularidad española, hecho que contrasta fuertemente con una significativa presencia de citas a trabajos científicos de nuestros investigadores en documentos de patentes USA (Albert et al., 2001, 2002) (6,7). Diversos resultados de dicho estudio fueron dados a conocer en el *8th International Conference on Scientometrics and Informetrics* celebrado en 2002 en Sydney, Australia y en el *7th International S&T Indicators Conference 2002* realizado en Karlsruhe, Alemania.

2 Posibilidades, limitaciones y dificultades

El objetivo de esta contribución es dar a conocer algunas de las posibilidades que ofrece el análisis de la literatura científica citada en patentes, así como las limitaciones y dificultades que presenta este tipo estudios.

2.1 Sobre el significado real de la literatura citada en patentes

Existen notables diferencias entre las patentes registradas en la *US Patent and Trademark Office* (USPTO) y en la *European Patent Office* (EPO) en cuanto a las citas recogidas en las mismas, debido a las diferentes políticas empleadas. En los EE.UU. no existen unas normas tan restrictivas como en Europa a la hora de exigir al solicitante de una patente no haber hecho público previamente, bajo ningún medio o procedimiento de comunicación, los descubrimientos para los que se solicita protección y además se exige citar toda la información relevante para la patente. Por este motivo, las patentes registradas en la USPTO muestran, por lo general, un mayor número de citas y una mayor diversidad de conexiones temáticas. Por el contrario, en las patentes tramitadas a través de la EPO las citas muestran un menor número de conexiones entre campos del conocimiento (Meyer, 2000) (9, 10), aunque puede atribuirse a éstas un mayor grado de interrelación. A lo anterior, hay que añadir que las citas en patentes USA corresponden, en un elevado porcentaje, a las aportadas por el solicitante, si bien esta información es también confirmada y complementada por el examinador, lo que da mayor credibilidad, en cuanto a relevancia y exhaustividad, a esta información.

Otro aspecto de interés es la abrumadora mayoría de inventores norteamericanos que registran sus invenciones en la USPTO, lo que conlleva una mayor presencia, en términos proporcionales, de citas a patentes y publicaciones científicas de autores norteamericanos (McMillan, 2000) (11). Precisamente por este motivo, las citas a literatura científica y técnica de otros países, en patentes USA, tienen un valor especialmente significativo.

2.2 Información que puede obtenerse del análisis de la literatura citada en patentes y de las patentes citantes

En cuanto a las citas, cabe distinguir dos tipos de documentos; las citas a otras patentes y las citas a documentos no-patente, entre las que se encuentran las correspondientes a literatura científica. Respecto a este último conjunto, su análisis nos permite, entre otros aspectos, obtener información acerca de los autores/investigadores, instituciones, revistas citadas, cooperación, carácter básico/aplicado de la investigación, las temáticas y la proximidad en el tiempo entre la patente citante y la literatura citada.

Por otra parte, el análisis de las patentes citantes permite conocer, entre otros aspectos, la identidad, ubicación y tipología de las empresas o entidades solicitantes o/y titulares de las patentes, así como la identidad y adscripción empresarial o institucional del inventor o inventores.

2.3 Indicadores obtenidos

Tras establecer una muestra constituida por 87 investigadores españoles del sector público de I+D, tomando como criterio de selección el hecho de haber sido líderes de grupo en proyectos de investigación financiados por el Plan Nacional en el área de Biotecnología, se han realizado una serie de búsquedas en la base de datos de patentes de la USPTO, a fin de identificar publicaciones científicas de dichos investigadores que hayan sido citadas en patentes USA entre 1980 y 2000.

Como paso previo al análisis de citas en patentes se ha analizado la producción científica (promedio anual de publicaciones por investigador) de los investigadores a lo largo de tres años (1996-1998) a partir de los datos de publicaciones científicas recogidas en la base de datos *Science Citation Index* (SCI). La finalidad de este indicador es la de conocer si los componentes de esa muestra son realmente representativos de un colectivo con una buena productividad. Los datos obtenidos permiten afirmar que la productividad media es ligeramente superior a los 3 artículos al año, lo que representa un valor muy competitivo si lo consideramos en el contexto europeo.

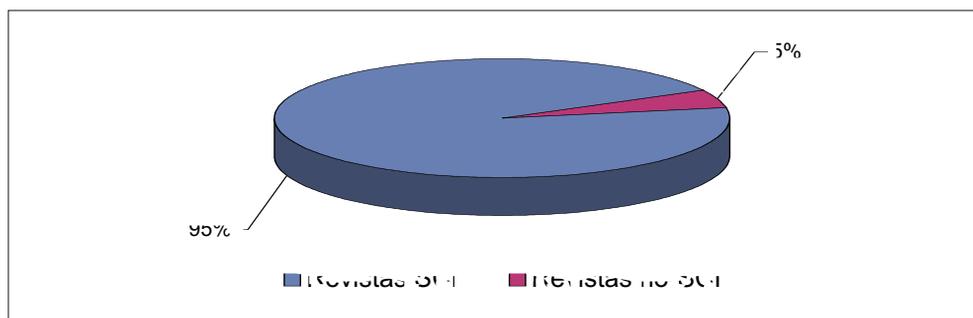
El análisis de las citas a los investigadores estudiados en patentes USA ha permitido identificar 214 patentes que citan a 35 de los 87 investigadores que componen la muestra, valor que constituye el 40%. El total de citas recibidas por éstos es de 257, lo que representa un promedio de 7,3 citas por investigador.

El análisis de las citas, considerando la adscripción institucional de los investigadores citados en patentes americanas, permite determinar que el mayor porcentaje de las citas son recibidas por investigadores del CSIC.

Con el fin de tener una aproximación en cuanto a la calidad científica de los artículos citados en patentes americanas, se ha comprobado si dichos artículos corresponden a revistas cubiertas por la base de datos SCI o no, ya que esta se considera receptora de revistas científicas de gran calidad que constituyen la «corriente principal de la ciencia».

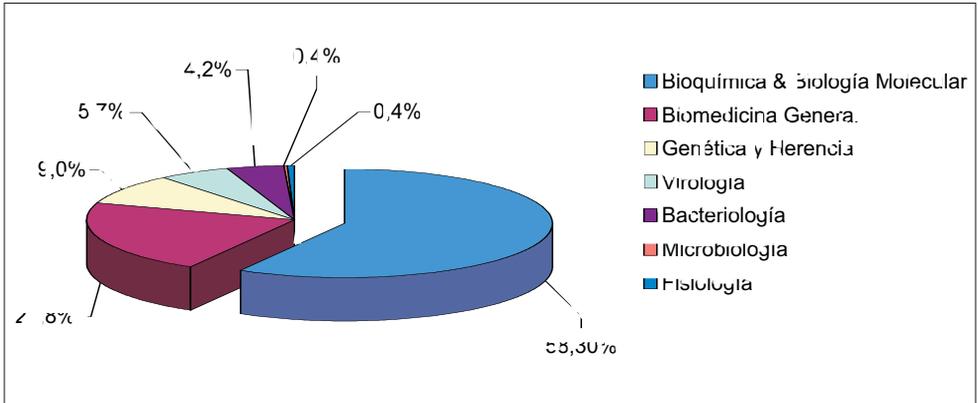
Se ha comprobado que de las 257 citas detectadas, 245 corresponden a revistas SCI, lo que representa el 95% (fig. 1).

Figura 1
Publicaciones de autores españoles citados en patentes USA



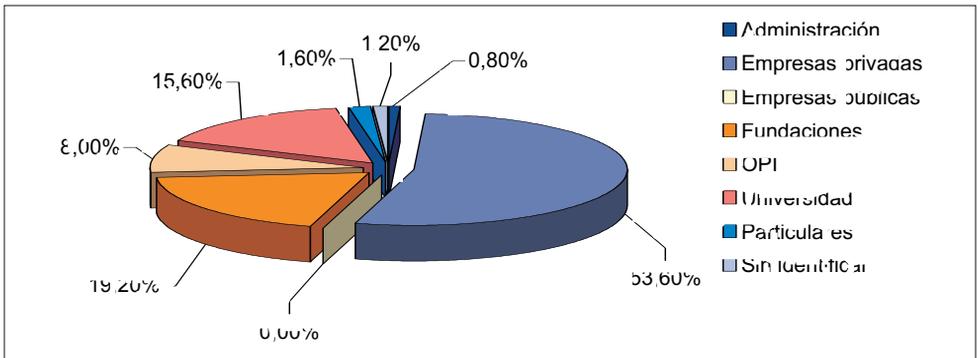
Estas citas corresponden a trabajos de investigación publicados en un total de 34 títulos de revista. Un análisis temático de estas últimas permite identificar 25 títulos de interés en el área de Investigación Biomédica, los cuales contienen un 94% de las citas en Biotecnología. A las revistas de Biología, Física, Medicina Clínica y Química les corresponde el resto de las citas (5,3%), distribuyéndose casi a partes iguales entre los títulos. En las revistas de Investigación Biomédica, las citas se distribuyen de la siguiente forma: Bioquímica & Biología Molecular (58,29%), Biomedicina General (21,80%), Genética y Herencia (9,00%), Virología (5,69%), Bacteriología (4,26%), Microbiología (0,47%) y Fisiología (0,47%) (figura 2).

Figura 2
Distribución de las citas de investigación biomédica



El análisis de la titularidad de las patentes citantes, permite señalar un amplio dominio del sector privado, ya que a éste le corresponden 134 de las 215 patentes detectadas, valor que representa el 53,6% (figura 3).

Figura 3
Patentes citantes atendiendo al tipo de entidad titular de la patente

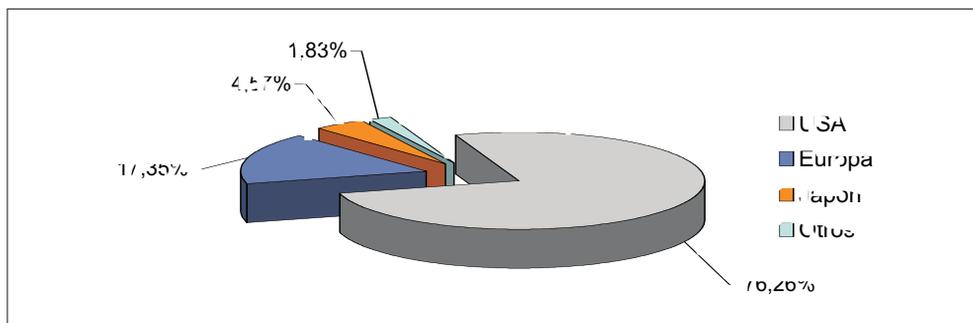


Las fundaciones médicas y universitarias constituyen el segundo grupo en importancia por el número de patentes, seguidas por las universidades y los organismos públicos de investigación (OPI).

Cabe señalar que, entre las empresas biotecnológicas identificadas, se encuentran buena parte de las más importantes de este sector a escala mundial como son: *Bristol Myers Squibb Co.*, *Envirogen Inc.*, *Life Technologies Inc.*, *Monsanto Co.*, *Novartis Finance Co.*, *Plant Genetics Systems*, *Stratagene*, *Virogenetics Corp.*, *Zeneca Ltd.*, *Merck & Co. Inc.*, *Rhone-Poulenc Pharmaceuticals Inc.*, *Eli Lilly & Co.*

La mayor parte de las patentes citantes (76,2%) corresponden a titulares con residencia en los EE.UU. El número de patentes de residentes en Europa representa el 17,3% (figura 4).

Figura 4
Origen geográfico de las patentes citantes de trabajos españoles de interés en Biotecnología

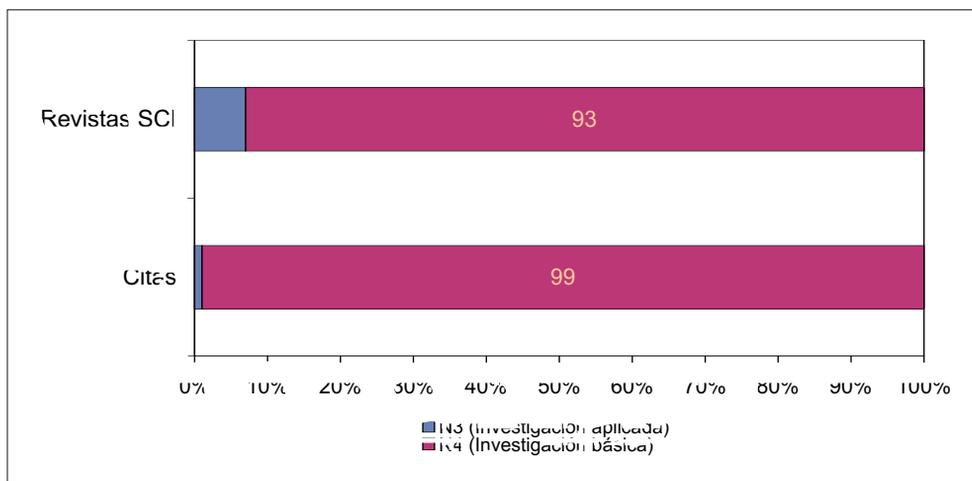


Por otra parte, se ha analizado el carácter básico o aplicado de los artículos publicados por los investigadores y que han sido citados en las patentes americanas, para determinar sobre qué tipo de conocimientos se sustentan dichas patentes.

El análisis se ha realizado cuantificando la producción científica citada en las patentes, considerando el número de citas recibidas por las revistas en las que han sido publicados los trabajos. De estas citas se han seleccionado las correspondientes a las revistas cubiertas por la base de datos SCI y se ha comprobado, para cada caso, el nivel mediante el que se define el carácter básico o aplicado de dichas publicaciones (Noma, 1986).

El 99% de las 245 citas en patentes USA a los artículos publicados en revistas SCI por los investigadores, corresponde a trabajos de investigación científica básica (nivel 4) mientras que el 1% restante corresponde a artículos de investigación aplicada (nivel 3) (figura 5).

Figura 5
Carácter básico/aplicado de los trabajos citados



Buena parte de los trabajos de autores españoles citados en las patentes corresponde a artículos publicados en revistas con un factor de impacto (FI) modesto. El 43% de estos trabajos han sido publicados en revistas con $FI < 4$. El 37% de los artículos citados están en revistas de bioquímica y biología molecular (con interés biomédico) con un $FI > 4$. El 20% restante corresponde a trabajos publicados en *Nature* y *Science*. Este hecho viene a matizar la idea de una estrecha relación entre la importancia tecnológica de un artículo científico y un elevado factor de impacto de la revista en que éste ha sido publicado.

2.4 Dificultades metodológicas

La metodología más empleada en el análisis de la literatura científica citada en documentos de patente, consiste en la identificación, en los registros incluidos en las bases de datos de la USPTO, de las citas a revistas *SCI* y, a partir de éstas, la determinación del carácter básico o aplicado de las publicaciones, la identificación de los autores, su adscripción institucional, país de origen y otros parámetros derivados de los anteriores.

Sin embargo, esta no es la única aproximación posible. A diferencia del método anterior, las experiencias aquí descritas consisten en una identificación previa de los investigadores que trabajan en un área específica y que además se ajustan a unos estándares mínimos de calidad en la investigación. A partir de una relación nominal así establecida, se procede a la búsqueda de la literatura citada en patentes, correspondiente a dichos investigadores.

Los principales problemas que plantea esta metodología es la identificación y verificación de la identidad de los investigadores (autores) citados. Estos problemas derivan de la existencia de homonimias y del hecho de que la mayor parte de las citas sólo recogen el nombre del primer autor firmante, con lo que sólo puede detectarse un investigador si es primer firmante del trabajo citado. Para resolver estos inconvenientes es imprescindible, por una parte, verificar si las publicaciones citadas correspondientes a un autor dado, se refieren efectivamente al mismo, para lo que se debe confrontar la cita detectada con la información contenida en bases de datos bibliográficas y, por otra parte, a fin de detectar las posibles citas a los trabajos de un autor cuando este no firma en primer lugar («citas ocultas»), es imprescindible localizar previamente su producción científica e identificar a los investigadores que como coautores, suelen firmar los trabajos en primera posición y efectuar las correspondientes búsquedas y comprobaciones en la base de datos de patentes.

En este caso se ha procedido a un análisis considerando los coautores habituales, lo que ha permitido determinar que, como promedio, el número de «citas ocultas» en patentes que recibe un autor es de 2,5 veces el valor obtenido en una primera búsqueda. Por otra parte, se ha comprobado que un 45% de los autores aparentemente no citados, sí habían sido citados, pero las citas resultaban indetectables al no figurar como primeros autores (Albert y Plaza, 2004) (12).

Conviene señalar asimismo que, mientras el método iniciado en EE.UU., y seguido por otros investigadores, está orientado a la obtención de macroindicadores, el método utilizado aquí se enmarca en el ámbito de los microindicadores. De hecho, el objetivo fundamental, al que se aplica esta metodología, es el de identificar los recursos huma-

nos del sistema público español de I+D cuyas capacidades y actividades investigadoras son de potencial interés para el desarrollo tecnológico.

Por otra parte, el método norteamericano se basa en el análisis de las citas a revistas cubiertas por el SCI, mientras que las citas a otras fuentes no son consideradas. A este respecto, conviene tener en cuenta que según el ámbito científico objeto de estudio, se deben considerar las citas a otros tipos de publicaciones, aunque sólo las realizadas a revistas SCI son las que permiten una aproximación fiable al análisis del carácter básico o aplicado de los trabajos citados.

3 Aplicabilidad de los datos e indicadores

Respecto al tipo de actividades y entornos en los que se pueden aplicar los indicadores obtenidos mediante análisis de citas en patentes resulta obligado citar la vigilancia tecnológica, la inteligencia competitiva y la investigación estratégica, que en cualquier caso son de potencial interés en la planificación de la I+D+i, tanto en el sector público como en el privado, y en la evaluación de la investigación. Por tanto, existe una notable variedad de agentes públicos y privados o instituciones que pueden ser beneficiarios de este tipo de información, (organismos responsables de políticas científicas y tecnológicas, evaluadores de I+D, empresarios, oficinas de transferencia de tecnología, centros tecnológicos, fundaciones universidad-empresa, investigadores, agencias financiadoras, etc).

En el caso de España, país que a pesar de mostrar una deficitaria balanza de pagos tecnológica, muestra una productividad y calidad científica de buen nivel en numerosos campos de la investigación, es importante conocer la intensidad y las características del flujo de conocimientos científicos que, generados por nuestros investigadores del sistema público, van a formar parte del estado de la técnica de las invenciones patentadas por instituciones y empresas de otros países y, por tanto, ajenas a nuestros intereses nacionales. Conforme a estos objetivos, el análisis de la literatura científica (citada en patentes) cuya autoría corresponde a nuestros investigadores, permite identificar y cuantificar los recursos humanos cuya actividad es de potencial interés tecnológico, aunque el hecho de ser citados en documentos de patente no implica, necesariamente, una relación directa entre la investigación citada y la invención patentada. Así el análisis de las patentes citantes permite asimismo conocer hacia qué países, instituciones, empresas y sectores tecnológicos fluye ese conocimiento. A partir de la identificación de las empresas o entidades titulares de las patentes citantes, es posible analizar sus características (tamaño de la empresa, objetivos de mercado, proyección internacional), su cartera de patentes y determinar el peso específico de las mismas en el sector tecnológico correspondiente.

En este marco de intereses, y junto al análisis de la literatura citada en documentos de patente, resulta igualmente importante el estudio de las patentes de titularidad nacional y extranjera en cuya generación han participado investigadores españoles, en calidad de inventores, ya que esto es imprescindible para determinar en qué medida nuestros recursos humanos y el conocimiento científico generado por ellos influyen en una y otra dirección y si son o no convenientemente utilizados en beneficio de nuestros sectores productivos.

En definitiva, el objeto de un análisis de esta naturaleza es el de comparar los resul-

tados obtenidos con los correspondientes al análisis de los mismos sectores tecnológicos en el ámbito nacional. Así se puede obtener una información útil para definir las políticas y actuaciones necesarias para que las capacidades de los investigadores del sistema público de I+D y los conocimientos científicos generados por ellos puedan ser captados y utilizados para objetivos tales como incrementar el número de patentes de titularidad española, intensificar la cooperación con la industria nacional y lograr la creación de nuevas empresas de base tecnológica.

Nota: Parte de los indicadores recogidos en este artículo han sido obtenidos por los autores del mismo en el «Estudio sobre la Biotecnología en el sistema público español de I+D. Indicadores de actividad científica y tecnológica», financiado por la Fundación para el Desarrollo de la Investigación en Genómica y Proteómica.

Referencias

1. BROAD, W. (1997). Study finds public science is pillar of industry. *The New York Times, Science Times*, may 13.
2. NARIN, F.; NOMA, E. (1988). Is technology becoming science? *Scientometrics* 7 (3-6), 369-381.
3. NARIN, F.; OLIVASTRO, D. (1988). Patent citation analysis: new validation studies and linkage statistics. En van Raan, Nederhoff & Moed (Eds.), *Science Indicators: their use in science policy and their role in science studies*, DSWO Press, The Netherlands, pp. 14-16.
4. NARIN, F.; OLIVASTRO, D. (1992). Status report – linkage between technology and science. *Research Policy* 21 (3), 237-249.
5. Investigación científica e innovación en la Comunidad de Madrid. Un estudio para incentivar la generación de patentes. Informe final, Proyecto 06/131/99. Dirección General de Investigación, Comunidad de Madrid.
6. ALBERT, A.; PLAZA, L.M.; SANCHO, R. (2001) A study to promote a more efficient transfer of knowledge from Spanish accademical research to technological development. *Proceedings of the 8th International Conference on Scientometrics and Informetrics*, vol. 2. Sydney, Australia.
7. ALBERT, A.; PLAZA, L.M.; SANCHO, R. (2002). La transferencia de conocimientos entre el sector público de I+D y el sector productivo. Obstáculos y posibles soluciones. *Boletín de la Sociedad Española de Bioquímica y Biología Molecular* (132), 12-16.
8. PLAZA, L.M.; ALBERT, A. (2002). La ciencia básica al servicio del desarrollo tecnológico. Principales indicadores para países de América Latina. En *Indicadores de Ciencia y Tecnología en Iberoamérica. Agenda 2002.*, 282-.297, RICYT, Buenos Aires, Argentina.
9. MELLER, M. (2000a) Wath is special about patent citations?. Differences between scientific and patent citations. *Scientometrics* 49 (1), 93-123.
10. MELLER, M. (2000b) Does science push technology? Patents citing scientific literature. *Research Policy* 29, 409-434.
11. McMILLAN, G.S.; NARIN F.; DEEDS D. (2000) An analysis of the critical role of public science in innovation: the case of biotechnology. *Research Policy* 29, 1-8.
12. ALBERT A.; PLAZA, L.M (2004) The transfer of knowledge from the Spanish public R&D system to the productive sectors in the field of Biotechnology. *Scientometrics* 59 (1), 3-14.