

Temas de actualidad

Metagenoma: acceso a los recursos potencialmente ilimitados de microorganismos no cultivables

Manuel Ferrer

CSIC, Instituto de Catálisis y Petroleoquímica, Campus Cantoblanco, 28049 Madrid.
E-mail: mferrer@icp.csic.es

Avances en Microbiología: del cultivo directo a las genotecas

La diversidad microbiana es una fuente importante de productos y procesos biotecnológicos. Sin embargo, los métodos tradicionales de cultivo de microorganismos han limitado el análisis a aquellos que son capaces de crecer en condiciones de laboratorio. Por tanto, uno de los retos más importantes de la microbiología y biotecnología actuales es el estudio de toda esta diversidad. El aislamiento directo de DNA de una determinada muestra nos permite el acceso a los recursos potencialmente ilimitados de microorganismos no cultivables. La clonación de todo este DNA genómico (el "metagenoma") en complejas genotecas, puede constituir una herramienta importante en la búsqueda de nuevas actividades enzimáticas de interés industrial y medioambiental. Conseguir de forma colectiva todos los genomas microbianos presentes en un determinado hábitat es un objetivo de los programas de descubrimiento de nuevas enzimas y encaminado a la utilización de estos recursos, evitando el problema asociado al cultivo tradicional.

Diversidad microbiana como fuente de nuevas enzimas

Existe en la actualidad una gran demanda de productos y procesos biotecnológicos. Los procesos catalizados por enzimas representan una alternativa a los procesos químicos convencionales, bien porque se produce un menor número de subproductos, o bien porque son procesos más selectivos hacia el producto de interés. La dependencia actual de productos derivados del petróleo, y la tendencia a la utilización de otras fuentes de carbono (menos contaminantes) para obtener energía y productos de interés farmacéutico, clínico y alimentario, entre otras, abren nuevas expectativas a una química verde dominada por la biotecnología. Sin embargo, el desarrollo de nuevos productos y procesos está limitado a las enzimas

que se conocen hasta la fecha. En la actualidad se conocen aproximadamente 4.300 enzimas (secuenciadas), de las que 300 están comercializadas. Estas a su vez se han utilizado en multitud de procesos, encaminados a la obtención de productos de alto valor añadido, por ejemplo, biopolímeros formados por unidades de amino ácidos o azúcares, compuestos quirales con actividad biológica, bioetanol y biodiesel, etc.; sin embargo, la fuerte demanda de nuevos productos con actividad biológica que impera en la sociedad, requiere la búsqueda de nuevas enzimas.

Si bien las propiedades y actividades de las enzimas conocidas pueden ser modificadas mediante técnicas de ingeniería de la reacción (mezclas de disolventes, inmovilización, modificación química, diseño de enzimas semisintéticas, etc.) o ingeniería del biocatalizador (evolución molecular dirigida, diseño combinatorio y computacional, mutagénesis dirigida y saturada, etc.), éstas parten de enzimas cuyas estructuras y actividades son conocidas, lo que limita a priori el diseño de novedosos procesos enzimáticos. Estudios recientes han demostrado que el número de proteínas (enzimas) existentes en el Universo (físico), asumiendo una longitud media de 200 amino ácidos, puede ser de 20^{200} , de las cuales se conocida actualmente una mínima parte (Koonin *et al.*, 2002). En este punto, la biodiversidad presente en nuestro planeta, en particular la microbiana, es la fuente mayor de recursos biotecnológicos, y por otro lado es la menos desarrollada (Handselman *et al.*, 1998). Los métodos tradicionales para obtener nuevos biocatalizadores por cultivo directo de cepas puras de microorganismos es una estrategia estándar y potente, todavía en uso (Ogawa y Shimitzu, 1999). De esta forma, grandes compañías, como BASF, tienen una amplia colección de microorganismos que utilizan como productores no sólo de biocatalizadores, sino también de metabolitos de alto valor añadido. Sin embargo, la biosfera está dominada por los microorganismos y la mayoría no han sido aún estudiados (Whitman *et al.*, 1998).

El desarrollo de métodos moleculares aplicados

a la ecología microbiana ha permitido estimar que la diversidad existente en la naturaleza se encuentra entre 10^5 y 10^7 especies (siendo el concepto de especie en procariotas más amplio que en eucariotas; Roselló-Mora y Amann, 2001). Su importancia radica en el papel clave del mundo procarionta en las transformaciones bio-geoquímicas que ocurren en la naturaleza. Se estima que sólo el 1% de los microorganismos han sido cultivados y caracterizados, dejando una gran cantidad de ellos (más de 10.000 especies diferentes) sin explorar (Tabla 1). El desconocimiento de esta enorme diversidad es una de las carencias más importantes que ha tenido el estudio de la Biología, Microbiología y Enzimología. Se puede decir que los métodos tradicionales para el cultivo de los microorganismos han limitado nuestro análisis a los que son capaces de crecer en condiciones de laboratorio (Pace, 1997; Suzuki *et al.*, 1997). Estas limitaciones han dado lugar a un surgimiento de nuevas técnicas de Biología Molecular, basadas en el aislamiento de todo el ADN genómico y encaminadas principalmente al (i) estudio de la diversidad microbiana a un nivel genético, mediante la extracción directa del ARN o DNA, y su posterior análisis, y (ii) a la puesta en marcha de programas de secuenciación de microorganismos.

Análisis de ARN y ADN y su utilización en la identificación y discriminación de microorganismos

La composición de comunidades microbianas complejas se ha estimado principalmente de acuerdo a la diversidad de las secuencias de DNA que codifican los rRNA que forman parte de las subunidades del ribosoma, existiendo en la actualidad una enorme base de datos. Todo ello las convierte en útiles herramientas identificadoras y

filogenéticas mediante técnicas de *Genetic Fingerprint* (huella genética) (Muyzer, 1999). Esta técnica proporciona un perfil de la diversidad microbiana basándose en la separación física de especies únicas de ácido nucleico. La estrategia general consiste en: primero, extracción de ácido nucleico (DNA o RNA) de una muestra dada; segundo, amplificación de genes que codifican el rRNA 16S y tercero el análisis de los productos de PCR en una electroforesis sometida a un gradiente desnaturante bien con urea o bien de temperatura (*Denaturing Gradient Gel Electrophoresis*; DGGE, o *Temperature Gradient Gel Electrophoresis*; TGGE) (Muyzer *et al.*, 1993; Heder *et al.*, 2001; Villadas *et al.*, 2002).

Con la identificación de organismos por técnicas genotípicas se ha puesto de manifiesto la existencia de un gran número de microorganismos, tanto más diferentes, en cuanto a la complejidad del ADN extraído y ha permitido mostrar el potencial genético "ilimitado" de recursos de los microorganismos no-cultivables, sin que su cultivo en el laboratorio sea necesario (Amman *et al.*, 1995; Borneman y Triplett, 1997). Así, por ejemplo, se estima que más de un 80% de los amplificados presentan una identidad inferior al 90% con secuencias descritas en las bases de datos, siendo además la mayoría de las homologías descritas frente a bacterias no cultivables presentes en otras muestras. Además, estudios recientes indican que los ambientes extremos (suelos tropicales, fondos marinos, fuentes termales, terrenos contaminados, ambientes acidófilos, alcalófilos o halófilos, etc.), se encuentran entre los ambientes de mayor diversidad microbiana descrita (Borneman y Triplett, 1997; Whitman *et al.*, 1998; Madrid *et al.*, 2001). Esto se debe fundamentalmente a la presencia de elementos móviles en microorganismos que contribuyen a la plasticidad y al poder adaptativo que da lugar a esta diversidad. Hoy en día estos elementos constituyen una

Tabla 1. Número de especies estimadas (y accesibles) en la Naturaleza.

| Organismo | Especies estimadas | Accesibilidad (% del total) |
|-----------------------------------|--------------------|-----------------------------|
| Animales (mamíferos, aves, peces) | $3,5 \times 10^4$ | >90 |
| Artrópodos/invertebrados | 10^6-10^7 | 10 |
| Nemátodos | 5×10^5 | 3 |
| Plantas superiores | $2,7 \times 10^5$ | >90 |
| Algas | 10^4-10^5 | 70 |
| Hongos | $1,5 \times 10^6$ | 5 |
| Bacterias | 10^4-10^5 | 1-10 |
| Arqueas | 10^5-10^6 | 0,1-1 |
| Virus | 10^5-10^6 | 4 |

herramienta molecular de primer orden en la identificación y discriminación de cepas bacterianas.

Programas de secuenciación de microorganismos

Desde el primer genoma bacteriano secuenciado (*Haemophilus influenzae*, 1995) se ha publicado la secuencia de más de un centenar de microorganismos y existen nuevos proyectos en marcha. El primer microorganismo secuenciado se eligió de acuerdo a su importancia médica. Este enfoque se ha ido ampliando a otros microorganismos y en la actualidad ya están disponibles los genomas de bacterias modelo importantes en otras áreas de la biología tanto básicas como aplicadas. Una de las principales utilidades de los genomas es su utilidad como fuente de procesos y productos. Sin embargo, la secuenciación total de un genoma es un proceso caro y costoso como estrategia de búsqueda de nuevas actividades, por lo que se hace necesario el desarrollo de estrategias más directas. Es en este punto donde las genotecas metagenómicas representan una alternativa reciente y de gran proyección como fuente de información genética de organismos tanto cultivables como no cultivables.

Metagenoma: DNA como fuente de recursos biotecnológicos

Los métodos actuales para el descubrimiento de nuevas enzimas se basan fundamentalmente en conseguir de forma colectiva todos los genomas microbianos presentes en un determinado hábitat y su posterior clonación en complejas genotecas. Handelsman y colaboradores (1998) han creado el término "metagenoma" para definir este DNA. El DNA de todo un conjunto de microorganismos es extraído de una fuente natural (terrestre o acuática) y puesto en complejas genotecas de DNA recombinante para subsiguientes protocolos de búsqueda (*screening*). Esta estrategia es factible y está teniendo un gran impacto en la literatura reciente (Cottrell *et al.*, 1999; Henne *et al.*, 2000; Rondon *et al.*, 2000; MacNeil *et al.*, 2001; Lorenz *et al.*, 2003; Schloss y Handelsman, 2003; Eggert *et al.*, 2004). A pesar de los enormes recursos derivados de la diversidad molecular presente en un DNA metagenómico, el número de actividades enzimáticas novedosas es todavía limitado, probablemente debido a problemas tecnológicos y logísticos derivados de (i) la utilización de métodos de búsqueda adecuados y (ii) el éxito en la clonación de los fragmentos de ADN (Handselman *et al.*,

1998). Es por ello que el éxito de estos métodos depende en gran medida de la capacidad de clonar fragmentos de DNA suficientemente grandes en los vectores apropiados y de su capacidad para expresar la información genética en hospedadores heterólogos de forma eficiente (Rondon *et al.*, 2000; Lorenz *et al.*, 2001) y la búsqueda de protocolos adecuados de *screening*. Optimizando ambos factores, se han podido encontrar nuevas moléculas funcionales con una homología de secuencia en el rango 0–40% a cualquier gen publicado en las bases de datos (Lorenz *et al.*, 2001). Los aspectos más destacados de las técnicas metagenómicas son (i) el acceso al gran reservorio genético (metagenoma) de los microorganismos no cultivables y (ii) la identificación de genes y enzimas novedosas que pueden trasladarse a nuevas aplicaciones biotecnológicas. La Figura 1 muestra de forma esquemática las diferencias entre trabajar con ADN de organismos individuales y ADN metagenómico para el aislamiento de nuevas enzimas. A grandes rasgos, el potencial de éste último radica en que se puede trabajar con genomas de entre 1 y 15.000 organismos (la mayoría de ellos no cultivables) y 10^5 – 10^9 clones/día.

Genotecas metagenómicas: información genética vs. funcional

La creación de genotecas metagenómicas puede realizarse bien para obtener el máximo grado de información genética o bien para obtener enzimas o proteínas con un alto grado de expresión y funcionalización. Desde este punto de vista conviene señalar tres tipos de genotecas:

1) Genotecas en plásmidos, cósmidos y/o fósmodos. Una de las estrategias más directas para la clonación del DNA metagenómico, lo constituyen la utilización de plásmidos, cósmidos y fósmodos utilizando, normalmente, *E. coli* como huésped heterólogo (Lorenz *et al.*, 2003). Estas genotecas se basan en la clonación directa de fragmentos de DNA de pequeños (3–10 kbp: plásmidos) o grandes (15–120 kbp: cósmidos o fósmodos) fragmentos de DNA en *E. coli*. Aunque presentan problemas asociados al nivel de expresión y toxicidad celular (durante la expresión génica), éstas se han demostrado útiles para la búsqueda de una alta variedad de enzimas: 4-hidroxitirato dehidrogenasas, lipasas, esterases, proteasas, amilasas, oxigenasas, oxidoreductasas, etc. (Lorenz *et al.*, 2003).

2) Genotecas de cromosomas artificiales. Una de las estrategias para la clonación de fragmentos de DNA se basa en la construcción y

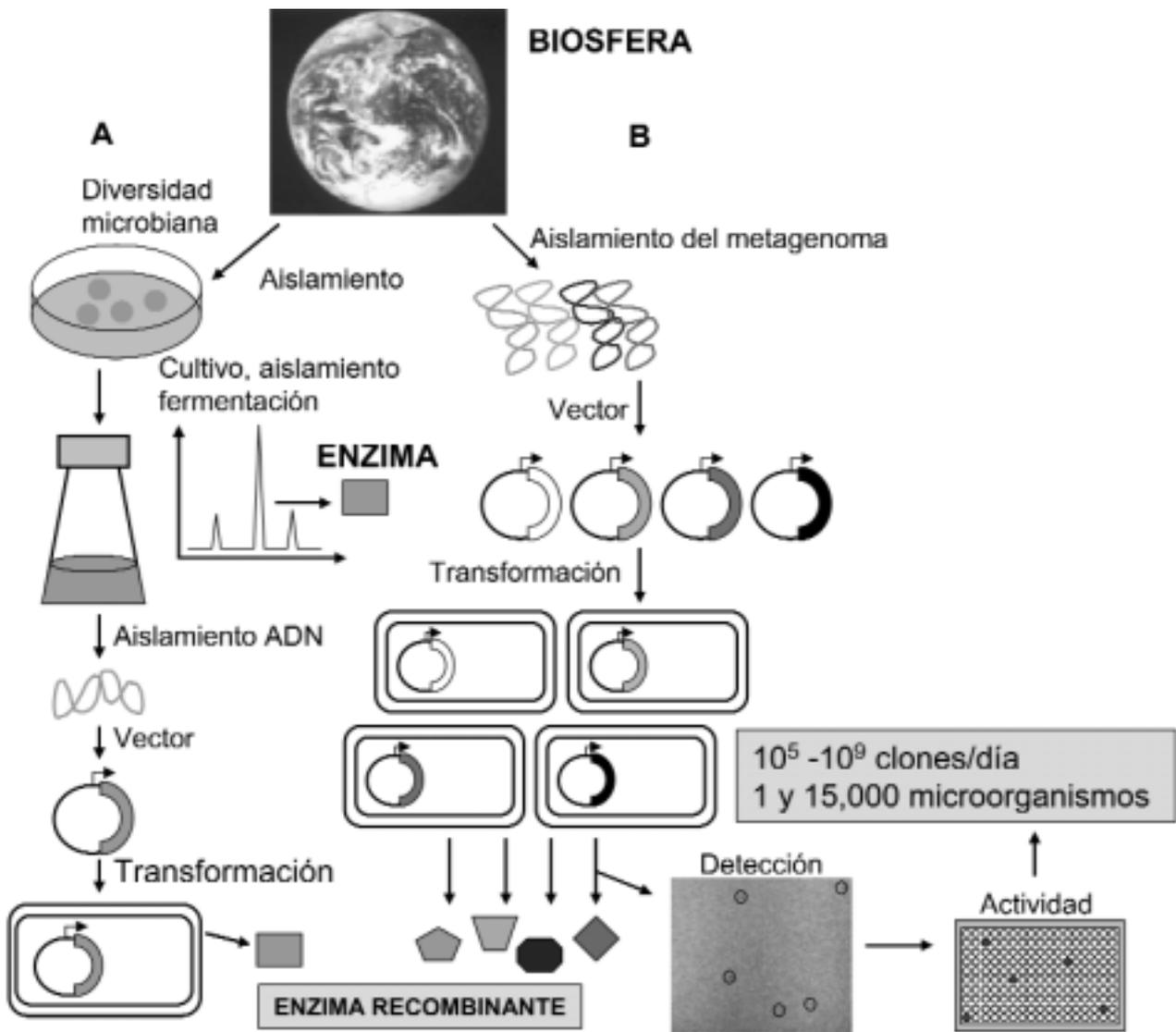


Figura 1. Obtención de enzimas recombinantes a partir de ADN de organismos puros (A) o a partir de ADN metagenómico (B). (Lorenz *et al.*, 2000).

expresión en *E. coli* de genotecas de cromosomas artificiales bacterianos (*Bacterial Artificial Chromosome Library*) conteniendo fragmentos de DNA genómico (5–120 Kpb) aislado directamente de los microorganismos de un ecosistema particular: el metagenoma. El escrutinio de estas genotecas para la identificación de metabolitos o enzimas de interés ha permitido encontrar toda una lista de actividades enzimáticas descubiertas: lipasas, esterases, amilasas, proteasas, oxidoreductasas, L-amino ácido oxidasas, hidroxilasas, dioxigenasas, nucleasas, quitinasas, xilanasas, etc (Gillespie *et al.*, 2002; Rondon *et al.*, 2000). Para una representación lo más completa posible del metagenoma de una muestra de suelo, por ejemplo, se ha estimado que se necesitan unos 10⁶ clones de cromosomas artificiales bacterianos con un tamaño de inserto medio de 100 kb (Handselman

et al., 1998). Genotecas de este tamaño no son fáciles de construir, ni de conservar, así como tampoco de realizar búsquedas de actividades enzimáticas concretas. Estos vectores replican como copia única en *E. coli* y pueden mantener de forma estable insertos de hasta 300 kb (Kouker y Jaeger, 1987). En la actualidad hay descritos vectores BACs mejorados capaces de aceptar grandes insertos permitiendo genotecas genómicas más completas (MacNeil *et al.*, 2001). Esta estrategia de búsqueda incrementará, en un futuro próximo, el gran número de compuestos naturales relevantes por su aplicación industrial. La construcción de genotecas en cromosomas artificiales (*BAC-libraries*) permite que junto con un adecuado sistema de *screening* lo suficientemente selectivo pueda usarse como estrategia de acceso a nueva información genética de microorganismos cultiva-

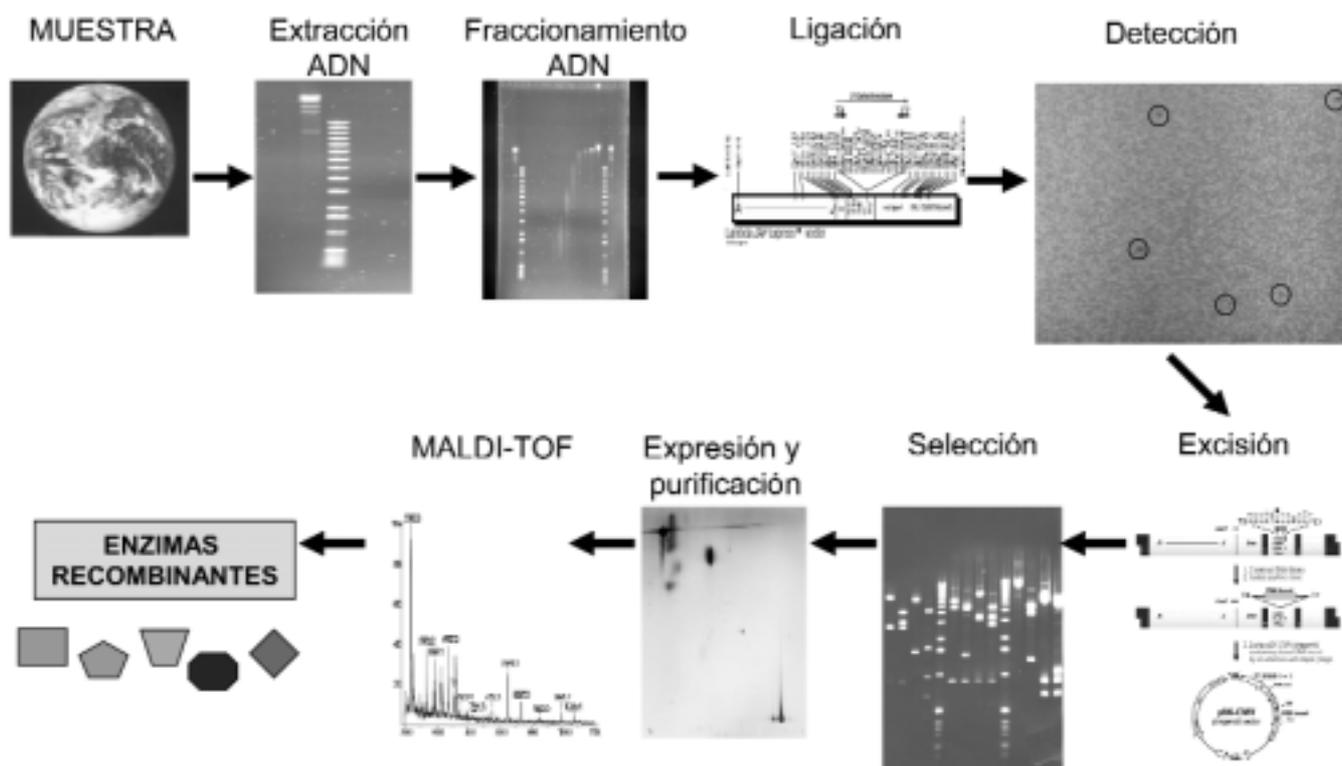


Figura 2. Metodología para el aislamiento de nuevas enzimas a partir de genotecas metagenómicas creadas en fagos *lambda*.

bles o no presentes en un determinado nicho.

3) Genotecas en fagos *lambda*. Aunque las genotecas en BACs proporcionan una gran información genómica, su principal inconveniente derivada de la baja expresión que, en ocasiones, se

produce de un determinado gen localizado en un fragmento de DNA de gran tamaño (> 20–30 Kpb). Este problema puede ser resuelto mediante la expresión en fagos utilizando vectores *lambda* capaces de aceptar insertos de hasta 12 Kb tanto de procariontes como de eucariotes. La ventaja de este sistema, es una alta expresión de la actividad deseada, así como la facilidad de construcción y posterior expresión en un huésped, por ejemplo, *E. coli* (ver Figura 2). El escrutinio de genotecas *lambda* ha permitido encontrar nuevas enzimas: lipasas, esterases, endoglucanasas, amilasas, ciclodextrinasas, oxidoreductasas, lacasas, quitinasas, DNAsas, etc. (Cottrell *et al.*, 1999; Ferrer *et al.*, no publicado).

Genotecas metagenómicas: metodología

Debido a la inherente fragilidad del DNA, el aislamiento de pequeños o grandes fragmentos con los que construir la genoteca del metagenoma exige la utilización de protocolos secuenciales capaces de separar las muestras de DNA por diferentes propiedades físico-químicas, así como en distintos tamaños (Mac Neil *et al.*, 2001). Una vez determinadas las muestras así como las condiciones de extracción de un DNA de suficiente calidad se comienza la construcción de la(s) genoteca(s),

Manuel Ferrer es Licenciado en Ciencias Químicas por la Universidad de Granada y Doctor en Química por la Universidad Autónoma de Madrid. Realizó su Tesis Doctoral sobre utilización de lipasas comerciales y aisladas de hongos filamentosos para la síntesis de ésteres de azúcares y ácidos grasos bajo la dirección del Prof. A. Ballesteros Olmo y del Dr. F.J. Plou Gasca, y una estancia postdoctoral en el Departamento de Microbiología del Prof. K.N. Timmis en el *National Research Center for Biotechnology* (GBF), Alemania, donde estudió la búsqueda de nuevas enzimas de interés industrial y medioambiental a partir de genotecas metagenómicas y la modificación de la temperatura de crecimiento de organismos mediante la expresión de chaperonas de organismos psicrófilos. En 2003 se incorporó al Departamento de Biocatálisis aplicada del Instituto de Catálisis (CSIC) de Madrid donde es Investigador del Programa "Ramón y Cajal". Su investigación actual se centra en el descubrimiento de enzimas de microorganismos no cultivables utilizando herramientas metagenómicas, así como, la adaptación molecular, funcional y estructural de enzimas aisladas de ambientes extremos.

que tendrá que hacerse, naturalmente, en un huésped heterólogo, esencialmente *E. coli*, *Bacillus* sp. o *Streptomyces lividans*. Como material de partida se necesita la mínima cantidad de muestra de la que se pueda aislar mediante protocolos establecidos entre 200 y 500 mg de DNA total con fragmentos de DNA del metagenoma del mayor tamaño posible (entre 50 y 200 Kb). Hasta 1Gbp de DNA del metagenoma presente en las muestras (el genoma medio de un microorganismo ronda los 3–4 Mbp), gran parte de él representado por fragmentos de microorganismos no cultivables, con una media de tamaño de inserto superior a las 30 kb (BACs) o inferior a 20 kb (fagos *lambda*), es sometido a continuación a procesos de búsqueda enzimática. Evidentemente el aislamiento de nuevas enzimas conlleva el desarrollo de 8 nuevos sistemas de screening, bien sea para la búsqueda de nuevas enzimas de amplio espectro de reactividad, o bien enzimas selectivas o específicas, punto este último que requiere la puesta a punto de sustratos más específicos (ver Figura 3).

Futuro de las herramientas metagenómicas: cohesión con técnicas tradicionales

Los hechos arriba expuestos auguran la existencia de nuevas enzimas en distintos ecosistemas. La búsqueda de estas enzimas puede ser combinada con técnicas de evolución dirigida, herramienta, que se ha mostrado muy eficaz para el desarrollo de nuevas actividades partiendo de enzimas ya conocidas (Alcalde, 2003). Esta técnica consiste en incrementar o variar las propiedades de un biocatalizador mediante ciclos sucesivos de mutagénesis aleatoria o recombinación del DNA. De esta forma se consigue acortar el proceso de la selección natural desde miles de millones de años hasta tan sólo semanas o meses de trabajo en el laboratorio. Si esta herramienta se combina con la diversidad enzimática aún sin explorar, nos podemos encontrar en una situación ventajosa que implicaría un avance sustancial en la Biotecnología y Microbiología actuales. Proyectos conjuntos que contemplen (i) el desarrollo de nuevos métodos de cultivo de microorganismos; (ii) el descubrimiento de nuevas enzimas a partir de todos los genomas de organismos (metagenoma) de una muestra y (iii) su posterior mejora mediante técnicas de evolución dirigida o ingeniería de la reacción y del biocatalizador (por ejemplo, técnicas de inmovilización), son por tanto proyectos ambiciosos y de gran interés, tanto de un punto de vista básico (conocimiento de nuevas estructuras y actividades), como de un punto de vista aplicado

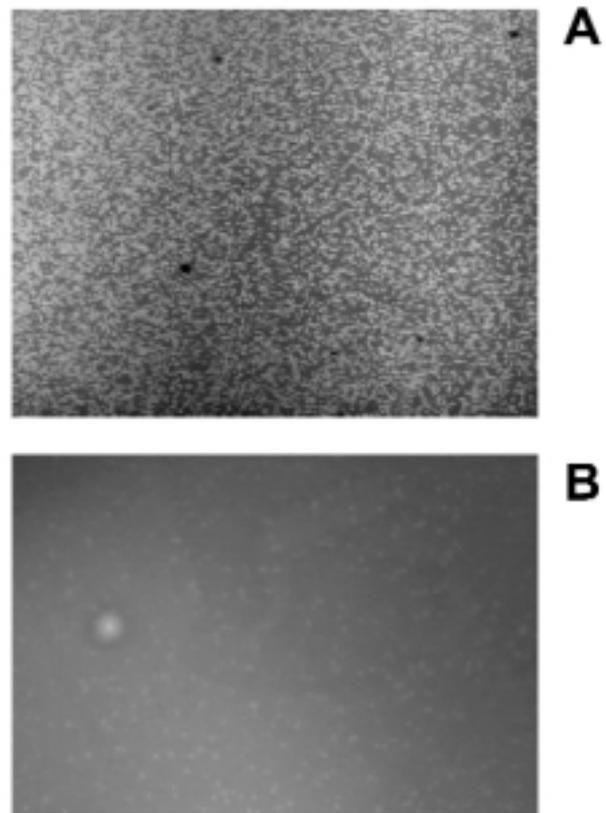


Figura 3. Detección de actividad esterasa (A) y celulasica (B) en genotecas metagenómicas.

(nuevas actividades que pueden trasladarse a una industria química verde libre de subproductos y respetuosa con el medioambiente; ver Figura 4). Éstas son, propuestas multidisciplinarias que engloban áreas tan diversas como microbiología, biotecnología, estructura de proteínas, bioquímica y química orgánica, así como en líneas de investigación emergentes, por ejemplo, genómica y proteómica.

Por último, señalar que estudios recientes reve-



Figura 4. Combinación de técnicas metagenómicas, evolución dirigida, e ingeniería del biocatalizador para el aislamiento de nuevas y mejores enzimas.

lan que el futuro de la microbiología será la creación de bancos de genes, es decir, organismos heterólogos (*E. coli*, *Bacillus*, *Streptomyces*, etc.) que expresen un gen particular que codifique una o varias enzimas con actividad determinada, obtenidas a través de búsquedas en genotecas metagenómicas (u otras herramientas similares). Estos bancos coexistirán con los bancos de microorganismos que actualmente existen (DSMZ, ATCC, etc.). La creación de estos bancos de genes será el punto de partida para acceder a la información genética, estructural y funcional de genes obtenidos de microorganismos no cultivables, cuyo potencial es ilimitado.

Bibliografía

- Alcalde M (2003) Evolución molecular dirigida. *Investigación y Ciencia* 11:69-77.
- Borneman J, Triplett EW (1997) *Appl Environ Microbiol* 63:2647-2653.
- Cottrell MT, Moore JA, Kirchman DL (1999) *Appl Environ Microbiol* 65:2553-2557.
- Eggert T, Leggewie C, Puls M, Streit W, Van Pouderoyen G, Dijkstra BW, Jaeger K-E (2004) *Biocatal Biotrans* 22:139-144.
- Gillespie DE, Brady SF, Betterman AD, Cianciotto NP, Liles MR, Rondon MR, Clardy J, Goodman RM, Handelsman J (2002) *Appl Environ Microbiol* 68:4301-4306.
- Handselman J, Rondon MR, Brady SF, Clardy J, Goodman, RM (1998) *Chem Biol* 5:245-249.
- Henne A, Smith RA, Bömeke M, Gottschalk G, Daniel R (2000) *Appl Environ Microbiol* 66:3113-3116.
- Koonin EV, Wolf YI, Karen GP (2002) *Nature* 420:218-223.
- Kouker G, Jaeger KE (1987) *Appl Environ Microbiol* 53:211-213.
- Lorenz P, Liebeton K, Niehaus F, Eck J. (2002) *Curr Opin Biotechnol* 13:572-577.
- Lorenz P, Liebeton K, Niehaus F, Schleper C, Eck J (2003) *Biocatal Biotrans* 21:87-91.
- MacNeil IA, Tiong CL, Minor C, August PR, Grossman TH, Loiacono KA, Lynch BA, Phylips T, Narula S, Sundaramoorthi R, Tyler A, Aldredge T, Long H, Gilman M, Holt D, Osburne MS (2001) *J Mol Microbiol Biotechnol* 3:301-308.
- Madrid VM, Aller JY, Aller RC, Chistoserdov AY (2001) *FEMS Microbiol Ecol* 37:197-209.
- Muyzer G (1999) *Curr Opinion Microbiol* 2:317-322.
- Muyzer G, Waal EC, Uitterlinden AG (1993) *Appl Environ Microbiol* 59:695-700.
- Ogawa J, Shimizu S (1999) *Trends Biotechnol* 17:13-21.
- Pace NR (1997) *Science* 276:734-740.
- Rondon MR, August PR, Betterman AD, Brady SF, Grossman TH, Liles MR, LoiaconoKA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman R (2000) *Appl Environ Microbiol* 66:2541-2547.
- Roselló-Mora R, Amann R (2001) *FEMS Microbiol Rev* 25:39-67.
- Schloss PD, Handelsman J (2003) *Curr Opin Biotechnol* 14:303-310.
- Suzuki MT, Rappe MS, Haimberger ZW, Winfield H, Adair N, Strobel J, Giovannoni SJ (1997) *Environ Microbiol* 63:983-989.
- Villadas PJ, Martínez-Abarca F, Toro N (2002) *Anal Biochem* 300:101-103
- Whitman WB, Coleman DC, Wiebe WJ (1998) *Proc Natl Acad Sci USA* 95:6578-6583.