

Análisis de datos de microarreglos de ADN

Parte I: Antecedentes de la tecnología y diseño experimental

Jamilet Miranda, ✉ Ricardo Bringas

Centro de Ingeniería Genética y Biotecnología, CIGB
Ave. 31 e/ 158 y 190, Cubanacán, Playa, AP 6162, CP 10600, Ciudad de La Habana, Cuba
E-mail: ricardo.bringas@cigb.edu.cu

RESUMEN

Los microarreglos de ADN han emergido como la tecnología más utilizada para la cuantificación masiva de la expresión de genes y han sido aplicados a temas muy diversos entre las investigaciones biológicas en los últimos años. Un elemento fundamental para la aplicación exitosa de esta tecnología es el conocimiento de los pasos a seguir para la obtención y análisis de los datos de expresión. En el presente trabajo se hace un recuento del surgimiento de la tecnología, su evolución y algunas de sus aplicaciones más comunes, se subraya la necesidad de definir claramente en el diseño, los objetivos del experimento y se exponen las diferentes fuentes de variabilidad a tener en cuenta en el diseño y los tipos de diseño más comunes.

Palabras clave: microarreglos de ADN, expresión de genes, diseño experimental

Biotecnología Aplicada 2008;25:82-89

ABSTRACT

Analysis of DNA microarray data. Part I: Technological background and experimental design. DNA microarrays have emerged as the most widely used technology for the massive quantification of gene expression and have been applied to a very diverse range of topics in molecular biology research over the last several years. One key element for a successful application of this technology is a thorough understanding of the steps to be followed in order to obtain and analyze expression data. In the present article we review the origins of the technology, its evolution and some of its more common applications, highlighting the importance of a clear definition of the objectives for the design of the experiment, the different sources of variability to be considered and the most common experimental setups.

Keywords: DNA microarrays, gene expression, experimental design

Introducción

La secuenciación completa de genomas ha develado la estructura primaria de las cadenas de ADN de los cromosomas que contienen todos los genes de un organismo, así como todos los componentes que intervienen en su regulación. Ello constituye la información primaria, cuyas variaciones entre individuos definen la forma diferenciada en que estos interactúan con el entorno, por lo que resulta de vital importancia el estudio de estas diferencias en la solución de problemas relacionados con la salud humana, la salud animal y la agricultura. Por otro lado, la disponibilidad de la secuencia completa de genomas ha permitido el desarrollo de tecnologías capaces de analizar en un solo experimento todos los elementos identificados de un genoma. Una de estas tecnologías es la de microarreglos de ADN, colecciones de segmentos de ADN que son fijados a una superficie sólida para su utilización en la cuantificación de niveles de ARN o ADN en muestras biológicas. Estos segmentos o sondas de ADN se diseñan de forma que sean complementarios a las regiones del ADN o ARN que se desea cuantificar, lo que permite medir los niveles de expresión de cada gen de acuerdo con la cantidad de ADN o ARN que se hibridan con las sondas impresas en el arreglo.

Los microarreglos de ADN se han convertido en la tecnología más ampliamente utilizada para generar perfiles de expresión a escala genómica. Sin embargo, su uso no se ha limitado al estudio de perfiles de expresión de genes, sino que se ha extendido a estudios sobre la variabilidad genética entre individuos, con el

empleo de microarreglos diseñados para el estudio de Polimorfismos genéticos de un solo nucleótido (SNPs, del inglés *Single Nucleotide Polymorphisms*) [1]. Por otra parte, su combinación con técnicas como la inmunoprecipitación de cromatina ha permitido la identificación de regiones reguladoras reconocidas por factores de transcripción, mediante el uso de arreglos que contienen sondas complementarias a las regiones promotoras de todos los genes conocidos de un organismo [2]. Otra aplicación de esta tecnología ha sido el estudio del empalme diferencial mediante el diseño de arreglos con sondas específicas a los diferentes exones que contiene cada gen [3].

La tecnología de referencia, al permitir el conocimiento de la expresión de los genes en diferentes condiciones experimentales, ha posibilitado el estudio de las bases y mecanismos moleculares de múltiples enfermedades tales como: infecciones virales [4, 5], esquizofrenia [6-10], cáncer de próstata [11-15] y cáncer de mama [16, 17]. También se han realizado experimentos de farmacogenómica dirigidos al estudio de los cambios a nivel molecular inducidos por el uso de medicamentos en diferentes afecciones. Algunos ejemplos son: el estudio de la resistencia al tamoxifeno en cáncer de mama [18] y el efecto de IL2 en cáncer [19-22].

La aplicación de los microarreglos ha demostrado además ventajas sobre otros métodos tradicionalmente usados para el diagnóstico de enfermedades complejas. Alizadeh *et al.* [23] identificaron dos subtipos de linfomas de células B, que resultaban muy difíciles de diferenciar a partir de pruebas histológicas, las que a

1. Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, *et al.* Accessing genetic information with high-density DNA arrays. *Science* 1996;274:610-4.
2. Lee TI, Rinaldi MJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 2002;298:799-804.
3. Castle J, Garrett-Engle P, Armour CD, Duenwald SJ, Loerch PM, Meyer MR, *et al.* Optimization of oligonucleotide arrays and RNA amplification protocols for analysis of transcript structure and alternative splicing. *Genome Biol* 2003;4:R66.
4. Wang X, Yuan ZH, Zheng LJ, Yu F, Xiong W, Liu JX, *et al.* Gene expression profiles in an hepatitis B virus transfected hepatoblastoma cell line and differentially regulated gene expression by interferon- α . *World J Gastroenterol* 2004;10:1740-5.
5. Yang J, Bo XC, Yao J, Yang NM, Wang SQ. Differentially expressed cellular genes following HBV: potential targets of anti-HBV drugs?. *J Viral Hepat* 2005;12:357-63.
6. Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, *et al.* Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. *Proc Natl Acad Sci USA* 2001;98:4746-51.
7. Vawter MP, Barrett T, Cheadle C, Sokolov BP, Wood WH, Donovan DM, *et al.* Application of cDNA microarrays to examine gene expression differences in schizophrenia. *Brain Res Bull* 2001;55:641-50.

su vez definían un conjunto de genes cuyos perfiles de expresión mostraban un claro patrón asociado a la supervivencia.

A la par del desarrollo de la tecnología ha sido inevitable la continua evolución, la adaptación y el desarrollo de métodos estadísticos y matemáticos para garantizar los análisis de las matrices de valores de expresión, que son generadas como resultado de su aplicación y tienen como característica que el número de variables o genes (g) es mucho mayor que el número de muestras o tejidos (n) que se analizan; $g \ll n$.

El tipo de método estadístico que se utiliza en el análisis depende de los objetivos definidos en el experimento, aunque muchas veces para realizar la interpretación de los resultados se hace necesario usar combinaciones de métodos por la complejidad de los sistemas biológicos. La comparación, la predicción y el descubrimiento de clases experimentales [24-26] son los objetivos más frecuentes que se plantean en este tipo de experimentos.

En el presente trabajo se aborda el uso de los microarreglos de ADN en el estudio de los perfiles de expresión de genes a escala de genomas completos y se hace énfasis en los antecedentes de la tecnología y en la etapa de diseño experimental.

Antecedentes de la tecnología

Surgimiento

Los comienzos de la tecnología de referencia pueden ubicarse a mediados de los 90 con los estudios de Schena *et al.* [27] y Lockhart *et al.* [28].

Schena *et al.* [27] describieron por primera vez el desarrollo de un microarreglo para el monitoreo de la expresión de múltiples genes en paralelo. Las muestras de una placa de 96 pocillos fueron impresas en cristales de microscopía en una área de 3.5 x 5.5 mm. Las muestras una vez depositadas sobre el cristal fueron tratadas química y térmicamente para fijar el ADN a la superficie y desnaturalizarlo. Se estudió la expresión de un total de 45 genes de *Arabidopsis thaliana*, más 3 genes control de otros organismos. Cada muestra se duplicó en pocillos adyacentes para validar la reproducibilidad de los procesos de impresión e hibridación. A partir del ARN total de *A. thaliana*, por transcripción reversa se obtuvo ADN complementario (ADNc) marcado por fluorescencia. Se usaron dos colores de marcaje para poder cuantificar 2 muestras a la vez. Las muestras analizadas, provenientes de tejidos de hojas y raíces, reportaron un total de 27 genes diferencialmente expresados entre las 2 muestras. En este trabajo pionero, se expusieron algunos de los principios fundamentales que dieron lugar al desarrollo de esta tecnología, en particular, los microarreglos de ADNc y el uso de la doble fluorescencia en la cuantificación relativa de 2 muestras en un solo experimento.

Lockhart *et al.* [28] desarrollaron una tecnología que permitía medir la expresión de miles de genes en paralelo. El método desarrollado se basa en la cuantificación de la proporción relativa del ARNm por hibridación de poblaciones enteras de ARNm sobre arreglos de alta densidad de sondas de ADN. Los microarreglos contenían miles de oligonucleótidos de longitud 20, los que fueron diseñados como sondas específicas a regiones 3' de genes humanos conocidos. La síntesis se rea-

lizó *in situ* sobre la superficie de placas de vidrio, combinando métodos de la química de ácidos nucleicos y de fotolitografía [29, 30], similares a los usados en la industria de la microelectrónica, que permiten la síntesis en paralelo de todos los oligonucleótidos del arreglo. Cada reacción de síntesis se realizó en un área de 50 μm de lado, por lo que en una superficie de poco más de 1.6 cm^2 se imprimieron más de 65 000 sondas diferentes. Los oligonucleótidos fueron sintetizados en pares. Cada par lo constituye una secuencia de 20 nucleótidos donde la primera es una secuencia complementaria perfecta a un segmento de un ARN mensajero (ARNm) y se conoce como PM (*Perfect Match*) y el segundo componente del par es otra secuencia de 20 bases que se diferencia de la primera en solo un nucleótido en una de las posiciones centrales y por tanto su complementariedad respecto al mensajero no es perfecta, se conoce como MM (*Mis Match*). El análisis de los pares PM-MM representa un control interno, ya que en caso de una correcta hibridación, el componente PM del par debe ofrecer una señal más intensa que el componente MM. La forma más simple utilizada para estimar el nivel de expresión de cada gen es promediar las diferencias PM-MM después de haberle realizado una corrección del fondo a cada valor de PM y MM, aunque otros métodos se han utilizado para estimar estos valores [31-33].

En el trabajo de Lockhart *et al.* [28], para la obtención de las imágenes de los microarreglos se usó un microscopio de barrido confocal especialmente diseñado al efecto. Las imágenes se obtuvieron con una resolución de 7.5 μm , lo que permite obtener como promedio unos 45 valores de intensidad en el área de 50 x 50 μm que corresponde a cada sonda impresa. Estos valores de intensidad se combinan para obtener un valor por cada celda o sonda impresa en el arreglo. Mientras más valores sean colectados por cada celda, mayor será la precisión de la intensidad resultante. El número de valores de intensidad que se obtendrán por cada celda depende del tamaño de la celda y de la resolución del equipo que se emplee para su lectura. En la figura 1 se ilustra la influencia de la relación tamaño de la celda/resolución del escáner. En el caso de que esta relación sea de 10:1

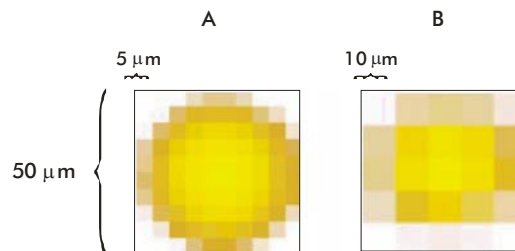


Figura 1. Ilustración de la lectura de las señales de un microarreglo con diferente resolución del escáner. En ambos casos se representa una señal de aproximadamente 50 μm de diámetro, en a) se ilustra la lectura a una resolución de 5 μm y en b) se ilustra la lectura a una resolución de 10 μm . En ambos casos las lecturas que corresponden a los bordes de las manchas no reflejarán un valor que corresponda solo con la intensidad de la señal sino con el área que ocupa la señal dentro del cuadrado donde se efectuó la lectura. Ello indica que los valores más exactos de intensidad corresponden a los valores en el interior de la celda, por lo que se evidencia que en el caso de una mayor resolución se obtiene una lectura mucho más exacta de la intensidad de la señal.

8. Vawter MP, Ferran E, Galke B, Cooper K, Bunney WE, Byerley W. Microarray screening of lymphocyte gene expression differences in a multiplex schizophrenia pedigree. *Schizophr Res* 2004;67:41-52.

9. Tsuang MT, Nossova N, Yager T, Tsuang MM, Guo SC, Shyu KG, *et al.* Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: A preliminary report. *Am J Med Genet B Neuropsychiatr Genet* 2005;133B:1-5.

10. Glatt SJ, Everall IP, Kremen WS, Corbeil J, Sasik R, Khanlou N, *et al.* Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. *Proc Natl Acad Sci USA* 2005;102:15533-8.

11. Luo J, Duggan DJ, Chen Y, Sauvageot J, Ewing CM, Bittner ML, *et al.* Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. *Cancer Res* 2001;61:4683-8.

12. Welsh JB, Sapinoso LM, Su AI, Kern SG, Wang-Rodriguez J, Moskaluk CA, *et al.* Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res* 2001;61:5974-8.

13. Ashida S, Nakagawa H, Katagiri T, Furihata M, Iizumi M, Anazawa Y, *et al.* Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: genome-wide gene-expression profiles of prostate cancers and PINs. *Cancer Res* 2004;64:5963-72.

14. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci USA* 2004;101:811-6.

15. Zhao H, Lai F, Nonn L, Brooks JD, Peehl DM. Molecular targets of docetaxin in human prostatic stromal cells. *Prostate* 2005;62:400-10.

16. Van 't Veer LJ, Dai H, Van de Vijver MJ, He YD, Hart AA, Mao M, *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-6.

(Figura 1a), el promedio de valores o píxeles que se obtiene por cada sonda en el arreglo, es de 100, mientras que si esta relación es de 5:1 este promedio será de 25. Los valores que van a reflejar mejor la intensidad de la señal, serán aquellos que se encuentran más cercanos al centro de la mancha, que tendrán señal en el total del área de la cuadrícula que ocupan. La proporción de estos valores será mayor al usar una mayor resolución como se evidencia en la figura 1.

En un experimento inicial, Lockhart *et al.* [28] diseñaron cientos de pares de sondas por cada uno de los genes que evaluaron, con la intención de medir su sensibilidad y especificidad cuando son usados en una muestra compleja de ARN celular. De este estudio se derivaron reglas para la selección de las sondas. En un segundo estudio donde se analizó la expresión de 118 genes, se diseñaron un promedio de 300 pares PM-MM de sondas por gen, que fueron seleccionadas de los extremos 3' de las regiones codificadoras. Del total de sondas de cada gen se seleccionaron al azar 10 conjuntos de 20 pares PM-MM. Los cambios en los patrones de hibridación fueron comparados en los 10 conjuntos aleatorios y en el total de sondas diseñadas; se concluyó que un conjunto de 20 sondas es suficiente para medir cambios de expresión incluso para ARN de baja expresión. Este trabajo sentó las bases para el desarrollo de la tecnología de la firma Affymetrix [34], líder en la producción de microarreglos de ADN. Estas tecnologías han tenido un continuo desarrollo, con un aumento paulatino de la densidad de los *chips*, que en la actualidad es de 5 micras, y de la resolución del escáner para su lectura.

El uso de microarreglos a escala genómica

DeRisi *et al.* [35] utilizaron microarreglos que contenían aproximadamente 6400 secuencias diferentes de ADN correspondientes a igual número de marcos abiertos de lectura, identificados en el genoma de la levadura *Saccharomyces cerevisiae*. En ese trabajo se estudió la expresión de casi la totalidad de los genes de este organismo en un cultivo con medio rico en glucosa. El cultivo incluyó un periodo inicial de fermentación anaeróbica en el cual, a medida que era agotada la glucosa como fuente de carbono, se iniciaba un proceso aeróbico que empleaba etanol como fuente de este nutriente.

La expresión de los genes fue medida en intervalos de 2 horas. Durante la fase anaeróbica, el patrón de expresión de los genes fue bastante estable, con muy pocos genes que mostraban expresión diferencial. Sin embargo, a medida que la concentración de glucosa en el medio disminuyó se observó un mayor número de genes con cambios significativos en sus niveles de expresión, alrededor de 710 genes mostraron un aumento de al menos dos veces en su expresión mientras que en aproximadamente 1030 se observó una reducción de al menos un factor de dos. De estos genes diferencialmente expresados, aproximadamente la mitad eran de función desconocida. Los conjuntos de genes como aquellos relacionados al citocromo C y los involucrados en el ciclo del ácido tricarbóxico (TCA)/glioxilasa se indujeron de forma coordinada según se agotaba la glucosa, mientras que otros genes como los involucrados en la síntesis de proteínas presentaron una reducción coordinada de sus niveles de expresión.

En este trabajo se pudo observar que los conjuntos de genes funcionalmente relacionados podían ser agrupados por la similitud de sus perfiles de expresión e incluso dentro de estos conjuntos de genes fue posible inferir mecanismos de regulación comunes al identificar secuencias reguladoras presentes en sus promotores.

Los resultados obtenidos evidenciaron el valor del uso de métodos de agrupamiento (*clustering*) en el análisis de datos de microarreglos. Otros aspectos evidenciados fueron la correspondencia de los perfiles de regulación de los genes con la presencia de patrones de regulación comunes en las regiones promotoras y la influencia de genes reguladores en los niveles de expresión de los genes que estos regulan.

Diseño experimental

A continuación se abordan los aspectos relacionados con la definición de la pregunta biológica y la elaboración del diseño del experimento. El diagrama que representa los diferentes pasos en el diseño y el análisis de datos de experimentos de microarreglos se muestra en la figura 2.

Definición de la pregunta o interés biológico

El uso de los microarreglos, como es común en las investigaciones, requiere formular las preguntas que se desea responder con cada uno de los experimentos. Una de estas preguntas puede ser, por ejemplo, qué conjunto de genes están diferencialmente expresados entre dos o más condiciones experimentales o si es posible relacionar las muestras experimentales de acuerdo con diferencias o similitudes de sus perfiles de expresión. Tener claridad sobre la pregunta es importante para definir el tipo y el número de muestras por grupos experimentales y llegar a establecer una estrategia de análisis. A continuación debe evaluarse la necesidad de usar esta tecnología y no otra. Esta

17. Van de Vijver MJ, He YD, Van't Veer LJ, Dai H, Hart AA, Voskuil DW, *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.

18. Jansen MP, Foekens JA, van Staveren IL, Dirkszwaiger-Kiel MM, Ritstier K, Look MP, *et al.* Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol* 2005;23:732-40.

19. Diehn M, Alizadeh AA, Rando OJ, Liu CL, Stankunas K, Botstein D, *et al.* Genomic expression programs and the integration of the CD28 costimulatory signal in T cell activation. *Proc Natl Acad Sci USA* 2002; 99:11796-801.

20. Mao M, Biery MC, Kobayashi SV, Ward T, Schimmack G, Burchard J, *et al.* T lymphocyte activation gene identification by coregulated expression on DNA microarrays. *Genomics* 2004;83:989-99.

21. Panelli MC, White R, Foster M, Martin B, Wang E, Smith K, *et al.* Forecasting the cytokine storm following systemic interleukin (IL)-2 administration. *J Transl Med* 2004; 2:17.

22. Kovanen PE, Young L, Al-Shami A, Rovella V, Pise-Masison CA, Radonovich MF, *et al.* Global analysis of IL-2 target genes: identification of chromosomal clusters of expressed genes. *Int Immunol* 2005;17:1009-21.

23. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, *et al.* Distinct type of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.

24. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from array to consolidation and consensus. *Nat Rev Genet* 2006;7:55-65.

25. Simon R, Radmacher MD, Dobbin K. Design of studies using DNA microarrays. *Genet Epidemiol* 2002;23:21-36.

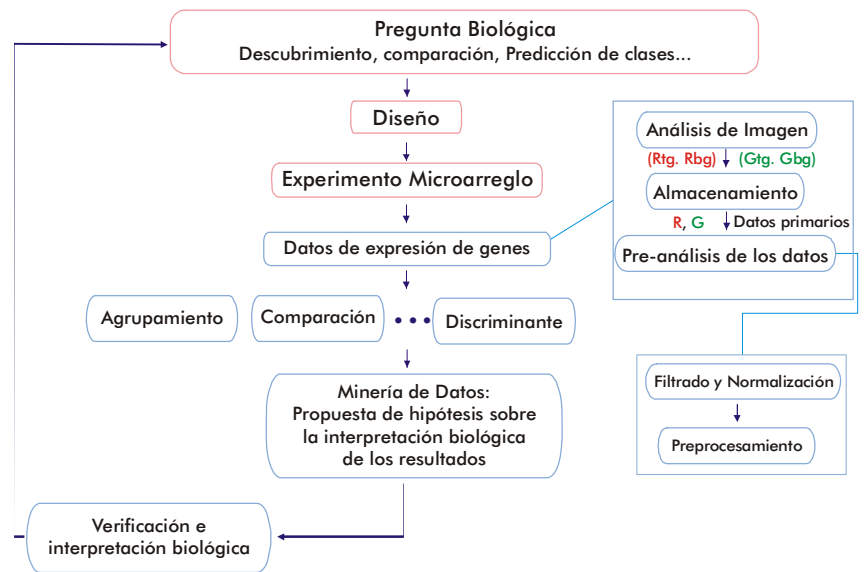


Figura 2. Esquema general para la realización de un experimento de microarreglos y el análisis de sus datos. En rojo se enmarcan los pasos de la primera etapa que requiere: la definición de la pregunta biológica, la elaboración del diseño experimental y la realización del experimento. En azul se enmarca la segunda etapa que comienza con el análisis de las imágenes y concluye con la interpretación de los resultados y la propuesta de verificación biológica.

necesidad se justifica cuando se requiere realizar experimentos a escala genómica donde por lo general se conoce poco del comportamiento individual de los genes y se pretende identificar procesos celulares o metabólicos asociados a la pregunta definida. Los resultados que se obtengan dependen en buena medida de la capacidad de análisis e interpretación de datos, en los que deben desempeñar un papel fundamental el uso del enfoque de la Biología de Sistema.

Diseño

En la planificación de un experimento de microarreglos hay diferentes factores que deben tenerse en cuenta para lograr el diseño más apropiado de acuerdo con las preguntas que se desea responder y los recursos de que se dispone. Como regla general, los experimentos de microarreglos analizan gran número de variables (miles de genes) en un reducido número de condiciones (decenas o cientos de muestras), lo cual obliga a ser rigurosos en la etapa de diseño para aumentar la probabilidad de éxito en los análisis de sus datos. Además, existen varias fuentes de variabilidad de los datos que deben considerarse en el diseño, como son aquellas derivadas de las propias diferencias entre individuos y las asociadas a posibles errores en los múltiples pasos requeridos para el uso de la tecnología. Otro aspecto a considerar es la elección de la tecnología más adecuada, ya que existen tecnologías que se adecuan mejor a un diseño que otras, aunque el análisis costo/beneficio es también determinante. Por estas razones, en el uso de los microarreglos resulta de suma importancia la adecuada planificación y diseño del experimento.

Definición de los objetivos del estudio de acuerdo con la pregunta biológica

La pregunta biológica que se propone responder el experimento es la base para definir los objetivos del estudio. En el caso de estudios por microarreglos se han identificado tres objetivos frecuentes, que se exponen seguidamente asociándolos a posibles preguntas biológicas para su mejor comprensión.

Pregunta 1: ¿Cuales genes se expresan diferencialmente entre dos o más grupos de muestras?

Objetivo: Comparación de clases

Con mayor frecuencia se comparan muestras de tejido sano y tejido enfermo, muestras de células tratadas y no tratadas con un medicamento o muestras de células salvajes y células mutadas. Para responder este tipo de pregunta debe plantearse como objetivo la comparación de clases, que consiste en comparar los perfiles de expresión de diferentes grupos de muestras. Las clases que serán comparadas deben definirse previamente con independencia de los perfiles de expresión. Un ejemplo donde se ilustra este tipo de interrogante es el experimento de Lapointe *et al.* [36], en el cual se estudiaron tejidos sanos y tumorales de pacientes con cáncer de próstata. A partir de clases definidas de acuerdo con características clínicas conocidas tales como grado del tumor, etapa y recurrencia de la enfermedad, los autores obtuvieron conjuntos de genes que cambiaron su expresión significativamente al realizar comparaciones entre estas clases.

En el presente estudio, se realizó un análisis de los datos de este experimento que se obtuvo de la base de datos de microarreglos de Stanford (SMD; <http://smd.stanford.edu/>).

Primeramente se definieron como clases los tipos de tejidos: sano, tumoral y metástasis. A continuación se realizaron comparaciones de tumores vs tejido sano y se determinaron un conjunto de genes diferencialmente expresados a los que se aplicaron métodos de agrupamiento (Figura 3). Posteriormente en comparaciones de clases dos a dos se encontraron genes que van disminuyendo paulatinamente su expresión de tejidos sanos a tumor y de tumor a metástasis. Un ejemplo es el caso del gen SYNPO2, reportado anteriormente como reprimido en cáncer de próstata avanzado [37] y más recientemente como predictor de metástasis en cáncer de próstata [38].

Pregunta 2: ¿A partir del conocimiento del perfil de expresión de un conjunto de genes en diferentes tipos de muestras puede conocerse si una nueva muestra pertenece a uno u otro tipo?

Objetivo: Predicción de clases

La pregunta asociada al objetivo de predicción de clases, se propone encontrar una función multivariada basada en la expresión de genes que permita clasificar con determinada precisión una nueva muestra o tejido en los grupos predefinidos sobre la base de los niveles de expresión de genes clave. Estos trabajos, como regla general, identifican una firma molecular o predictor representada por un conjunto de genes cuyo perfil de expresión permite discriminar con alta probabilidad si una muestra pertenece a uno u otro grupo y limitar el

26. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531-7.

27. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;270:467-70.

28. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 1996;14:1675-80.

29. Fodor SP, Rava RP, Huang XC, Pease AC, Holmes CP, Adams CL. Multiplexed biochemical assays with biological chips. *Nature* 1993;364:555-6.

30. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 1994;91:5022-6.

31. Efron B, Tibshirani R, Storey JD, Tusher V. Empirical bayes analysis of a microarray experiment. *J Am Stat Assoc* 2001;96:1151-60.

32. Li C, Wong WH. Model based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc Natl Acad Sci USA* 2001;98:31-6.

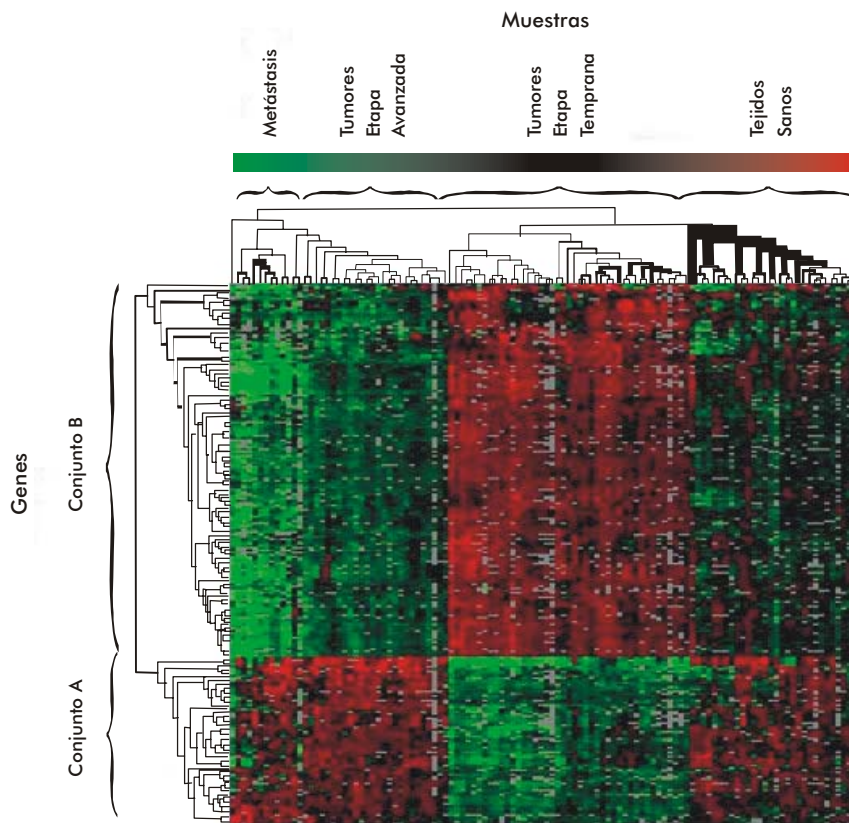


Figura 3. Agrupamiento bidimensional de 242 genes expresados diferencialmente entre tejidos sanos y tumorales de próstata. Los datos fueron obtenidos del estudio de Lapointe *et al.* [35]. Este análisis exploratorio muestra cómo los genes del conjunto A se reprimen en las etapas tempranas del tumor y luego vuelven a aumentar su expresión con el avance de la enfermedad. Los genes del conjunto B muestran un comportamiento contrario.

análisis a los perfiles de expresión de solo aquellos genes incluidos en el predictor, lo que contribuye a su abaratamiento mediante el diseño de arreglos a la medida. Esto constituye una herramienta importante en el diagnóstico y pronóstico de enfermedades. Uno de los ejemplos más elocuentes lo constituye su empleo para predecir el desarrollo de metástasis a partir de la expresión en tumores primarios de mama [39]. Un predictor también puede ser usado para tomar otras decisiones clínicas como la selección de un tratamiento y definir grupos de riesgo.

Pregunta 3: ¿Se pueden definir nuevos subtipos en muestras estudiadas a partir de los patrones de expresión, que puedan asociarse a características encontradas en ellas?

Objetivo: Descubrimiento de clases

El tipo de pregunta asociada al descubrimiento de clases, consiste en la identificación de nuevos subtipos dentro de una población. La diferencia fundamental con los demás objetivos descritos es que en este caso no se usan clases predefinidas. Un ejemplo ilustrativo fue el de Alizadeh *et al.* [23] que estudiando diferentes muestras de linfoma grande y difuso de células B (DLBCL, del inglés *Difuse large B-cell lymphoma*) lograron distinguir dos subgrupos o subclases de DLBCL a partir de la expresión diferencial de cientos de genes diferentes, estas diferencias a nivel molecular estaban acompañadas de diferencias clínicas entre estos subgrupos. Ello sugirió que los dos subgrupos se debían considerar como enfermedades diferentes.

Selección de la plataforma experimental apropiada

Antes de definir el tipo de diseño experimental es recomendable decidir el tipo de plataforma o microarreglo que será utilizado porque ello determinará el número de variables (genes) cuyos perfiles de expresión serán estudiados. Este número influye directamente en el diseño experimental y de forma particular, en la definición del número mínimo de muestras a incluir.

El uso de microarreglos de cobertura genómica, es decir, aquellos que incluyen todos o la mayoría de los genes de un genoma, solo se justifica si la pregunta que se desea responder requiere un análisis masivo de todos los genes del organismo en estudio. Si por el contrario, solo interesa el análisis de un número limitado de genes, otras tecnologías pueden resultar más económicas y precisas, como pudiera ser Reacción en cadena de la polimerasa con transcripción reversa (RT-PCR, del inglés *Reverse Transcription-Polymerase Chain Reaction*).

Otra alternativa sería emplear arreglos que incluyen un número menor de genes, pero relacionados con la pregunta biológica que se pretende responder. Este tipo de arreglos se ofrece por varias firmas comerciales e incluye conjuntos de genes previamente identificados por su relación con una enfermedad, vía metabólica u otra función biológica. Su empleo tiene ventajas por el menor número de variables a analizar y la segura reducción de costos del experimento, su desventaja principal es que el estudio se reduce a un número de genes previamente identificados por su relación con el hecho en estudio y limita las posibilidades de obtener resultados originales. Es también posible

hacer un diseño propio de un arreglo de genes, definido con las herramientas de análisis del investigador, pero este tema no será abordado en el presente trabajo.

Breve descripción de las tecnologías más empleadas

A pesar de que existen varias tecnologías para la producción de microarreglos, las de ADNc [27] y *Affymetrix* [28, 34] son las más utilizadas por los investigadores.

ADNc. Las sondas de ADN de cada gen se imprimen en un soporte sólido, más comúnmente vidrio, en forma de arreglo bidimensional mediante el uso de robots. Se usan 2 muestras de ARN, cada una con un marcaje fluorescente diferente (usualmente con los fluoróforos Cy5 y Cy3). Ambas muestras se hibridan simultáneamente en el arreglo. Una vez efectuada la lectura de la señal, los valores de intensidades generados son una medida de los niveles de expresión de los genes que están representados. Las 2 muestras utilizadas pueden corresponder a dos condiciones que se deseen comparar para determinar la expresión relativa entre ellas o se pueden aparear las muestras de interés con una única muestra control. De esta forma se obtiene la expresión de cada una respecto al control, lo que permitirá a su vez comparar la expresión de cualquier par de muestras. Existen también tecnologías ADNc que utilizan sondas impresas sobre membranas de nylon y marcaje radioactivo de la muestra.

Affymetrix. Por cada gen se selecciona un conjunto de sondas representativas y por cada una de estas sondas se imprime un par PM-MM como se describió anteriormente. En este caso, la impresión ocurre sobre un *chip* de silicón [40]. Cada muestra a analizar es marcada e hibridada individualmente en un arreglo. Después de un proceso de lavado, el nivel de expresión de cada gen es estimado por un algoritmo que analiza las intensidades de los conjuntos de pares PM-MM correspondientes a cada gen. En esta tecnología, la longitud de las sondas empleadas es más pequeña que en los arreglos de ADNc. La firma *Affymetrix*, en unos de sus productos más recientes, el *Human Gene 1.0 ST Array*, emplea sondas de 25 pb para cerca de 29 000 genes humanos, cada uno representado por aproximadamente 26 sondas diferentes distribuidas a lo largo del gen correspondiente, lo que hace un total de más de 750 000 sondas diferentes. Cada sonda se imprime en un área de 5 micras de lado (<http://www.affymetrix.com/>).

Para decidir el uso de una u otra plataforma debe tenerse en cuenta las ventajas y desventajas de ambas. La tecnología *Affymetrix* es más confiable pero más costosa que la de ADNc. Además, la posibilidad de doble marcaje en ADNc puede abaratar el experimento si se reduce el número de arreglos en dependencia del tipo de diseño que se seleccione. *Affymetrix* es más precisa y reproducible que ADNc. Las sondas en los arreglos de *Affymetrix* son más homogéneas y menos variables que en los arreglos ADNc y logra menor variabilidad entre arreglos debido a que minimiza los efectos por la distribución de las muestras dentro del arreglo durante la hibridación, mientras que en ADNc esta distribución no es uniforme. Estos elementos hacen que varios laboratorios se propongan los primeros experimentos con la tecnología *Affymetrix* sobre un *chip* con cobertura del genoma completo, hacen

33. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; 4:249-64.

34. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet* 1999;21 Suppl 1:20-4.

35. DeRisi JL, Iyer VR, Brown PO. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 1997;278:680-6.

36. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K, *et al.* Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci USA* 2004;101:811-6.

37. Lin F, Yu YP, Woods J, Cieply K, Gooding B, Finkelstein P, *et al.* Myopodin, a synaptodin homologue, is frequently deleted in invasive prostate cancers. *Am J Pathol* 2001;159:1603-12.

38. Yu YP, Tseng GC, Luo JH. Inactivation of myopodin expression associated with prostate cancer relapse. *Urology* 2006; 68:578-82.

39. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415:530-6.

40. Fodor SPA, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed, spatially addressable parallel chemical synthesis. *Science* 1991;251:767-73.

una selección de cientos de genes, que luego analizan con otras tecnologías más baratas y menos precisas, como la de ADNc, en mayor número de replicas, preferiblemente biológicas.

Selección del tipo de diseño

Una vez establecida la pregunta biológica, el objetivo del experimento y el tipo de plataforma es posible seleccionar el diseño experimental que mejor corresponda. Hay una diferencia importante entre, por ejemplo, utilizar microarreglos del tipo *Affymetrix* y del tipo ADNc de doble marcaje. En el primer caso, cada muestra es marcada e hibridada en arreglos independientes mientras que en arreglos de ADNc, generalmente se usa doble marcaje de manera que en un mismo arreglo se combinan 2 muestras diferentes que pueden ser una de las muestras en estudio y una muestra control o dos muestras provenientes de condiciones que se desee comparar.

En los experimentos de ADNc, a causa de la diferencia de eficiencia en el marcaje de los dos fluoróforos [41, 42] (al evaluar una misma muestra con dos marcadores, las intensidades de expresión difieren según el fluoróforo utilizado) se hace necesario hacer replicas con marcaje reverso. Dobbin *et al.* [41, 43] plantean que no es necesario hacerlas para cada par de muestras A y B, sino que puede hacerse el reverso usando replicas biológicas de A y B, es decir, un marcaje balanceado para ganar en eficiencia. Posteriormente al realizar una normalización a partir del marcaje reverso se eliminan los sesgos promedios por marcaje aunque pueden quedar sesgos específicos a genes en particular.

Existen varios tipos de diseños de experimentos que se repiten en los estudios con microarreglos de ADNc, el más usado y recomendado es el de referencia. A continuación se describen brevemente los más importantes.

Diseño de referencia. Se basa en el uso de una muestra de referencia, preferiblemente universal, que se hibrida en cada arreglo al igual que la muestra de interés (Figura 4a). Para facilitar la posterior comparación entre experimentos, se recomienda que los laboratorios usen la misma muestra de referencia en todos los experimentos [44, 45]. Este tipo de diseño facilita la inclusión de nuevas muestras pasado un periodo de tiempo. Su mayor desventaja es que la mitad de las hibridaciones se realizan con la muestra de referencia, lo que aumenta los costos.

Diseño en bloque balanceado. Esta es una alternativa de diseño propuesta por Dobbin y Simon [46] que consiste en aplicar una muestra diferente de cada grupo en cada arreglo, alternando el orden de la asignación de fluoróforos a las muestras según el grupo a que pertenecen (Figura 4b). Este diseño puede utilizarse en situaciones simples donde se quieren comparar 2 tipos de muestras. Es un diseño muy eficiente en cuanto al número de arreglos a emplear, por cada par de muestras a comparar se utiliza un solo arreglo, pero tiene como desventajas que se dificultan el uso de los métodos de agrupamiento y las comparaciones de perfiles de expresión entre diferentes arreglos y grupos experimentales.

Diseño en lazo. Propuesto por Kerr y Churchill [47], en este diseño los pares de muestras a comparar se

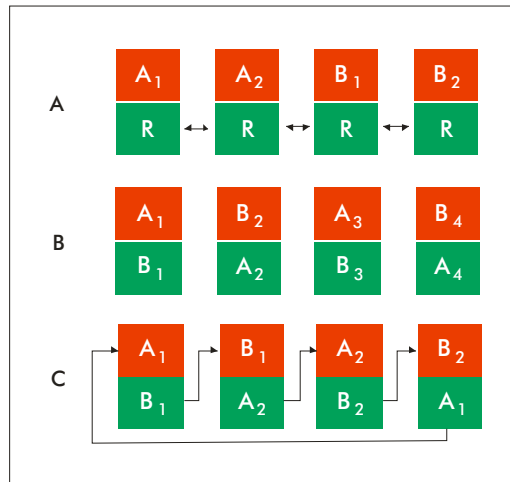


Figura 4. Tipos fundamentales de diseños de experimentos. a) Diseño de referencia; b) Diseño en bloque balanceado; c) Diseño en lazo. A_x, B_x representan dos conjuntos de muestras en estudio y R representa una muestra de referencia. En el caso a) siempre en cada chip se combina una muestra de interés con una de referencia, mientras que en b) y c) se proponen alternativas de diseños donde las mismas muestras en estudios pueden funcionar como referencia.

distribuyen de forma tal que cada muestra se hibrida en dos arreglos diferentes, en cada caso con un fluoróforo diferente, en la forma que se muestra en la figura 4c. Este diseño requiere del doble de arreglos que el de bloque balanceado, dificulta el uso de métodos de agrupamiento y requiere métodos de análisis más complejos que el del diseño de referencia. Por estas razones es poco usado.

Fuentes de variabilidad en el diseño

Algunas fuentes de variabilidad a tener en cuenta en el diseño de experimento de microarreglos ya se han referido. A continuación se enumeran las que se consideran más importantes:

- La heterogeneidad biológica de la población y las muestras en estudio.
- El proceso de colección y de manipulación de muestras.
- La extracción de ARN y la amplificación de ARN en el caso de que se realice.
- El marcaje de las muestras (eficiencia del marcaje, propiedades físicas del fluoróforo).
- La hibridación y lectura en dependencia del voltaje PMT y potencia del láser.

Pequeñas variaciones en las condiciones señaladas pueden inducir cambios significativos en los valores de expresión de los genes en estudio, que darían como resultado predicciones erróneas. Sin embargo estos sesgos en los datos de microarreglos pueden atenuarse con controles adecuados, un número de replicas en correspondencia con los niveles de variabilidad estimados y la normalización estadística [48].

La contaminación del tejido que será analizado también es una fuente de variabilidad a tomar en consideración. Con el objetivo de obtener una mayor precisión en la selección del tejido de las muestras a estudiar se han utilizado técnicas como el LMM (*Laser Microbeam Microdissection*) [13].

41. Dobbin K, Shih JH, Simon R. Statistical design of reverse dye microarrays. *Bioinformatics* 2003;19:803-10.

42. Rosenzweig BA, Pine PS, Doman OE, Morris SM, Chen JJ, Sistrare FD. Dye bias correction in dual-labeled cDNA microarray gene expression measurements. *Environ Health Perspect* 2004;112:480-7.

43. Dobbin K, Shih JH, Simon R. Questions and answers on design of dual-label microarrays for identifying differentially expressed genes. *J Natl Cancer Inst* 2003; 95:1362-9.

44. Novoradovskaya N, Whitfield ML, Basehore LS, Novoradovsky A, Pesich R, Usary J, et al. Universal Reference RNA as a standard for microarray experiments. *BMC Genomics* 2004;5:20.

45. Khan RL, Gonye GE, Gao G, Schwaber JS. A universal reference sample derived from clone vector for improved detection of differential gene expression. *BMC Genomics* 2006;7:109.

46. Dobbin K, Simon R. Comparison of microarrays designs for class comparison and class discovery. *Bioinformatics* 2002; 18:1438-45.

47. Kerr MK, Churchill GA. Statistical design and analysis of gene expression microarray data. *Genet Res* 2001;77:123-8.

48. Fan J, Ren Y. Statistical analysis of DNA microarray data in cancer research. *Clin Cancer Res* 2006;12:4469-73.

Tabla. Tipo de diseños de experimento recomendado según los objetivos del experimento

Objetivos / Tipo de Diseño	Referencia*	Bloque balanceado	Lazo
Comparación de clases	Se recomienda este tipo de diseño para comparar más de dos clases	Se recomienda este tipo de diseño para comparar dos clases	El diseño en Lazo es menos eficiente que el bloque balanceado y requiere métodos de análisis más complejos que el de referencia, generalmente no se recomienda su uso
Predicción de clases	Se recomienda un diseño de referencia para predicción	No se recomienda	No se recomienda
Descubrimiento de clases	Se recomienda un diseño de referencia para ejecutar métodos de clustering	No se recomienda	No se recomienda

*Puede observarse cómo el diseño de referencia funciona bien para todos los objetivos definidos, además de permitir futuras comparaciones entre experimentos.

Existe una relación de dependencia muy estrecha entre el tipo de diseño, el objetivo y el método estadístico para el análisis de los datos generados a partir del diseño. En la tabla se puede observar una propuesta para la elección del tipo de diseño de acuerdo con el objetivo que se propone el experimento.

Selección de las muestras

El principio de homogeneidad es el principio básico para la selección del material biológico en experimentos de microarreglos [49], que se satisface si se construye una población de controles como una muestra aleatoria extraída de la misma población que los casos. También se debe ser cuidadoso al seleccionar los casos que serán incluidos en el estudio de manera que sus características queden representadas. Estos estudios de expresión de genes además deben conducirse de manera que se cumpla el principio de comparabilidad de la exactitud de las mediciones entre casos y controles, por ejemplo, tumores vs tejidos sanos. Este principio trata de disminuir los sesgos por errores en las mediciones, es decir, que puedan ser diferentes en casos y controles.

Otro aspecto importante es la realización de un número adecuado de replicas, que pueden ser técnicas o biológicas. Las replicas técnicas van dirigidas a medir la variabilidad experimental, realizadas ya sea ubicando múltiples copias de un mismo gen en el arreglo o evaluando una muestra en varios arreglos, pero no están relacionadas con el fenómeno biológico en investigación. Por otra parte, las replicas biológicas se proponen medir la variabilidad biológica entre los individuos de una población. A pesar de que las replicas técnicas son importantes para aumentar la precisión, no dan una medida de la variabilidad biológica, por lo que varios autores plantean que es más provechoso y eficiente hacer replicas biológicas que técnicas [50, 51]. Existen casos particulares en los que las replicas técnicas serían de mucha utilidad, por ejemplo, al evaluar un predictor de genes para el diagnóstico médico de un paciente.

En el presente trabajo no se aborda el tema de la determinación del tamaño de muestra, que debe definirse dentro del diseño del experimento por estar ampliamente divulgado en la literatura existente[52-56].

Realización del experimento de microarreglos

Pasos fundamentales

Con independencia de la tecnología utilizada, hay pasos comunes que son imprescindibles en la realización

de un experimento, estos son: extracción de ARNm de las muestras, hibridación de cada muestra en uno o varios arreglos, lectura del arreglo y obtención de la imagen, localización en la imagen del área correspondiente a cada uno de los pocillos y asignación de las intensidades a cada gen como medida de su expresión en la muestra hibridada.

En la figura 5 pueden observarse las principales etapas para la obtención de los datos primarios de un experimento de microarreglos. Si se considera que se cuenta con chips impresos previamente, el experimento consiste en la realización de procesos continuos de marcaje, hibridación y lectura de las muestras en similares condiciones experimentales y a continuación cada chip se somete al análisis de imagen, que permite conocer las intensidades de cada elemento del chip en las muestras hibridadas.

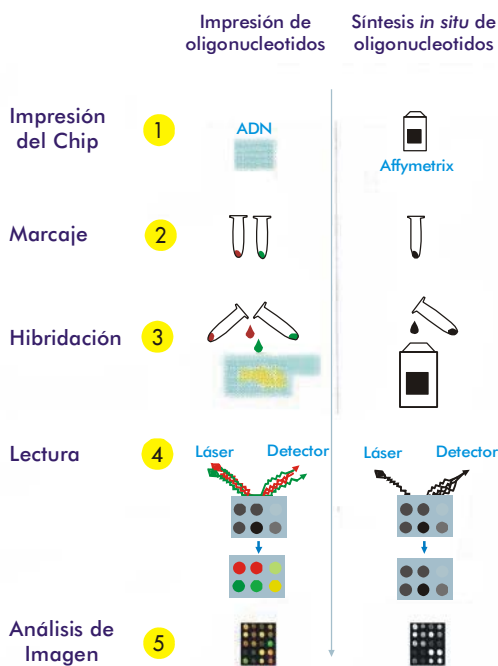


Figura 5. Principales etapas para la obtención de los datos primarios de un experimento de microarreglos. Se ilustran las dos tecnologías más utilizadas, ADNc a la izquierda y Affymetrix a la derecha. Ambas tecnologías requieren de: la impresión del chip, el marcaje de las muestras, hibridación de las muestras en cada chip, lectura de las placas y empleo de un escáner para los análisis de las imágenes. La diferencia fundamental entre ellas está en el diseño e impresión de los chips.

49. Repsilber D, Fink L, Jacobsen M, Blasing O, Ziegler A. Sample selection for microarray gene expression studies. *Methods Inf Med* 2005;44(3):461-7.

50. Kerr, MK. Design considerations for efficient and effective microarray studies. *Biometrics* 2003;59:822-8.

51. Landgrebe J, Bretz F, Brunner E. Efficient two-sample designs for microarray experiments with biological replications. *In Silico Biol* 2004;4:461-70.

52. Simon R, Radmacher MD, Dobbin K. Design of studies using DNA microarrays. *Genet Epidemiol* 2002;23:21-36.

53. Pavlidis P, Li Q, Noble WS. The effect of replication on gene expression microarray experiments. *Bioinformatics* 2003;19:1620-27.

Otras recomendaciones

- Debe ser realizado cada experimento por un único investigador.
- Debe aleatorizarse la asignación de arreglos en cada hibridación.
- Es necesario antes de aplicar un tratamiento conocer las diferencias de base entre los tejidos a comparar.
- Considerar que si la variabilidad de la población es alta, se obtiene poca ventaja al usar un tipo de diseño que no sea el de referencia.

Conclusiones

En el resumen de las diferentes etapas involucradas en el diseño y la realización de un experimento de microarreglos, en el que se describen brevemente las principales tecnologías y los objetivos fundamentales que han tenido los experimentos de microarreglos en

las investigaciones biomédicas, se ha destacado la importancia de la etapa de diseño experimental y dentro de ella, la necesidad de la definición de la pregunta biológica para establecer los objetivos del experimento y el diseño más adecuado. Se han expuesto además algunas ideas y recomendaciones sobre esta etapa que pueden ayudar al investigador a obtener resultados más precisos y confiables. También se ha propuesto el uso de datos públicos de microarreglos para integrar este conocimiento a la búsqueda de los mecanismos moleculares de enfermedades complejas como el cáncer. Este punto se ejemplificó mediante un análisis realizado con datos públicos de cáncer de próstata. Es previsible el futuro desarrollo de la tecnología, que implicará una mayor precisión y generación de información, lo que sin dudas modificará aspectos del diseño experimental y ampliará sus posibles aplicaciones.

54. Zien A, Fluck J, Zimmer R, Lengauer T. Microarrays: how many do you need?. *J Comput Biol* 2003; 10:653-67.

55. Dobbin K, Simon R. Sample size determination in microarray experiments for class comparison and prognostic classification. *Biostatistics* 2005;6:27-38.

56. Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A. False discovery rate and sample size for microarray studies. *Bioinformatics* 2005;21:3017-24.

Recibido en octubre de 2007. Aprobado en mayo de 2008.