
La revolución de los microarrays en la investigación biosanitaria: tipos de plataformas, usos y perspectivas en oncología

The microarray revolution in biomedical research: types of platforms, uses and perspectives in oncology

J.C. Cigudosa

RESUMEN

Mediante la introducción, uso y manejo de herramientas de base genómica, la investigación sobre las alteraciones genéticas que están en el origen de enfermedades tan comunes como el cáncer han sufrido una revolución técnica comparable a la incorporación del microscopio en los laboratorios. Ahora, estudiar la relación gen-enfermedad no está basado en analizar un gen único y sus efectos sino en analizar el comportamiento de miles de genes de forma simultánea. Estos sistemas, denominados genéricamente como matrices, arrays, microarrays o biochips, están cambiando nuestra forma de plantear problemas y extraer conclusiones de los experimentos ya que nos ofrecen una foto compleja del conjunto del genoma. Los análisis de expresión mediante microarrays de cDNA o de oligos ya son accesibles a la comunidad científica española. Los resultados, además, fascinan a los investigadores ya que son bastante reproducibles y aportan una gran cantidad de información sobre la regulación de la expresión génica en condiciones normales y patológicas.

Palabras clave. Microarray. Perfil de expresión. Genómica. Cáncer.

ABSTRACT

Through the introduction, use and management of genome-based tools, research into genetic alterations that give rise to diseases as common as cancer has undergone a technical revolution comparable to the introduction of the microscope in laboratories. Now, study of the gene-disease relationship is no longer based on analysing a single gene and its effects, but on analysing the behaviour of thousands of genes in a simultaneous form. These systems, generically called matrices, arrays, microarrays or biochips, are changing the way we pose problems and draw conclusions from experiments, since they offer us a complex photo of the genome as a whole. Analyses of expression through microarrays of cDNA or oligos are now accessible to the Spanish scientific community. The results have proved fascinating to researchers since they can be reproduced easily and contribute a great quantity of information on the regulation of the gene expression in normal and pathological conditions.

Key words: Microarray. Profile of expression. Genome. Cancer.

An. Sist. Sanit. Navar. 2004; 27 (1): 11-20

Programas de Biotecnología y Patología Molecular. Centro Nacional de Investigaciones Oncológicas

Correspondencia:

Dr. Juan C. Cigudosa
Unidad de Citogenética
Centro Nacional de Investigación Oncológicas
C/ Melchor Fernández Almagro, 3
28029 Madrid
Tfno. 91 2246900
Fax: 91 2246923

INTRODUCCIÓN

El cáncer es una enfermedad genética. Esta frase sencilla y simple tiene un alcance que ha sobrepasado cualquier tipo de aproximación que la ciencia biosanitaria ha realizado hasta la fecha para estudiar el fenómeno oncológico. La principal consecuencia de la definición del cáncer como enfermedad genética es que la piedra angular para la investigación oncológica ha de estar basada en herramientas de análisis genético.

Partiendo de ese enfoque genético y basado en la biología celular, ¿cómo se investiga la relación gen-enfermedad? El esquema de investigación habitualmente seguido para caracterizar esa relación consiste, en esencia, en estudiar un gen y su efecto en el modelo biológico estudiado, en determinar las relaciones con otros genes y, por último, en diseñar modificaciones (inhibidores o activadores) de su expresión para regular su efecto en condiciones patológicas. Este sistema ha sido muy útil pero se adivina un tanto insuficiente si tenemos en cuenta que el número de genes es de unos 30.000, que se traducen en unas 100.000 proteínas y que las interacciones gen-gen, proteína-proteína y gen-proteína, por simple combinatoria, alcanzan números casi inimaginables.

En los últimos años, no más de diez, las herramientas de análisis genético han sufrido una revolución técnica comparable a la incorporación del microscopio en los laboratorios. Han aparecido sistemas de análisis genómico masivo que permiten analizar en un solo experimento el estatus de miles de genes. Es decir, ahora, estudiar la relación gen-enfermedad no está basado en analizar un gen único y sus efectos sino en analizar el comportamiento de miles de genes de forma simultánea. Estos sistemas, denominados genéricamente como matrices, arrays, microarrays o biochips, están cambiando nuestra forma de plantear problemas y extraer conclusiones de los experimentos ya que nos ofrecen una foto compleja del conjunto del genoma. Este conjunto de datos sobre expresión génica y relaciones entre genes tratados de forma sistemática mediante herramientas de análisis masivo, constituye el cuerpo esencial de lo que denominamos genómica. La genómica intenta proporcionar a los investigadores en biología el equivalente de la Tabla Periódica de los Elementos, es decir, un inventario de todos los genes que se emplean para conseguir una criatura viviente, asociado a un sistema de clasificación que genere bloques o subconjuntos de genes que se comportan de forma coordinada.

Estudios genómicos masivos, como los realizados con microarrays de expresión, nos ofrecen información rápida y reproducible sobre el nivel de expresión de un número elevado de genes (generalmente varios miles) en la muestra estudiada en un solo experimento. Esa información tiene varios

niveles de complejidad. El nivel más bajo corresponde a datos simples sobre genes concretos. Un array de expresión puede identificar el nivel de expresión de un gen individual asociado a un determinado fenotipo de forma similar a experimentos más tradicionales. Los niveles más alto de complejidad, sin embargo, permiten examinar la figura completa del conjunto del transcriptoma, es decir todos los genes que se expresan de forma asociada a ese fenotipo. Este nivel de complejidad es imposible de alcanzar con otros sistemas distintos a los biochips. Además, los arrays de expresión han acelerado en varios órdenes de magnitud el descubrimiento de marcadores moleculares de utilidad en medicina clínica. Las investigaciones recientes procedentes de varios laboratorios están revelando principios fundamentales que ayudan a clasificar una enfermedad de acuerdo con su perfil de expresión específico. Esos principios se basan en datos que proceden del análisis de microarrays y que son sorprendentemente reproducibles, a menudo esclarecedores y algunas veces inesperados.

Sin duda, la mayor contribución de los estudios con arrays tiene su nicho natural en la taxonomía. Clasificar un determinado tumor con un nombre y apellidos lo más certero posible permite ajustar el diagnóstico y el tratamiento casi de forma individual. Sin embargo, como todos sabemos, realizar una filiación adecuada de un caso real depende de muchas variables: circunstancias de recogida de la muestra, ser riguroso en la descripción, disponibilidad de herramientas de diagnóstico adecuadas y, finalmente, la experiencia del profesional. Todos estos elementos tienen como final el desarrollo de una clasificación de los tumores que sea universal y efectiva. Este desarrollo es paralelo a las estrategias prácticas que se han desarrollado en el campo de la patología. En primer lugar, se da el reconocimiento de una diferencia en la morfología de una manera reproducible y sistemática. En segundo lugar, se intenta establecer la relación entre las características del tumor y las del tejido normal del que se origina. En tercer lugar, se busca y se trata de establecer asociaciones entre la evolución del paciente o la respuesta al tratamiento y la clasificación propuesta, de forma que ésta tenga una utilidad clínico-terapéutica. La tecnología de microarrays puede usarse para derivar clasificaciones, establecer asociaciones con la línea celular de origen y ayudar en el pronóstico, pero con la ventaja añadida de ser capaz de identificar los genes que determinan esa clasificación¹.

DEFINICIÓN Y TIPOS DE MICROARRAYS

Los ensayos de hibridación en microarrays, descritos a finales de los años ochenta, se basan en la disposición de material genético sobre un sustrato (plástico, cristal, membranas), en posiciones conocidas. En la actualidad existen en

el mercado científico herramientas de análisis genómico (matrices o microarrays) diseñadas a partir de ácidos nucleicos en forma de cDNA (DNA monocatenario y complementario al RNA mensajero que contiene la secuencia codificante de un gen), secuencias cortas de nucleótidos (oligonucleótidos), o DNA de doble cadena. En todos los casos, cada material genético que se coloca en el array está referido a un gen de situación y características conocidas o predichas con cierta verosimilitud. Es decir, cada uno de los puntos del microarray representa un gen y la información que nos genera se refiere al estatus de ese gen en la muestra analizada. Los microarrays más extendidos y utilizados son, sin embargo, los dos primeros tipos (cDNA y oligos) y los experimentos en los que son empleados son, casi exclusivamente, estudios sobre análisis masivo o global de la expresión genética de organismos.

El material obtenido de la muestra problema, convenientemente tratado y marcado (radioactividad, fluorescencia) se pone en contacto con el array, hibridando en aquellas posiciones en que se complementa con las sondas. El patrón de hibridación es revelado empleándose un escáner basado en microscopía confocal, y la imagen resultante se convierte a valores numéricos que constituyen los resultados del ensayo. La cuantificación de la expresión génica diferencial es susceptible de un abordaje con la tecnología descrita. El nivel de expresión de un gen se refleja en el número de copias de mRNA presentes en la muestra problema y, por tanto, es proporcional al nivel de señal detectado.

Una de las ventajas significativas es que, debido a la miniaturización del sistema, la alta concentración del material permite la identificación de muestras presentes en un bajo número de copias. Además las técnicas de revelado por fluorescencia ofrecen una elevada relación señal/ruido incluso en casos de transcritos poco abundantes.

El alto grado de integración del array permite, en un solo ensayo, obtener multitud de valores de expresión génica relativa para distintas condiciones biológicas, lo que convierte a esta técnica en una herramienta de alto rendimiento para trabajos en el área de la genómica funcional. El gran volumen de datos generados debe ser tratado con herramientas y métodos bioinformáticos.

Aunque la aplicación más conocida de los biochips de DNA es la determinación de perfiles de transcripción, el formato microarray también ha sido utilizado de manera eficaz en otros ámbitos, como en la genómica estructural. Al este nivel genómico-cromosómico el formato del microarray se ha adaptado a la detección de polimorfismos puntuales (*Single Nucleotide Polymorphisms* o SNPs) y a la detección de variaciones en la dosificación o número de copias del genoma mediante

arrays de Hibridación Genómica Comparada (CGH array). Hasta la fecha, la técnica más empleada para la detección de los cambios en el número de copias de ADN a lo largo de todo el genoma, ha sido la Hibridación Genómica Comparada (CGH), desarrollada en 1992². La técnica de CGH se ha aplicado extensamente durante los últimos años a los más diversos tipos de neoplasias, mostrando su gran utilidad en la búsqueda de regiones cromosómicas.

La CGH, sin embargo, tiene una capacidad de resolución limitada; los cambios que impliquen segmentos inferiores a 7-10 megabases no son detectados. Segmentos de ese tamaño pueden contener decenas de genes cuyo comportamiento se escapa al análisis. Esta limitación puede salvarse mediante la utilización de los CGH arrays, una técnica que combina la tecnología de los microarrays con la CGH, y en la que la hibridación se realiza sobre segmentos de ADN de menor tamaño clonados en forma de Cromosomas Artificiales Bacterianos o Plasmídicos (BAC o PAC, respectivamente). El nivel de resolución pasa, por tanto, de cromosómico a bacteriano (150 Kb). Estos segmentos de ADN son secuencias genómicas conocidas cuya localización en el genoma es conocida y cuyo contenido génico es más o menos conocido³. La técnica se alimenta de la gran cantidad de datos aportados por el Proyecto Genoma Humano, sobre la ubicación, secuencia y contenido de genes de estos fragmentos de ADN, en buena medida, conocidos.

LOS MICROARRAYS DE EXPRESIÓN Y SU EMPLEO EN LA TAXONOMÍA MOLECULAR DEL CÁNCER

En general, como hemos comentado anteriormente, el análisis de los patrones de expresión tiene una incidencia directa en el diagnóstico molecular de los tumores. Cuanto mayor sea nuestra capacidad para clasificar un tumor determinado de forma no ambigua, acumularemos información más precisa referente a la biología de ese cáncer (capacidad de metástasis, resistencia al tratamiento, etc.) y, en definitiva, mayores serán nuestras posibilidades de éxito en la curación de ese paciente. El efecto clínico del empleo de los microarrays de expresión, aunque esté pendiente de validaciones a gran escala en grupos de pacientes homogéneos, está basado en dos evidencias que ya tenemos documentadas en la literatura. Por un lado, los tumores presentan patrones de expresión directamente relacionados con la línea celular de la que proceden y con características clonales específicas de pacientes. Por otro, los patrones de expresión nos ayudan a establecer diferencias en la biología y el pronóstico de los tumores más allá de las aproximaciones empleadas hasta la fecha.

Empleando el microscopio de luz transmitida, los tumores presentan un aspecto que recuerda a la célula normal en la

que se origina pero en la que se hubiera detenido el proceso de diferenciación celular. El resultado de esa similitud es que la taxonomía del cáncer está basada en los distintos tipo histológicos normales en los que se origina. Al aplicar los estudios mediante arrays de expresión a diferentes tipos de tumores, lo primero que observamos es que los perfiles de expresión se correlacionan bastante bien con los fenotipos histológicos de los tumores. En leucemias y linfomas, esta relación está comprobada y nos acerca mucho a una clasificación basada en el linaje celular^{4,6}. Otros autores han encontrado también diferencias en tumores de origen epitelial. Por ejemplo, tumores de mama primarios se agrupan en dos patrones de expresión semejantes a los encontrados en líneas celulares derivadas de dos linajes diferentes del tejido mamario (células basales y lumbales)⁷.

Es muy interesante comprobar que, al margen del tipo celular estudiado o del tipo de análisis, la proporción de genes que, en cualquier array, aparece diferencialmente expresada entre el tumor y el tejido normal de origen es de, aproximadamente, sólo un 5%. Las relaciones que pueden ser definidas mediante patrones de expresión pueden también contribuir a aumentar nuestro conocimiento de la progresión tumoral. Lesiones precancerosas, como las encontradas en adenomas de colon, pueden separarse mediante patrones de expresión del tejido normal y el carcinoma. Sin embargo, esos patrones de expresión génica del adenoma presentan un mayor grado de similitud con los carcinomas que con el tejido normal. En general, estos estudios nos demuestran que el linaje celular es un elemento determinante del patrón de expresión de un tumor.

Otros datos que apoyan este papel importante del linaje celular provienen de observaciones entre tejidos pareados. Estudios comparativos entre tumores y metástasis o entre momentos terapéuticos distintos que proceden del mismo paciente presentan unos patrones de expresión más similares entre sí que cuando se comparan el mismo tipo de muestra (tumores, metástasis, después del tratamiento) entre diferentes pacientes. Esto se ha comprobado en cáncer de mama⁸, melanomas⁹ y hepatocarcinomas¹⁰. Este compromiso clonal del patrón de expresión de un tumor, íntimamente relacionado con el paciente y su origen de linaje celular, permite obtener marcadores moleculares individuales y, con el tiempo, será muy útil en el desarrollo de terapias individualizadas.

LOS PATRONES DE EXPRESIÓN COMO HERRAMIENTA DE PRONÓSTICO Y PREDICCIÓN DE LA RESPUESTA AL TRATAMIENTO

Una vez establecido que los patrones de expresión reafirman en gran medida los patrones histológicos diferenciales

en los tumores, es necesario demostrar que la ventaja añadida de reconocer esos patrones es poder subclasificar los tumores en base a su comportamiento biológico. Esos subgrupos deberían estar definidos mediante marcadores moleculares sutiles y complejos que conlleven un impacto significativo en la predicción del pronóstico y la respuesta al tratamiento. Esta aproximación se ha conseguido en varios tipos de cáncer.

Leucemias agudas y linfomas son tipos de cáncer especialmente informativos por su accesibilidad, la abundancia de marcadores, el conocimiento adquirido de las células normales originales y la existencia de tratamientos muy efectivos. Por ejemplo, se han publicado ya análisis de expresión, con arrays de Afymetrix, de más de 300 pacientes infantiles con leucemia linfoblástica aguda (LLA)^{11,12}. Los resultados permiten clasificar esas leucemias en subgrupos que relacionan patrones de expresión diferentes con alteraciones citogenéticas y/o histológicas esenciales: LLA de células T, hiperdiploidía con más de 50 cromosomas, BCR-ABL, E2A-PBX, TEL-AML1, reordenamientos del gen MLL, y un grupo diferenciado que carece de marcador citogenético específico. Dado que cada uno de estos subgrupos conlleva un pronóstico asociado, el estudio mediante microarrays va a proporcionar un gran número de marcadores informativos sobre la biología de ese pronóstico y facilitará de forma enorme la monitorización del proceso terapéutico. En la siguiente década vamos a asistir a la validación de todos estos marcadores.

Otro ejemplo significativo son los estudios realizados en cáncer de pulmón y próstata y el establecimiento de grupos de pronóstico^{13,14}. Al menos tres estudios han ayudado a separar los adenocarcinomas de pulmón en tres subgrupos que expresan de forma diferencial un conjunto de genes. Esos subgrupos tienen diferentes patrones de supervivencia: los de peor pronóstico son aquellos adenocarcinomas que expresan de forma preferente genes de diferenciación neumocítica y tienen una alta expresión de genes implicados en procesos de remodelación tisular y actividad proliferativa. Hay publicados estudios de la misma naturaleza en próstata¹⁵. Un cuidadoso análisis que integra datos de cuatro estudios diferentes con microarrays ha identificado que las rutas de biosíntesis de las purinas y las poliaminas están completamente desreguladas a nivel transcripcional.

Finalmente, otro de los usos más prometedores de los microarrays es el empleo de los patrones de expresión como predictores de la respuesta al tratamiento. Se conocen datos confirmatorios en relación con el tratamiento de linfomas, leucemias agudas y cáncer de mama. El grupo de Alizadeh en colaboración con el *Lymphoma/Leukemia Molecular Profiling Project* ha estudiado, mediante microarrays de expresión, 240

pacientes con linfoma B difuso de células grandes¹⁶. Han confirmado la existencia de 3 subgrupos y han descrito cinco patrones de expresión diferentes (firmas moleculares) con pronósticos asociados diferenciales. Las firmas del centro germinal, MHC clase II y nódulo linfático están asociadas a buen pronóstico, mientras que las firmas de genes proliferativos y “otros genes” están asociadas a mal pronóstico. Como todos los pacientes recibieron el mismo tratamiento, los resultados de los perfiles de expresión también han podido ser utilizados para elaborar predictores de la respuesta al tratamiento estándar de este tipo de linfoma.

Los estudios realizados con muestras de cáncer de mama han sido realizados con diferentes plataformas y sobre diferentes poblaciones^{7,17}. Sin embargo, los resultados son muy concordantes. La primera conclusión es que el elemento de más impacto en el perfil de expresión del cáncer de mama es el estatus de receptor de estrógenos (RE), reforzando el papel fundamental del eje RE en la génesis del cáncer de mama humano y en la determinación de la respuesta a la manipulación hormonal. Además subdividen los grupos pronósticos con mayor precisión. Sotirou y col⁸ han realizado estudios de perfiles asociados a la respuesta al tratamiento y han podido determinar genes directamente relacionados con la respuesta farmacodinámica.

CONCLUSIONES Y PERSPECTIVAS

Los análisis de expresión mediante microarrays de cDNA o de oligos ya son accesibles a la comunidad científica española. Los resultados, además, fascinan a los investigadores ya que son bastante reproducibles y aportan una gran cantidad de información sobre la regulación de la expresión génica en condiciones normales y patológicas. Como consecuencia del éxito de estas herramientas, se pueden recoger en la literatura un gran número de experimentos realizados con arrays de expresión y que agrupan, relacionan o describen los perfiles de expresión de cientos o miles de genes según dos tipos generales de preguntas:

¿Puedo utilizar los perfiles de expresión para clasificar de una manera mejor los diferentes tipos de cáncer?

¿Puedo conocer el efecto, en el perfil de expresión, de estímulos extrínsecos (drogas, estrés, etc.) o internos (mutaciones, sobreexpresión de genes)?

La respuesta a ambas preguntas es afirmativa aunque, en general, este tipo de análisis han de ser validados con posterioridad mediante técnicas de PCR cuantitativa ya que existen problemas de estandarización y otras fuentes de variabilidad como:

- a) La naturaleza normalmente lábil y diluida de las muestras biológicas.
- b) Condiciones estocásticas y subóptimas de los procesos de marcaje y de interacción molecular con elementos microscópicos insuficientemente redundantes.
- c) Sensibilidad indeterminada a factores y procedimientos mal estandarizados y controlados (temperatura y volúmenes de disoluciones, concentración de ozono atmosférico).
- d) Dependencia en instrumentación calibrada de forma variable (baños termostatzados, escáneres).
- e) Diversidad de procedimientos de procesamiento de datos.
- f) Sistemas inapropiados de análisis de datos.
- g) Escasez de recursos bioinformáticos para el tratamiento y comprensión de datos masivos. A todo esto se añade la diversidad existente de alternativas y de plataformas tecnológicas, y la variabilidad del proceso de producción de plataformas “hechas en casa” por laboratorios de apoyo a la investigación de centros académicos.

En resumen, esta tecnología es revolucionaria y va a tener un gran impacto en la medicina. Sin embargo, ninguna tecnología es suficiente para contribuir por sí sola al avance de las grandes cuestiones abiertas en la investigación biosanitaria. Lo importante son las preguntas y establecer validaciones adecuadas que permitan la traducción de todos estos hallazgos a la práctica clínica de rutina. Lo veremos pronto.

BIBLIOGRAFÍA

1. MILLER LD, LONG PM, WONG L, MUKHERJEE S, McSHANE LM, LIU ET. Optimal gene expression analysis by microarrays. *Cancer Cell* 2002; 2: 353-361.
2. KALLIONIEMI A, KALLIONIEMI OP, SUDAR D, RUTOVITZ D, GRAY JW, WALDMAN F et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 1992; 258: 818-821.
3. PINKEL D, SEGRAVES R, SUDAR D, CLARK S, POOLE I, KOWBEL D et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 1998; 20: 207-211.
4. GOULB TR, SLONIM DK, TAMAYO P, HUARD C, GAASENBECK M, MESIROV JP et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; 286: 531-537.
5. ALIZADEH AA, EISEN MB, DAVIS RE, MA C, LOSSOS IS, ROSENWALD A et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; 403: 503-511.
6. SHIP MA, ROSS KN, TAMAYO P, WENG AP, KUTOK JL, AGUIAR RC et al. Diffuse large B-cell lymphoma outcome prediction by gene expression profiling and supervised machine learning. *Nat Med* 2002; 8: 68-74.

7. SORLIE T, PEROU CM, TIBSHIRANI R, AAS T, GEISLER S, JOHNSEN H et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001; 98: 10869-10874.
8. SOTIRIOU C, POWLES TJ, DOWSETT M, JAZAERI AA, FELDMAN AL, ASSERSOHN L et al. Gene expression profiles derived from needle aspiration correlate with response to systemic chemotherapy in breast cancer. *Breast Cancer Res* 2002; 4: R3.
9. WANG E, MILLER LD, OHNMACHT GA, MOCELLIN S, PEREZ-DIEZ A, PETERSEN D et al. Prospective molecular profiling of melanoma metastases suggests classifiers of immune responsiveness. *Cancer Res* 2002; 62: 3581-3586.
10. CHEN X, CHEUNG ST, GUAN XY, WOPNG SY, TAI LS, NG IOL et al. Gene expression profiling in human liver cancers. *Mol Biol Cell* 2002; 13: 1929-1939.
11. YEOH EJ, ROSS ME, SHURTLEFF SA, WILLIAMS WK, PATEL D, MAHFOUZ R et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002; 1: 133-143.
12. FERRANDO AA, NEUBERG DS, STAUNTON J, LOH ML, HUARD C, RAIMONDI SC et al. Gene expression signatures define novel oncogenic pathways in T-cell acute lymphoblastic leukemia. *Cancer Cell* 2002; 1: 75-87.
13. GARBER MED, TROYANSKAYA OG, SCHULUENS K, PETERSEN S, THAESLER Z, PACYNA-GENGELBAHC M et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci USA* 2001; 98: 13784-13789.
14. VIRTANEN C, HONJOH D, KIMURA M, SHIMANE M, MIYOSHI T, NOMURA H et al. Integrated classification of lung tumors and cell lines by expression profiling. *Proc Natl Acad Sci USA* 2002; 99: 12357-12362.
15. SINGH D, FEBBO PG, ROSS K, JACKSON DG, MANOLA J, LADD C et al. Gene expression correlates of clinical prostate cancer behaviour. *Cancer Cell* 2002; 1: 203-209.
16. ROSENWALD A, WRIGHT G, CHAN WC, CONNORS JM, CAMPO E, FISHER RI et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large B-cell lymphoma. *N Engl J Med* 2002; 346: 1937-1947.
17. VAN'T VEER LJ, DAI H, VAND VIJVER MJ, HE YD, HART AA, MAO M et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415: 530-536.