



UNIVERSIDAD DE JAÉN  
*Facultad de Ciencias Experimentales*

Trabajo Fin de Grado

# **“Microarrays” de ARN**

**Alumno: Torres Godino, Isabel**

**Junio, 2014**





UNIVERSIDAD DE JAÉN  
*Facultad de Ciencias Experimentales*

Trabajo Fin de Grado

# **“Microarrays” de ARN**

**Alumno: Torres Godino, Isabel**

**Junio, 2014**



## ÍNDICE

|                                                                                                                          |    |
|--------------------------------------------------------------------------------------------------------------------------|----|
| 1. INTRODUCCIÓN .....                                                                                                    | 1  |
| 2. OBJETIVOS.....                                                                                                        | 4  |
| 3. MATERIAL Y MÉTODOS.....                                                                                               | 5  |
| 3.1. Diseño experimental .....                                                                                           | 5  |
| 3.2. Preparación de la muestra.....                                                                                      | 6  |
| 3.2.1. Extracción y aislamiento de ARN .....                                                                             | 6  |
| 3.2.2. Purificación del ARN .....                                                                                        | 7  |
| 3.2.3. Comprobación integridad/ calidad .....                                                                            | 7  |
| 3.3. Procedimiento de “microarray” mediante Affymetrix.....                                                              | 9  |
| 3.3.1. Amplificación de ARN mensajero .....                                                                              | 9  |
| 3.3.2. Transcripción inversa y marcaje .....                                                                             | 9  |
| 3.3.3. Descripción del “microarray” (Affymetrix GeneChips’s) .....                                                       | 11 |
| 3.3.4. Hibridación en el “microarray” .....                                                                              | 14 |
| 3.4. Análisis de imagen del “microarray” .....                                                                           | 16 |
| 3.5. Normalización y análisis de datos .....                                                                             | 18 |
| 3.5.1 Análisis de la diferencia de expresión a nivel de ARN .....                                                        | 21 |
| 3.5.2 Análisis de datos .....                                                                                            | 22 |
| 3.5.3 Verificación de resultados .....                                                                                   | 23 |
| 3.5.4 Clasificación de muestras o genes .....                                                                            | 23 |
| 3.6 Relaciones funcionales. Bioinformática aplicada a los “microarrays”.....                                             | 26 |
| 4. RESULTADOS Y DISCUSIÓN.....                                                                                           | 27 |
| 4.1 Aplicación de los “microarrays” en el estudio de enfermedades. Cáncer.....                                           | 27 |
| 4.2 Aplicación de los “microarrays” para el estudio de enfermedades<br>infecciosas y generación de vacunas .....         | 30 |
| 4.3 Aplicación de los “microarrays” en estudios llevados a cabo en <i>Arabidopsis</i><br><i>thaliana</i> y plantas ..... | 36 |
| 4.4 Aplicación de los “microarrays” en estudios de <i>Drosophila melanogaster</i> .....                                  | 37 |
| 4.5 Aplicación de los “microarrays” en estudios realizados en <i>Caenorhabditis</i><br><i>elegans</i> .....              | 40 |
| 5. CONCLUSIONES .....                                                                                                    | 41 |
| 6. BIBLIOGRAFÍA .....                                                                                                    | 44 |



## RESUMEN

El descubrimiento en la última década de la tecnología del *microarray* ha supuesto una revolución en los ensayos llevados a cabo en ARN y ADN. En contraposición con los estudios biológicos tradicionales, los “microarrays” permiten medir los niveles de expresión de miles de transcritos de ARN mensajero (ARNm) para conocer el patrón global de expresión génica en una determinada célula, tejido u órgano. Desde su aparición a final del siglo XX, los análisis de “microarrays” de ARN han constituido una herramienta esencial tanto en estudios biológicos como biomédicos por su gran aplicación en diversos campos, siendo imprescindibles en estudios de enfermedades, en ensayos con animales modelo como *Drosophila melanogaster* y *Caenorhabditis elegans* e incluso en plantas. Este trabajo proporciona una visión general de la metodología y aplicaciones de los “microarrays” de expresión de ARN en varias áreas de estudio biológico y demuestra que se está convirtiendo en una metodología de vanguardia en la biotecnología y biomedicina.

**Palabras clave:** ARN, transcritos, expresión, “microarrays”, metodología, aplicaciones.



## **ABSTRACT**

The discovery in the last decade of microarray technology has supposed a revolution in RNA and DNA assays. In contrast with traditional biological researches, “microarrays” allow the simultaneous measure of the expression levels of thousands of mRNA transcripts with the aim to know the global mRNA expression level in a cell, tissue or organ. Since its appearance to end of 20th century, RNA microarray analysis has become an essential tool in biological and biomedical studies for its wide application in different fields, being essential in diseases studies, animals model assays as *Drosophyla melanogaster* and *Caenorhabditis elegans* and even in plants. This review provides an overview of the methodology and applications of RNA expression “microarrays” in various areas of biological research and demonstrates that microarray technology is at the vanguard methodology in biotechnology and biomedicine.

**Key words:** RNA, transcripts, expression, “microarrays”, methodology, applications.



## 1. INTRODUCCIÓN

Aunque los principios de la tecnología de “microarrays” fueron concebidos hace casi 20 años y su tecnología desarrollada a partir de la técnica de “Northern-blot”, no es hasta esta era post-proteómica cuando han cobrado verdadera importancia y utilidad [1]. Una búsqueda rápida en *Pubmed* con las palabras “microarray and 1995” da como resultado un total de 58 artículos, 5 de los cuales son artículos de revisión. Búsquedas similares pero más avanzadas en los años, por ejemplo “microarray and 2005” dan como resultado un total de 4755 artículos, de los cuales una cuarta parte corresponden a revisiones y una inmensa mayoría a aplicaciones de estos.

A pesar de este gran aumento de artículos relacionados con “microarrays”, existen pocas publicaciones recientes que aborden cuestiones básicas sobre el diseño, aplicaciones y análisis de datos posteriores de estos. Cubriendo esta falta de información y casi desconocimiento de la materia, esta revisión pretende ilustrar los puntos fuertes, las debilidades y las aplicaciones de los “microarrays”, especialmente para aquellos no familiarizados con la tecnología. Hoy en día, las micromatrices son altamente superiores a sus predecesores en cuanto a la calidad, densidad de sonda y diseño estructural [2,3]. Antes de tratar con estas y otras características, es importante discutir, con cierta extensión, que son los “microarrays”, así como los conceptos fundamentales de su tecnología. El término *microarray* es descriptivo y un tanto ambiguo, utilizándose comúnmente para describir una variedad de plataformas incluyendo “microarrays” de proteínas y “microarrays” de tejidos [3,4]. Un microarray se define típicamente como una colección de puntos microscópicos dispuestos en una matriz o en un formato de cuadrícula y adherido a una superficie sólida o membrana [4,5]. Estos puntos normalmente se refieren a sondas que se diseñan de tal forma que cada sonda se une a una secuencia de ácido nucleico específico correspondiente a un gen particular en un proceso denominado hibridación [3]. La secuencia que se une a la sonda, que es denominada a menudo como target o diana, es marcada con una molécula detectable o con un componente fluorescente denominado fluorocromo [4].

El nivel de unión entre la sonda y su secuencia target es cuantificada mediante la medida de fluorescencia o señal emitida por la molécula marcada

cuando se escanea. Esta señal, a su vez, proporciona una medida de la expresión del gen específico que contiene la secuencia diana [2,3].

Aunque existen diversos tipos de “microarrays”, el propósito de este texto es solo considerar los “microarrays” de ARN. Estos “microarrays” se refieren a menudo como “microarrays” de doble canal o “microarrays” de dos colores porque dos muestras, cada una marcada con un fluorocromo diferente, son hibridadas en una sola de estas micromatrices. Como resultado de la combinación de dos muestras en un solo dispositivo, los niveles de expresión relativos pueden ser determinados usando este tipo de “microarrays” [1]. Las sondas en estos “microarrays” son oligonucleótidos, ADN complementario (ADNc) o fragmentos de productos de PCR [6,7]; confiriendo cada tipo propiedades diferentes a la matriz. A pesar de estas diferencias, todos los “microarrays” de oligonucleótidos son similares en términos de construcción del array, preparación de las sondas y análisis de datos [3,4].

Aunque actualmente una multitud de “microarrays” están disponibles comercialmente, cada uno diseñado para una especie específica o familia general de organismos; estas matrices están limitadas por la información disponible en bases de datos genómicas [2,8]. Sin embargo, solo los genomas de unas pocas especies han sido completamente secuenciados y están disponibles para el público. Por ejemplo, revisando el sitio web de Affymetrix, una de las principales casas comerciales de “microarrays”, los genomas disponibles para sus “microarrays” son: *Bacillus subtilis*, *Escherichia coli*, *Pseudomonas aeruginosa*, miembros del género *Plasmodium*, *Staphylococcus aureus* y miembros del género *Saccharomyces* [9,10]. “microarrays”, sin embargo, se pueden diseñar para muchas especies más mientras que sus secuencias genómicas estén disponibles para ese organismo particular o esa familia de organismos.

La exploración continua del genoma ha dado lugar a la necesidad de actualizar y reorganizar frecuentemente las micromatrices. Por tanto, con información continuamente actualizándose, los “microarrays” se refinan continuamente para mejorar la reproductibilidad y la detección de las señales de estos. Por último, cualquier discusión sobre la tecnología de “microarrays” sería incompleta sin un examen detallado de las diferentes limitaciones y complejidades intrínsecas. Dicho debate es vital para realizar experimentos de “microarrays” adecuadamente y analizar los datos de estos; la superación de estas limitaciones está en proceso [10,11].

## Hipótesis y planteamiento del trabajo

La expresión global de ARN mensajeros que se transcribe en un sistema biológico concreto puede determinarse usando “microarrays” de oligonucleótidos. Affymetrix ha desarrollado una tecnología apropiada para ello. Para determinados organismos modelo como *Saccharomyces cerevisiae*, *Drosophyla melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, *Mus musculus* y *Homo sapiens*, existen diversos “chips” comerciales disponibles que permiten cuantificar el nivel de expresión de prácticamente todo el genoma completo de estas especies. Para una interpretación apropiada de los resultados, se requiere un análisis de los datos complejo ya que las fuentes de variabilidad biológica y técnica son altas y el volumen de datos que se maneja muy elevado. Se requiere la aplicación de programas estadísticos apropiados y la búsqueda en bases de datos específicas para la determinación y agrupación funcional de estos genes.

Esta tecnología se está utilizando en muchos ámbitos de la Biología y de la Medicina, en concreto, en el hombre se está utilizando para estudiar la expresión génica diferencial que se produce en determinadas patologías, en diferentes organismos o en respuesta a fármacos específicos.

Los resultados de expresión global de mensajeros complementan los resultados obtenidos mediante técnicas proteómicas y son un paso más en la comprensión del funcionamiento de un sistema biológico en una situación experimental determinada.

En el presente trabajo se realiza una revisión bibliográfica con la que se pretende revisar y poner al día las aplicaciones más novedosas de los “microarrays” de ARN, haciendo una revisión del procedimiento experimental utilizado por Affymetrix que en la actualidad se está convirtiendo en una de las plataformas más utilizada para estos estudios.

## 2. OBJETIVOS

El objetivo general de esta revisión bibliográfica es hacer una actualización de lo publicado hasta la fecha relativo a los “microarrays” de ARN y sus aplicaciones en Biología y Medicina. Primeramente hay que poner de manifiesto que, actualmente, la tecnología de los “microarrays” de ARN es una de las herramientas experimentales que tienen una mayor proyección de futuro y utilidad en la investigación biomédica y biotecnológica. Se ha planteado esta revisión con estructura de un trabajo de investigación en el que, además de las aplicaciones, se han explicado los aspectos metodológicos relativos a este procedimiento.

Así este objetivo general puede dividirse en los siguientes objetivos específicos:

1. Describir el procedimiento experimental a seguir para realizar un “microarray” de ARN mediante la tecnología Affymetrix.
2. Hacer una revisión bibliográfica actualizada sobre la utilización de los “microarrays” de ARN en estudios en humanos, tanto en los relativos a patologías humanas como el cáncer como en enfermedades infecciosas y generación de vacunas.
3. Hacer una revisión actualizada de algunas de las aplicaciones de los “microarrays” en *Arabidopsis thaliana* y plantas.
4. Hacer una revisión actualizada de algunas de las aplicaciones de los “microarrays” en *Drosophila melanogaster*.
5. Hacer una revisión actualizada de algunas de las aplicaciones de los “microarrays” en *Caenorhabditis elegans*.

### **3. MATERIAL Y MÉTODOS**

En los experimentos de análisis de perfiles de expresión génica por medio de la tecnología de “microarrays” de Affymetrix pueden distinguirse las siguientes etapas: Diseño del experimento, preparación de la muestra, realización de los “microarrays”, análisis de imagen y análisis de resultados [12].

#### **3.1. Diseño experimental**

Es uno de los pasos críticos del experimento, ya que es esencial para la consecución de los objetivos planteados por el investigador. Se realiza previo al experimento y para ello deben tomarse decisiones relativas a aspectos diversos implicados en el propio experimento. Los principales aspectos a considerar son [13]: 1) objetivos que se quieren cumplir mediante la utilización de este, 2) tipos de muestras a analizar: individuo completo, órgano, tejido, biopsia, línea celular, así como su forma de empleo, almacenaje y mantenimiento, 3) protocolo de extracción de ARN: reactivos a utilizar y condiciones necesarias de temperatura, esterilidad y almacenaje del producto obtenido, 4) sistema de marcaje de la muestra: determinación del método de fluorescencia a utilizar así como pasos a seguir para este marcaje; 5) tipos de controles a emplear y distribución de muestras entre “microarrays”: utilización de un número adecuado de controles que permitan reducir al mínimo la posible variabilidad que se produce en estos experimentos; 6) diseño del “microarray”: tipo de “microarray” a utilizar según la muestra a analizar; plataforma para la construcción e hibridación del “microarray”: método de hibridación así como material y aparatos necesarios para esta hibridación; lugar físico donde se va a llevar a cabo el experimento con “microarray”, condiciones físicas y químicas necesarias y tiempo necesario; 7) número de réplicas biológicas y técnicas: entendiendo como réplicas biológicas las distintas extracciones de ARN del mismo origen que sean estadísticamente independientes y puedan controlar la variabilidad natural y como réplicas técnicas aquellas que dan precisión a los resultados pero no significado biológico y que incluyen tanto la repetición de genes en el “microarray” como la repetición de varias hibridaciones con el mismo ARN; 8) método para la obtención de imágenes y resultados y para su análisis: estableciendo qué aparatos de obtención de imágenes serán necesarios y que programas/softwarees serán de

más utilidad para el análisis de estas; y por último, 9) las limitaciones físicas y coste de estos ensayos incluyendo en estas limitaciones el número de arrays necesarios con respecto a los posibles y la cantidad de material necesario en contraposición al material del que se dispone.

Por lo tanto, de este diseño experimental saldrá la forma en la que se realizará el experimento y los métodos estadísticos que se deberán aplicar para obtener unos resultados satisfactorios

### **3.2. Preparación de la muestra**

#### *3.2.1. Extracción y aislamiento del ARN*

El proceso se inicia con la extracción del ARN de la muestra que se quiere analizar (tejido, por ejemplo). Este ARN es muy inestable y se degrada en pocos minutos, por lo que los tejidos deben ser frescos o congelados inmediatamente tras su obtención. Para este paso es imprescindible que se obtenga un ARN de alta calidad y pureza [14]. El método que da mejores resultados es la extracción de este ARN con TRIzol (Invitrogen). Este es un reactivo que se utiliza en el aislamiento de ARN de células y tejidos, es una solución mono-fásica de fenol e isotiocianato de guanidina que durante la homogeneización o lisis de la muestra mantiene la integridad del ARN, al mismo tiempo que estabiliza las células y disuelve los componentes celulares. La adición de cloroformo seguida de centrifugación separa la muestra en dos fases, una de ellas acuosa y la otra orgánica, el ARN total permanece exclusivamente en la fase acuosa y puede ser recuperado por precipitación con alcohol isopropílico. Una vez retirada la fase acuosa, el ADN y las proteínas restantes en la muestra pueden ser recuperadas por precipitaciones secuenciales de la misma [15].

#### *3.2.2. Purificación del ARN*

La clave de la purificación de ARN radica en evitar su degradación por acción de las ribonucleasas. De manera que todos los protocolos existentes para la purificación del ARN se basan en la rápida inactivación de dichas enzimas. Existen diversos métodos para llevar a cabo esta purificación. Uno de los más utilizados, es

el método de purificación a través del kit RNeasy de Quiagen. Este es utilizado para la purificación de ARN total de células de animales, tejidos animales, bacterias, levaduras y para la limpieza del ARN [16].

El procedimiento del kit RNeasy representa una tecnología bien establecida para la purificación del ARN. Esta tecnología combina las propiedades de unión selectiva de una membrana basada en sílice con la velocidad de la tecnología “microspin”. Las muestras biológicas son primero lisadas y homogeneizadas en presencia de un tampón altamente desnaturizante que contiene tiocianato de guanidina, que inmediatamente inactiva las RNasas para asegurar la purificación del ARN intacto. Posteriormente, es añadido etanol para proporcionar unas condiciones adecuadas, y la muestra es luego aplicada a una columna del “RNeasy Mini spin”, donde el ARN total se une a la membrana y los contaminantes son eliminados con eficacia. El ARN de alta calidad obtenido es diluido finalmente en 30-100 µl de agua. [17].

Este procedimiento es suplementado con un reactivo de estabilización del ARN, “RNAlater”. Este reactivo permeabiliza los tejidos sumergidos en él para estabilizar y proteger el ARN celular “in situ”. Este reactivo minimiza la necesidad de procesar inmediatamente las muestras de tejido o de congelar las muestras en nitrógeno líquido para el posterior procesamiento, pero no es válido para la estabilización del ARN de células animales, sangre, plasma o suero [18].

### *3.2.3. Comprobación integridad/ calidad*

La integridad/calidad de las muestras de ARN es fundamental en el contexto de los análisis de expresión génica mediante la tecnología de “microarrays”. El ARN debe estar libre de contaminación de proteínas, ADN u otro material celular como fenol, etanol y sales asociadas con el procedimiento de aislamiento de ARN ya que las impurezas pueden disminuir la eficacia de la transcripción inversa y reducir el nivel de la amplificación.

La calidad del ARN, es decir, la presencia o ausencia de contaminación en nuestra muestra, se puede medir mediante espectrofotometría. Una medida efectiva de la pureza de este ARN es el ratio de absorbancia leído a 260 y 280 nm ya que los ácidos nucleicos absorben con un pico máximo en 260 nm y las proteínas con un pico en 280 nm, por lo tanto una relación del cociente  $Abs_{260}/Abs_{280}$  de 1,8 o más

indicaría pureza de nuestra muestra. Sin embargo, un valor menor de 1,6 como resultado de ese cociente de absorbancias indicaría presencia de proteínas contaminantes o trazas de fenol. Por lo tanto, los valores obtenidos de absorbancia  $Abs_{260}$  a  $Abs_{280}$  deben estar en un rango de 1.7 -2.1 para una buena pureza. Una vez asegurada la calidad de nuestro ARN este debe ser suspendido en tampón Tris-EDTA, TE (10 mM Tris-HCl, 1 mM EDTA)

La absorbancia a 260 nm y la relación  $A_{260}/A_{280}$  sólo nos da una indicación de la cantidad de ARN y del grado de contaminación de impurezas orgánicas e inorgánicas, no nos aporta datos sobre el nivel de degradación de ARN. Para determinar el grado de degradación de este ARN, es decir, su nivel de integridad, se utiliza el método denominado “Número RIN”. RNA Integrity Number (RIN) [19] es un número que aplica el algoritmo del software del Bioanalizador 2100 Bioanalyzer de Agilent Technologies y que es indicador del grado de degradación de la muestra de ARN (Figura 1). Este instrumento permite la caracterización cualitativa y cuantitativa de muestras de ARN total o ARN mensajero (ARNm) basándose en los patrones de movilidad electroforética de las muestras en una matriz de gel que se introducen en unos chips formados por una serie de microcanales y depósitos en miniatura interconectados a través de los cuales migran electroforéticamente los ácidos nucleicos.[20]



Figura 1. Pasos en la medida de la integridad del ARN mediante el Bioanalizador Agilent 2100.

Imagen procedente de: <http://www.gene-quantification.de/rna-integrity.html>

Además, la introducción del número de integridad del ARN/ RNA Integrity Number (RIN) facilita la determinación objetiva de la calidad del ARN que se utilizará en experimentos posteriores. El RIN es una herramienta diseñada para ayudar a los científicos a estimar la integridad del ARN total de las muestras. La extensión RIN asigna automáticamente un número de integridad a la muestras de ARN total eucariota analizada en el Bioanalizador Agilent 2100. Este valor que va desde el 1 al 10, se basa en la cuantificación de productos de degradación del ARN a lo largo de toda su migración electroforética. Un ARN intacto posee un RIN de 9-10, mientras que un ARN degradado tendría un RIN <6. Para las aplicaciones de PCR a tiempo Real y microarrays se recomienda un RIN mínimo de 7, aunque depende del tipo de muestra [21].

### **3.3. Procedimiento de “microarray” mediante Affymetrix**

#### *3.3.1 Amplificación del ARN mensajero*

Una vez realizada la extracción del ARN total de las muestras que se quieren comparar y comprobada su calidad e integridad se lleva a cabo la amplificación de ARN mensajero, ya que normalmente la concentración que se obtiene es demasiado baja para realizar la posterior hibridación con los “microarrays” donde se medirán la expresión [22].

#### *3.3.2. Transcripción inversa y marcaje*

Una vez obtenido, purificado y amplificado el ARN total se procede a la síntesis de la primera hebra de ADNc, esta se lleva a cabo mediante el método tradicional de transcripción inversa, utilizando para ello nuestro ARN aislado como molde, la enzima transcriptasa inversa, cebadores de ADN, la enzima RNAasa, desoxinucleótidos trifosfato (dNTP) y un tampón para proporcionar las condiciones óptimas [23].

Es en este instante, en el momento de obtención del ADNc, cuando se procede al marcaje de las moléculas para la posterior hibridación y lavado (Figura 2). Según las necesidades experimentales que el ensayo requiera, el tipo de marcaje y

las moléculas utilizadas para este serán diferentes, distinguiéndose así dos tipos de marcajes, marcajes de un solo color o marcajes de dos colores.

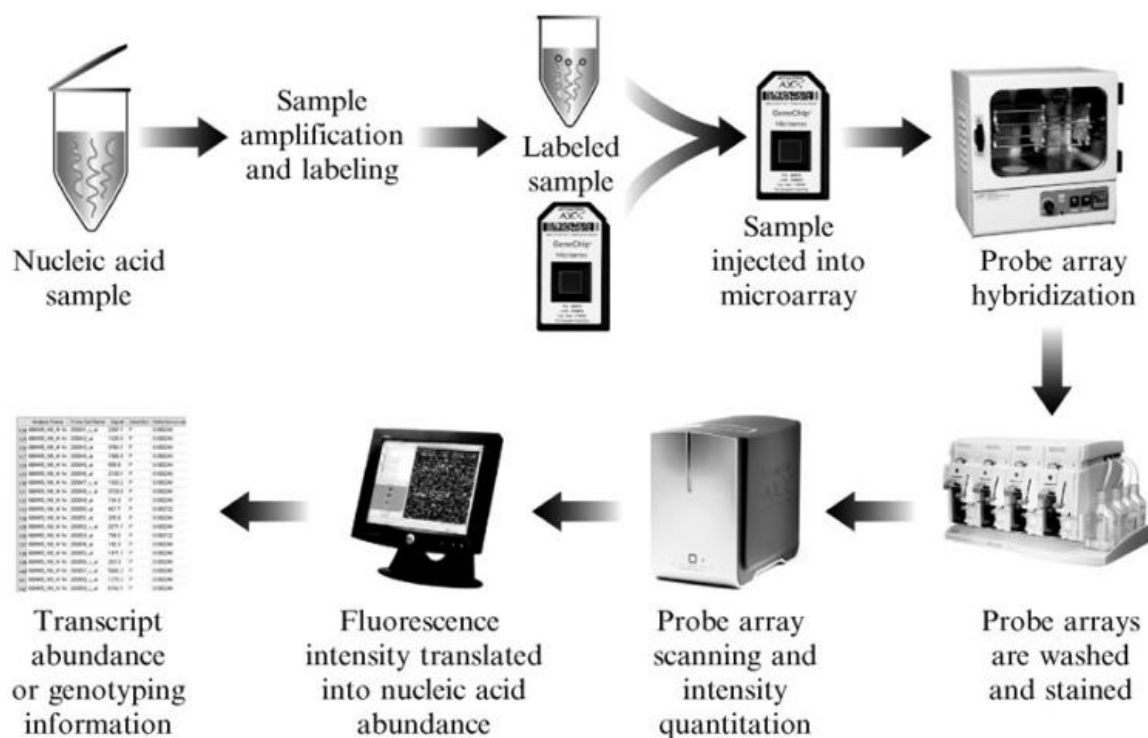


Figura 2. Imagen ilustrativa de los pasos a seguir en el procedimiento de microarray mediante "Affymetrix".

Imagen procedente de:

[http://www-nmr.cabm.rutgers.edu/academics/biochem694/reading/DalmaWeiszhausz\\_2006.pdf](http://www-nmr.cabm.rutgers.edu/academics/biochem694/reading/DalmaWeiszhausz_2006.pdf)

Estudios de comparación de expresión génica en los que se miden la abundancia relativa de transcritos entre dos muestras diferentes, una muestra experimental y una muestra de referencia o control, requieren de la utilización de dos colores distintos para el marcaje, pues están basados en la hibridación competitiva con un ADNc diana de 2 sondas marcadas con diferente color, por tanto, se marca con un fluoróforo la muestra experimental o "enferma" y con otro fluoróforo la muestra de referencia/ control o "sana" [24]. En este caso, los fluoróforos más utilizados y que proporcionan mejores resultados son los fluoróforos cianinas Cy3, utilizado habitualmente para el marcaje de la muestra experimental, que reporta color verde y cuyas longitudes de onda de absorción y de excitación son 532 nm y 550-600 nm respectivamente; y el fluoróforo cianina Cy5, usada con frecuencia para

marcar la muestra control, que reporta color rojo y que posee sus longitudes de onda para absorción y excitación en 635 nm y 655-695 nm [25]. Estos fluoróforos cumplen una serie de condiciones que los convierten en útiles para la comparación de las muestras: son fotoestables, tienen una elevada intensidad de fluorescencia, presentan un espectro de fluorescencia bien separado y son, por tanto, resistentes al *fotobleaching* o apantallamiento [26].

Por otro lado, sin embargo, estudios de medida del nivel o de presencia o ausencia de expresión de una sola muestra solo requieren el uso una molécula que aporte fluorescencia en su marcaje, esta molécula puede ser bien un solo fluoróforo como el Cy3 o una técnica no radiactiva que utiliza la molécula biotina. En este caso hablamos de “microarrays” de un color, en los cuales, en cada “array” solamente se hibrida una muestra, por lo que no se da la hibridación competitiva como sucede en los arrays de dos colores [26].

Una vez marcadas nuestras moléculas de interés se procede a la fragmentación de estas hasta obtener secuencias suficientemente pequeñas para que puedan hibridar en nuestro “microchip”, para ello, estas son incubadas en un tampón de fragmentación a 94°C durante 30 minutos produciendo fragmentos de un tamaño de entre 35 a 200 nucleótidos. Esta fragmentación reduce posibles interferencias de la estructura secundaria del ARN, y minimiza los efectos de múltiples interacciones con otras moléculas [27].

### 3.3.3. Descripción del microarray (Affymetrix GeneChips's)

Un microarray, Mikro (del griego, pequeño) y array (del inglés, distribución ordenada), también denominado GeneChip, “DNA-chip” o simplemente *array*, es un “chip” del tamaño aproximado de un sello formado por una matriz inerte de 1.28 cm<sup>2</sup> de distinta naturaleza (vidrio, plástico, silicona, etc.) dividida a su vez en múltiples celdas de 50µm<sup>2</sup> [28].

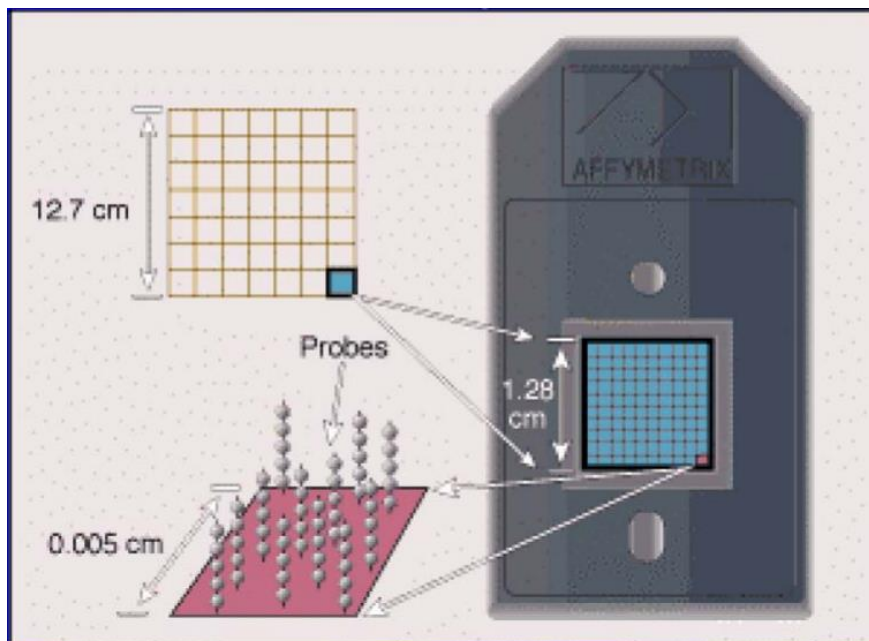


Figura 3. Descripción gráfica de un GeneChip de Affymetrix

Imagen procedente de: <https://www.msi.umn.edu/sites/default/files/IntroMicroarray.pdf>

Nos centraremos en un tipo específico de microarrays como son los microarrays de oligonucleótidos, de los cuales los más utilizados por su auge y comercialización en el mercado, son los “microarrays” GeneChip® de la compañía *Affymetrix*. Un oligonucleótido es una cadena de nucleótidos, que en el caso de los microarrays de *Affymetrix* son 25. Para cada gen, este microarray contiene un conjunto de sondas o “probe-set” que son oligonucleótidos de 11 tipos distintos (en el caso del “microarray” modelo U133 GeneChip Human); esto es, 11 hebras simples distintas de oligonucleótidos de tamaño 25. Cada oligonucleótido se encuentra miles de veces adherido en cada celda del “microarray” constituyendo una sonda o “probe” [29].

Para cuantificar el número de “falsas” hibridaciones de estas hebras simples cada celda del “microarray” está dividida en dos partes, es decir, para cada gen existen dos tipos de sondas, una de homología perfecta (PM, Perfect Match) de 25 bases y otra que presenta un error deliberado/mutación (MM, Mismatch) en la zona central. Es decir, las secuencias de las sondas PM y MM son las mismas, excepto por un cambio de nucleótido en la zona central, más exactamente en la posición número 13 de la PM, que en la sonda MM en su lugar tiene otro nucleótido diferente, su nucleótido complementario [30].

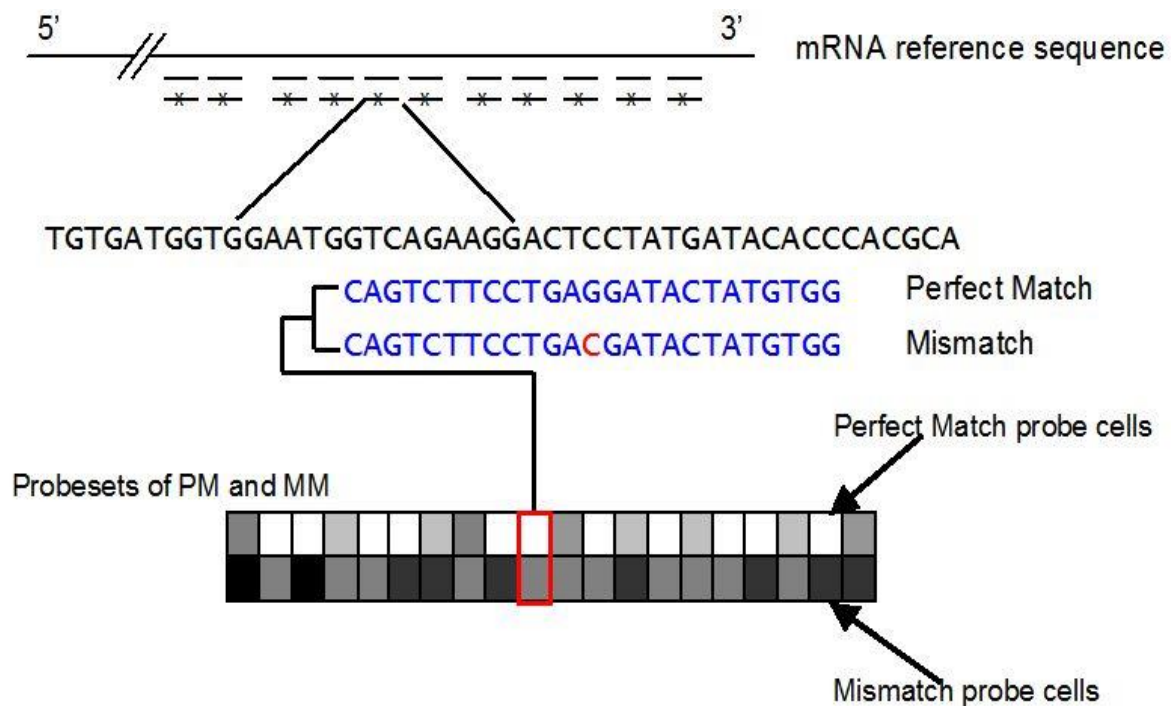


Figura 4. Descripción de las dos tipos de sondas de un microarray, una de homología perfecta (PM, Perfect Math) de 25 bases y otra de error deliberado/mutación (MM, MisMatch) en la zona central de la secuencia.

Imagen procedente de: <http://www.intechopen.com/books/thermodynamics-fundamentals-and-its-application-in-science/thermodynamics-of-microarray-hybridization>

Son múltiples los “GeneChips” comercializados por la casa comercial “Affymetrix”, todos ellos basados en las distintas secuenciaciones de los genomas de los organismos, destacando por su mayor uso [31] el GeneChip® Human Genome (Especie Humana), GeneChip® Rat Genome (Especie Rata), GeneChip® Genome Mouse (Especie Ratón), GeneChip® Genome *Arabidopsis* ATH1 Genome Array (Especie *Arabidopsis*), GeneChip® *C.elegans* Arrays (Especie *C.elegans*), GeneChip® *Drosophila* Genome Array (Especie *Drosophila*) y GeneChip® *E. coli* Genome Array (Especie *E.coli*)

### 3.3.4 Hibridación en el “microarray”

Los “microarrays” de Affymetrix usan la propiedad conocida como hibridación, es decir, la propiedad biomolecular fundamental del ADN que es la “complementariedad” de las bases nitrogenadas, de cómo A y T, y C y G, se unen específicamente por puentes de hidrógeno [27]. Durante la hibridación, las muestras de material genético marcadas con los distintos fluoróforos (target sequences/secuencias diana) se unirán a sus complementarias ya inmovilizadas en el soporte del chip, permitiendo la identificación y cuantificación del ADN presente en la muestra.

Para llevar a cabo la hibridación, se mezclan ambas poblaciones de ADNc marcados (una marcada con Cy3 y otra con Cy5) y se incuban conjuntamente en el *array* para que cada especie de ADNc hibride específicamente y de forma competitiva con su ADNc que ya se encuentra inmovilizado en el *array*. Cuanto mayor sea la hibridación de un determinado ADNc marcado con en ADNc del *array*, mayor será la expresión tisular original del ARNm correspondiente [27].

Una vez producida la correspondiente hibridación se procede al lavado del microchip mediante diversos lavados de astringencia, para eliminar las secuencias no adheridas en el microarray, y a su secado para el posterior análisis [27].

El nivel de hibridación existente se calcula determinando la longitud de onda emitida por cada uno de los ADNc incubados después de ser excitados por una luz fluorescente. Para ello, se utiliza un sistema lector del *array*, un escáner laser, que detecte las señales fluorescentes emitidas de ambas muestras. Estos escáneres láser son sistemas que emplean fuentes monocromáticas de excitación y que posteriormente recogen los fotones de emisión de los fluoróforos en un receptor en el que los fotones incidentes son transformados en electrones y amplificados mediante efectos fotoeléctricos. Para cubrir toda la superficie del microarray estos dispositivos cuentan con motores que, o bien mueven el microarray para permitir el barrido de toda la superficie a escanear, o bien mueven el sistema de excitación y detección [28].

El **escáner Agilent G2565CA Microarray** (Figura 5) es uno de los más utilizados en estos estudios de expresión génica. Este es un sofisticado escáner láser que actúa excitando cada molécula de ácido nucleico simple marcada con el fluoróforo mediante una luz monocromática producida por un láser que tiene incorporado y colectando la luz de emisión (fluorescencia), convirtiendo así la corriente de fotones en valores digitales que pueden ser almacenados en un ordenador como un archivo de imagen (.TIFF). Esta habilidad de medir fluorescencia de dos muestras marcadas simultáneamente facilita los estudios de “microarrays” de dos colores. Este escáner provee además un escaneado rápido, de alta calidad y automático [32].



*Figura 5. Agilent G2565CA Microarray Scanner*

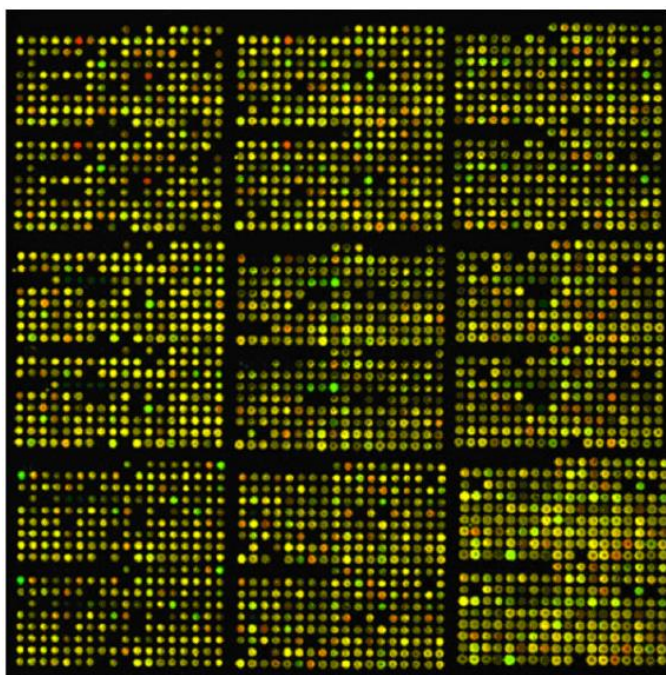
Imagen procedente de: <http://www.pacificlab.co.nz/shop/products/q2565ca-high-resolution-scanner-bundle-agilent/1/name>

Este sistema lector del array asigna un código informático de colores a la cantidad de fluorescencia emitida. Si hay mayor hibridación del ADNc de la condición patológica (roja), predominará el componente rojo de la emisión, y viceversa, cuando sea mayor la hibridación del ADNc de la condición control, el componente verde de la emisión será el que predomine. En el caso de que la hibridación de los dos tejidos estudiados sea similar, se asignará un color amarillo/anaranjado [33].

Hay que tener en cuenta que esta diferencia de emisión debe evaluarse en cada uno de los pocillos del *array*, es decir, para cada uno de los ADNc (genes) evaluados. De esta manera se compara la expresión de todos los genes representados en el microarray en el tejido de estudio frente al tejido control (*differential display*) [33].

### 3.4. Análisis de imagen de microarray

Para un típico experimento de microarray en el que se compara la expresión génica diferencial de dos situaciones, se producen dos archivos después del escaneado del microchip, una para cada tinte fluorescente. De estas dos imágenes se obtiene la imagen final (Figura 6) por superposición de estas [34].



*Figura 6. Imagen superpuesta de los dos canales de color de un chip, canal rojo y canal verde.  
Imagen procedente de:*

*Kelmansky DM (2007) Aspectos Estadísticos de “microarrays”. Dpto. de Matemáticas. Instituto de Cálculo.*

Un ejemplo de la imagen obtenida tras la lectura de un “microarray” es la que se muestra en la Figura 6. En ella se distinguen círculos coloreados que corresponden a los puntos del “array”, en los cuales el brillo y el color indica la

cantidad y tipo de fluorescencia retenida y por tanto, del comportamiento del gen correspondiente. Los colores de cada spot de esta imagen son representaciones artificiales de las intensidades, el color rojo indica que el gen está más expresado en el tejido patógeno (marcado con Cy5, color rojo) que en el sano (marcado con Cy3, color verde), el color verde indica la situación contraria, el gen se encuentra más expresado en el tejido sano que en el patógeno, y por último, el color amarillo indica que el gen se encuentra expresado con igual intensidad en ambos tejidos [35].

La cuantificación de la señal a partir de las imágenes es un proceso muy importante ya que determina los valores que posteriormente se analizarán. En la actualidad existen múltiples herramientas, tanto comerciales como de libre distribución, diseñadas exclusivamente para analizar las imágenes de “microarrays”. Las imágenes a analizar suelen ser 2 archivos en formato “TIFF” en escala de grises de 16 bits, es decir, cada pixel puede tener una intensidad de señal entre  $2^0$  y  $2^{16}$ , es decir, entre 1 y 65.535 bits [36].

El proceso de análisis de la imagen se puede dividir en 3 etapas [37]. En primer lugar se localizan los puntos a partir de los datos que proporciona el fabricante del “microarray”, es decir, cuántos puntos hay, como están agrupados, separación teórica entre puntos. A continuación se realiza la segmentación, consistente en identificar qué píxeles corresponden a puntos del “microarray” (“foreground”) y que píxeles son fondo (“background”). Por último, se procede a la cuantificación, a menudo esta se realiza como el promedio o la mediana de las intensidades de los píxeles que forman el punto. En este apartado del análisis llevado a cabo mediante un software informático, es importante que este software proporcione varias medidas que puedan ser empleadas como indicadores de la calidad del punto. Medidas típicas son los índices de circularidad, diámetro máximo, perímetro y homogeneidad de la señal del punto.

También es necesario obtener una medida del fondo, es decir, el nivel de señal en los píxeles reconocidos como fuera de los puntos. Este valor normalmente se sustrae de la intensidad en el punto para obtener la intensidad neta.

### 3.5 Normalización y análisis de datos

Una vez se han analizado las imágenes y almacenado los datos, la última parte del proceso consiste en realizar el análisis estadístico de estos [38]. Previo al análisis de datos se llevan a cabo dos técnicas de eliminación de posibles errores: Control de calidad de los datos y normalización de los datos.

#### *Control de calidad de los datos*

Técnica aplicada para detectar valores incorrectos para su posterior exclusión de los análisis. Estos valores incorrectos pueden surgir por problemas de calidad de los experimentos o por accidentes en su manipulación, como rasguños o imperfecciones en la superficie del “microarray”. Su detección puede realizarse empleando límites de tolerancia en los indicadores de calidad del punto. Es muy útil que el diseño del “microarray” incluya réplicas de una misma diana. La comparación de las réplicas permite identificar casos discordantes [39].

#### *Normalización de los datos*

Las dos imágenes de un “microarray” se obtienen por separado, cada una con una longitud de onda diferente (normalmente rojo y verde) y una potencia que debe ajustarse de manera independiente para evitar saturación por lo que el ajuste independiente hace que las 2 imágenes no sean comparables en cuanto a intensidad si no se normalizan previamente [39].

El objetivo de la normalización es identificar y eliminar fuentes de variabilidad sistemática introducida por el proceso técnico que no está relacionada con el nivel de expresión: Diferente eficiencia en el marcaje con el colorante, diferente cantidad de ARN en Cy3 y Cy5, diferentes parámetros de escáner, efectos espaciales del chip (aguja, zona...) [39].

Es necesaria para asegurar que las diferencias en intensidades se deben a diferencias de expresión real, no a artefactos de impresión, hibridación o escaneo. Existen múltiples métodos para la normalización. El método más recomendado emplea modelos de regresión de datos suavizados mediante técnicas no

paramétricas que capturen la no linealidad; como el método *Lowess/Loess* de *regresión robusta* [40].

En la Figura 7 se muestran los datos de un experimento en el que el ADNc de un mismo tejido se marcó con dos fluorocromos (R(CY5) y G(Cy3)). Tras la aplicación del método *Lowess*, se esperaría que la nube de puntos se situara alrededor de la recta con pendiente 1 que pasa por el origen. Puede apreciarse que la nube de puntos está desplazada por debajo de la línea teórica y muestra una desviación no lineal. La magnitud de la dispersión corresponde a la variabilidad de la técnica, pues no puede deberse a las muestras ya que ambas corresponden al mismo tejido [39].

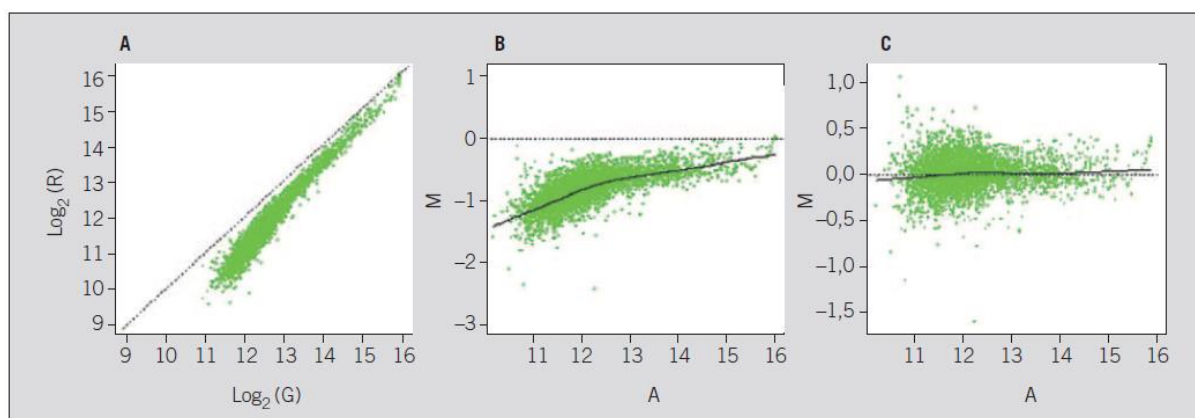


Figura 7. Normalización de las intensidades del canal rojo (R) y verde (G) obtenidas en un experimento de reproducibilidad, es decir, la misma muestra marcada con diferentes fluorocromos.

A: gráfica  $\log(R)$  frente a  $\log(G)$ , donde se aprecia menor señal en la muestra marcada con el fluorocromo rojo.

B: gráfica  $M = \log(R/G)$  frente a  $A = \log(R*G)/2$  sin normalizar. Es simplemente una transformación de la gráfica A donde se aprecia mejor que las diferencias siguen una curva no lineal en función de la intensidad.

C: gráfica  $M$  frente a  $A$  después de normalizar mediante regresión local robusta (*lowess/loess*)

Imagen procedente de:

<http://www.sc.ehu.es/ccwbayes/doctorado/UsodeChipsDeADN.pdf>

Para obtener una mejor visualización, normalmente se transforman los valores según las fórmulas:

$$A = \log(R \cdot G) / 2 \quad Y \quad M = \log(R / G)$$

El valor A corresponde al logaritmo de la intensidad media (geométrica) de los 2 colores y el M al logaritmo de la razón de las intensidades.

La Figura 7b muestra los mismos datos después de su transformación. La curva central a la nube de puntos muestra la estimación no paramétrica de la relación entre M y A, obtenida por el método lowess/loess  $L = \text{lowess}(A, M)$ .

La normalización consiste simplemente en restar a cada punto la diferencia entre M y L. El resultado es el que se apreciaría en la Figura 7c [39].

El modelo de normalización puede incluir covariables, además de la intensidad media, de las que dependa el valor de la señal. Se recomienda, por ejemplo, estratificar el modelo según la aguja del robot que produjo el “microarray”, o emplear información espacial para eliminar posibles heterogeneidades en la intensidad de la hibridación en diferentes zonas del “microarray” (normalización en 2 dimensiones). Otros métodos, útiles para normalizar múltiples “microarrays”, se basan en igualar no sólo el promedio de razón de intensidades de cada fluorocromo, sino la forma de la distribución (normalización por cuantiles). Este método se usa de forma rutinaria con “microarrays” de oligonucleótidos y puede emplearse también para los “microarrays” de ADNc [41].

### *Tratamiento de valores perdidos*

En el proceso de análisis de múltiples “*microarrays*” puede ser importante dar un tratamiento adecuado a los valores perdidos, ya que trabajar con los puntos con información completa en todos los “*microarrays*” puede suponer una pérdida importante de genes valorables. A menudo se emplean técnicas de imputación basadas en medias condicionales del gen respecto al conjunto de los “microarrays” o respecto a los valores de los puntos vecinos [42].

### 3.5.1 Análisis de la diferencia de expresión a nivel de ARN

Una vez que la imagen ha sido obtenida, corregida, normalizada y se han extraído los datos necesarios, se procede a determinar qué genes varían su nivel de expresión en función del tejido que ha sido analizado, por ejemplo, tejido normal y tejido tumoral.

Normalmente, un estudio de “microarray” típico emplea relativamente pocos casos y lo que se quiere investigar es si existen diferencias en miles de variables (genes), las técnicas estadísticas clásicas no son adecuadas sin las correspondientes modificaciones [43].

Por un lado, no puede asegurarse la normalidad de los datos que requieren estas pruebas estadísticas clásicas. Por otro lado, la tasa de resultados falsamente positivos puede ser muy elevada si no se emplean correcciones que tengan en cuenta la multiplicidad de hipótesis que se prueban. Las soluciones propuestas para estos problemas son emplear “tests” de permutaciones para evaluar empíricamente el nivel de significación para cada gen y, posteriormente, controlar la tasa global de resultados falsos positivos mediante un ajuste de los valores  $p$  que tengan en cuenta las múltiples comparaciones realizadas [44].

Se emplea un test estadístico clásico como la *t de Student* u otra prueba adecuada según el tipo de variable o diseño de estudio. Se calcula, para un gen, el valor del test que compara los grupos (valor observado) y se repite el mismo test múltiples veces de manera que cada vez la asignación del grupo, normal o tumor, se cambia al azar. De esta manera se simula la situación en la que no hay diferencias, pues la asignación de cada muestra a uno u otro grupo es aleatoria. Finalmente, el valor  $p$  se calcula a partir del percentil que ocupa el valor observado en la distribución de valores obtenidos por permutación. Con unos cientos de permutaciones suele ser suficientes para obtener el valor  $p$ , pero pueden precisarse varios miles si se desea diferenciar entre valores pequeños (muy significativos). Este proceso se repite por separado para cada gen [45].

Los métodos para evitar resultados falsos positivos corrigen los valores  $p$  para controlar el nivel global de significación. Existen varios métodos, desde el más sencillo de *Bonferroni*, que consiste en considerar significativos sólo aquellos valores  $p$  inferiores al cociente entre *alfa* y el número de tests. Este método es muy conservador pues asume que cada test es independiente, lo cual probablemente es

falso en el contexto de análisis de múltiples genes en “*microarrays*”, ya que la expresión de algunos genes puede estar correlacionada. Otros métodos de control de nivel global de significación emplean procedimientos adaptados y tienen en cuenta la correlación entre tests. Los más empleados en el contexto de “*microarrays*” utilizan también métodos de remuestreo como el *min-P* y *max-T* [44].

### 3.5.2 Análisis de datos

En cuanto al análisis de datos de “*microarrays*”, existe una amplia disponibilidad de softwares de uso libre que pueden ser utilizados para obtener datos a partir de la imagen del “*microarray*” una vez procesada, mereciendo una mención especial el proyecto *Bioconductor*.

**Bioconductor software:** Es un proyecto internacional de software libre para el análisis e interpretación de datos genómicos, escrito en el lenguaje estadístico R (<http://www.r-project.org>), que incluye gran número de procesamientos computacionales para el análisis de datos de expresión de genes [46]. El proyecto Bioconductor (<http://www.bio.conductor.org>) dispone de una serie de paquetes que contienen un amplio rango de aplicaciones estadísticas para varios tipos de análisis genómicos: para “*clustering*” de valores de expresión [47], “*multtest*” [48] y “*maanova*” [49] para comparación de múltiples condiciones experimentales, “*marray*” [50] que contiene las funciones *lowess/loess* para realizar una regresión local, “*samr*” [51] que implementa el método SAM para determinar expresión diferencial, entre otros.

**MeV** (*The Institute for Genomic Research*, EE.UU.): este es uno de los programas más populares para el análisis de datos de “*microarrays*”, debido a la gran variedad de métodos matemáticos que tiene implementado [52]. Entre sus opciones, permite las transformaciones básicas de los datos, el filtrado y la normalización, el agrupamiento de genes o de condiciones experimentales con el empleo de diferentes medidas de distancia y métodos tales como *k-Means* [53], HCL (*Hierarchical Clustering*) [54], SOM (*Self Organizing Maps*) [55] y SOTA (*Self Organizing Tree Algorithm*) [56] Para la búsqueda de genes diferencialmente expresados en determinadas condiciones experimentales cuenta con métodos como la prueba *t*, ANOVA [57] y SAM con control del FDR. Este software se encuentra disponible en <http://www.tigr.org/software/tm4/>.

### 3.5.3 Verificación de resultados

El estudio de decenas de miles de genes en cada experimento simultáneamente, favorece el aumento del número de falsos positivos en los resultados de los análisis estadísticos. De aquí que se utilicen otras técnicas experimentales para confirmar estos resultados.

**RT-PCR cuantitativa en tiempo real (Q-RT-PCR)** es una de las técnicas experimentales más sensibles para detectar y cuantificar el ARNm en tejidos. Esto hace que sea uno de los métodos más robustos y comúnmente usados para verificar la expresión de genes como resultado del análisis estadístico de datos de “microarrays” [58].

Otras técnicas como el análisis por “*Northern blot*” y los ensayos de protección de ribonucleasas también se emplean para estos propósitos.

### 3.5.4 Clasificación de muestras o genes

A partir de los genes que muestren expresión diferencial se puede buscar patrones con el objetivo de clasificar las muestras. Además posteriormente, tras evaluar las características de las muestras agrupadas, se pueden identificar “genes prototipo” que definan los grupos. También se pueden buscar grupos de genes que muestren un patrón de expresión diferencial similar, útil para asignar funciones a genes o secuencias que se expresen y que hasta ahora sean desconocidas [59].

Existen múltiples métodos de clasificación automática pero los más utilizados se basan en técnicas de análisis de conglomerado o *cluster jerárquicos*, el objetivo de esta técnica es el de formar grupos o clase de datos, de tal forma que los datos de un mismo grupo comparten una serie de características y similitudes mientras que los datos de grupos distintos tienen mayores diferencias [60].

El clustering jerárquico es usado para producir lo que se conoce como “mapas de calor”. Los *mapas de calor* presentan una cuadrícula de puntos coloreados donde cada color representa un valor de la expresión génica en la muestra.

Esta técnica genera un gráfico en forma de árbol (dendograma) con la jerarquía obtenida. El problema suele ser decidir por dónde cortar el árbol para definir el número de grupos identificados.

Los análisis de cluster se pueden aplicar para grupos muestras o agrupar genes, pudiendo realizarse también las 2 agrupaciones simultáneamente, lo que permite interpretar más fácilmente los resultados.

La Figura 8 muestra un ejemplo de una doble clasificación, en el cual las muestras se encuentran situadas en columnas y los diferentes genes en filas, donde se aprecian 2 grupos de muestras y dos de genes. La imagen en pseudocolor muestra en rojo genes con expresión aumentada y en verde genes con expresión disminuida [61].

Cuando las muestras están caracterizadas por variables que interesa diferenciar, los métodos de clasificación supervisada son más eficaces. Ejemplos de aplicación son la discriminación en función del tipo de tejido (normal o tumoral) o en función del tipo celular o pronóstico. Entre estos métodos destacan los de regresión discriminante con múltiples versiones (lineal, cuadrática, logística, etc.), las redes neurales, los árboles, etc. Una buena revisión de estos métodos aplicados al análisis de microarrays puede encontrarse en el libro de Hastie *et al* [62].

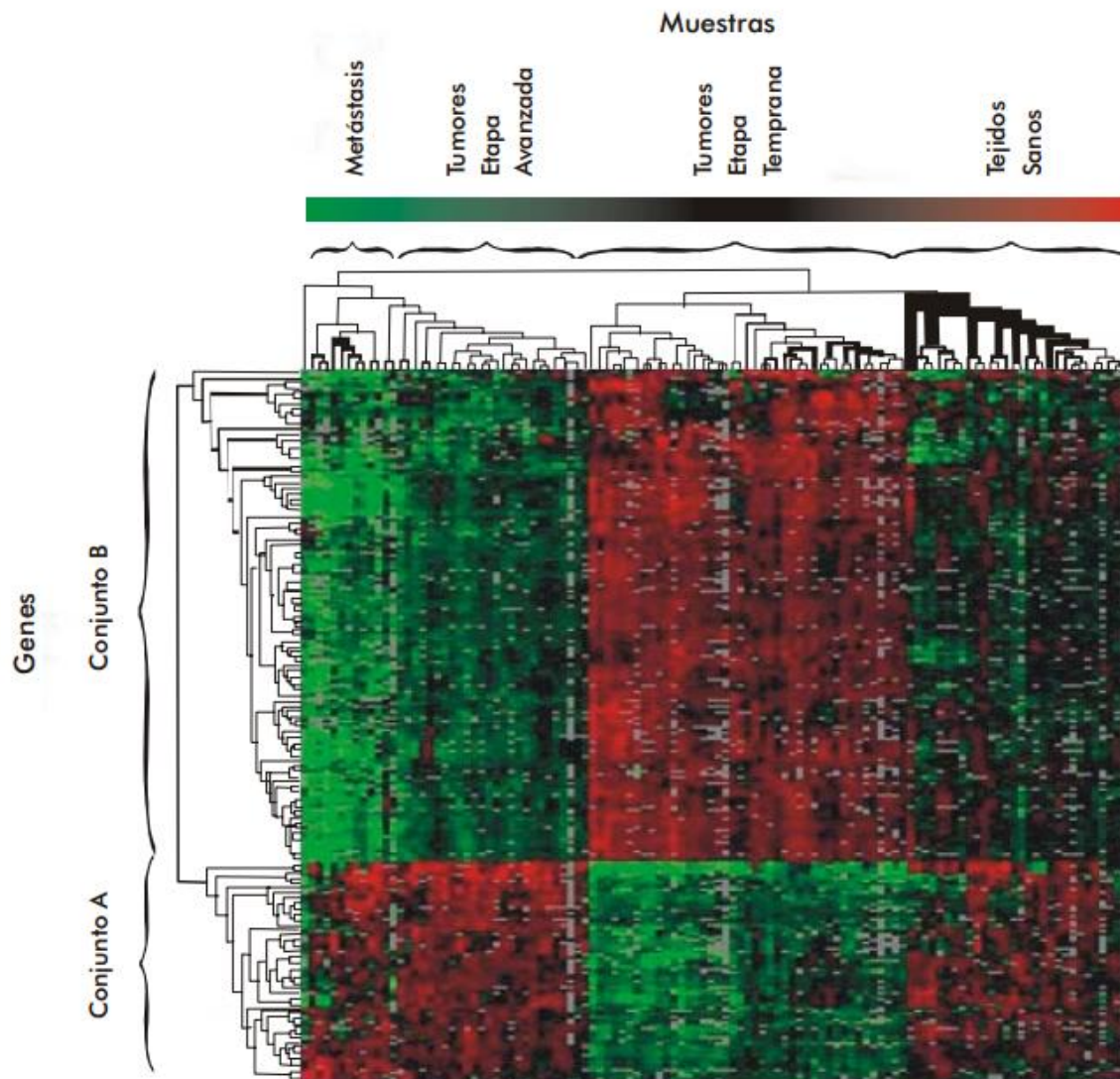


Figura 8. Ejemplo de un “mapa de calor”. Agrupamiento bidimensional de 242 genes expresados diferencialmente entre tejidos sanos y tumorales de próstata. Los datos fueron obtenidos del estudio de Lapointe *et al.* . Este análisis exploratorio muestra como los genes del conjunto A se reprimen en las etapas 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

Imagen procedente de: Miranda J, Bringas R (2008) Análisis de datos de microarreglos de AND. Parte I: Antecedentes de la tecnología y diseño experimental. Centro de Ingeniería Genética y Biotecnología, CIGB, 25:82-89

### **3.6 Relaciones funcionales. Bioinformática aplicada a los “microarrays”**

La Bioinformática se puede definir como la ciencia que maneja, procesa e interpreta la información procedente de las secuencias y estructuras biológicas. Uno de los principales destinos de esta masiva información biológica es la fabricación de bases de datos de secuencias. Los abundantes proyectos de secuenciación del genoma que se han llevado a cabo en los últimos años, han contribuido a un crecimiento exponencial de estas bases de datos. Además, los avances tecnológicos, sobre todo en el campo de los “microarrays” han llevado al campo de la genómica y la proteómica a estudiar genes, proteínas y procesos de expresión génica que regulan el funcionamiento celular. Con todo esto, una característica del campo de la Bioinformática es la avalancha de datos que existen hoy en día o que se prevé tener en un periodo corto de tiempo y que necesita un proceso eficaz para extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior. Por lo tanto, el avance continuo en la tecnología del “microarray” es debido, en parte, a los avances producidos en el campo de la Bioinformática, pues es ésta la que surte de información mediante las bases de datos y la que dispone de herramientas útiles para el manejo e interpretación de los datos empleados en los experimentos de “microarrays”.

## 4. RESULTADOS Y DISCUSIÓN

La herramienta de los “*microrrays*” es una tecnología muy reciente que se está aplicando, a nivel de investigación, en estudios realizados en especies modelo para las cuales se han diseñado “microchips” específicos. Son aquellas especies en las que se conoce muy bien su genoma, lo cual ha permitido tener sondas específicas de un gran número de genes para poder diseñar los “microchips”. En base a esto, se ha dividido esta sección en subsecciones referentes a las diferentes aplicaciones de los microarrays en los diferentes organismos. En el caso del hombre, esta metodología es tan reciente que todavía se encuentra limitada al ámbito de la investigación, sus aplicaciones clínicas aún están en fase de desarrollo.

### 4.1 Aplicación de los “microarrays” en el estudio de enfermedades. Cáncer.

El cáncer es una enfermedad altamente compleja que puede abarcar múltiples alteraciones genómicas, incluyendo mutaciones puntuales, translocaciones, amplificaciones génicas, modificaciones epigenéticas, deleciones, “splicing” aberrante y expresión génica alterada. Estos cambios pueden ser hereditarios o adquiridos somáticamente durante la progresión de una célula normal a una célula cancerígena. Los resultados de las investigaciones en este campo han distinguido como estas perturbaciones genómicas conducen a las células cancerígenas a sobrevivir alterando el mecanismo de control del ciclo celular, reparación de ADN, diferenciación, apoptosis, vascularización del tumor y metabolismo [63]

Son varios los ejemplos de aplicaciones del “microarray” Affymetrix en la investigación del cáncer [64], abarcando estas desde el descubrimiento a la validación y utilidad clínica, entre estas incluye:

- a) La comprensión de las bases moleculares de la carcinogénesis. Los “microarrays” han facilitado enormemente el estudio global de los patrones de expresión génica que conducen a la pérdida de la regulación del ciclo de división celular y de la muerte celular programada (apoptosis), como mecanismos esenciales de control involucrados en la transformación maligna. Particularmente en neoplasias inducidas por virus, los “*microarrays*” han permitido dilucidar

algunas de las vías de señalización que emplean para inducir la transformación [65,66].

- b) La clasificación y el pronóstico. El estudio y clasificación de estos tumores mediante el uso de “microarrays” ha facilitado la definición de patrones de expresión diferenciales que hacen posible un acercamiento más profundo a su origen molecular [67].
- c) Seguimiento de metástasis. En tumores tales como el melanoma se ha podido determinar la expresión diferencial negativa de genes relacionados en la adhesión celular y con el complejo mayor de histocompatibilidad, que hacen posible diferencias los melanomas con potencial metastático [68].

Como ejemplo de estos estudios he seleccionado el trabajo de Bratslavsky y col. [69] en el que se estudió el perfil de expresión génica global de tumores de diferentes tamaños en pacientes con tumores renales resultado de un daño en el mecanismo regulador del gen supresor del tumor de von Hippel-Lindau (VHL) con pérdida de función por mutación somática y silenciamiento.

Para una mejor caracterización biológica de los eventos asociados con este crecimiento tumoral, se evaluó la alteración en la expresión génica en 11 pacientes con dos tipos de tumores, con tumor renal avanzado (gran tamaño) y con tumor renal inicial (pequeño tamaño) (Tabla 1).

Los criterios de selección por los cuales se eligieron a los pacientes a analizar incluían: 1) haber sido intervenido quirúrgicamente de más de un tumor sólido procedente del mismo riñón durante la misma operación, 2) que al menos uno de los tumores sólidos fuera mayor de 3 cm de largo y al menos uno menor de 2 cm y 3) que ambos tumores sólidos, el grande y el pequeño, debieran tener un grado nuclear de Fuhrman (sistema de cuatro niveles, basado en la morfología del núcleo y en la presencia del nucléolo cuya utilidad es predecir la supervivencia específica al cáncer independientemente del estadio patológico) idénticos.

Tabla 1. Tamaños de los tumores de los 11 pacientes seleccionados para el estudio [69].

|            | Size  | Tumor size (cm) | Size Difference (cm) | Tumor volume (cc <sup>3</sup> ) | Volume Difference (cc <sup>3</sup> ) |
|------------|-------|-----------------|----------------------|---------------------------------|--------------------------------------|
| Patient 1  | large | 6               | 4.7                  | 84.8                            | 84.0                                 |
|            | small | 1.3             |                      | 0.8                             |                                      |
| Patient 2  | large | 5.5             | 3.7                  | 25.9                            | 24.5                                 |
|            | small | 1.8             |                      | 1.4                             |                                      |
| Patient 3  | large | 5.2             | 4                    | 34.3                            | 31.9                                 |
|            | small | 1.2             |                      | 2.4                             |                                      |
| Patient 4  | large | 4.5             | 2.5                  | 24.7                            | 22.0                                 |
|            | small | 2               |                      | 2.7                             |                                      |
| Patient 5  | large | 3.5             | 1.7                  | 11.9                            | 11.4                                 |
|            | small | 1.8             |                      | 0.5                             |                                      |
| Patient 6  | large | 3.5             | 2                    | 9.2                             | 8.9                                  |
|            | small | 1.5             |                      | 0.3                             |                                      |
| Patient 7  | large | 3.4             | 2.1                  | 15.9                            | 15.3                                 |
|            | small | 1.3             |                      | 0.6                             |                                      |
| Patient 8  | large | 3.1             | 1.1                  | 11.8                            | 10.3                                 |
|            | small | 2               |                      | 1.5                             |                                      |
| Patient 9  | large | 3               | 1                    | 4.7                             | 2.3                                  |
|            | small | 2               |                      | 2.4                             |                                      |
| Patient 10 | large | 3               | 1.6                  | 7.9                             | 6.8                                  |
|            | small | 1.4             |                      | 1.1                             |                                      |
| Patient 11 | large | 3               | 1.8                  | 14.1                            | 13.8                                 |
|            | small | 1.2             |                      | 0.3                             |                                      |

Para el análisis mediante “microarray”, el procedimiento de extracción de ARN, aislamiento, purificación fue similar al indicado en el material y métodos y todos los chips fueron preparados de acuerdo con las instrucciones, utilizándose para medir la expresión del ARNm el GeneChip ® Hg-U133 plus 2.0. y el escáner Affymetrix GeneChip 3000 para la obtención de la imagen.

Los archivos con los datos brutos (.CEL) fueron importados a los instrumentos BRB. Los datos fueron normalizados usando el algoritmo RMA. Se identificaron los genes que se expresaban diferencialmente y se agruparon en “clústeres” usando una correlación jerárquica basada en la diferencia en tamaño entre los tumores grandes y pequeños.

Los datos fueron separados en dos grupos para el análisis, un primer grupo de paciente con un par de tumores ambos T1a, es decir, con un tamaño menor a 4 cm y un segundo grupo que incluía pacientes con un tumor T1a (menor de 4 cm) y uno T1b (entre 4 y 7 cm).

Los resultados de la organización en *cluster* de los genes basada en la diferencia de volumen del par de tumores se puede observar en la figura:

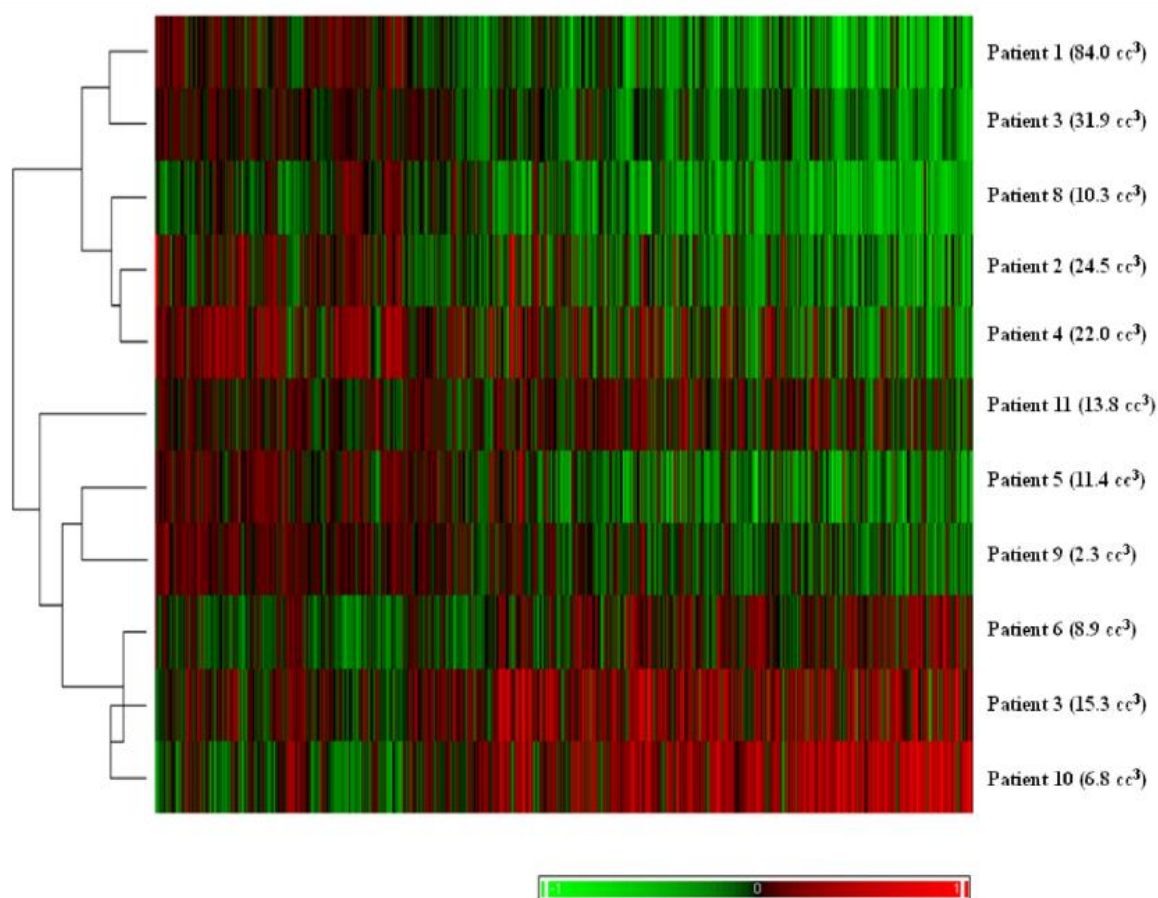


Figura 9. Cluster jerárquico del par de tumores de cada paciente basados en la diferencia de su tamaño. La diferencia de volumen entre el tumor grande y el pequeño se muestra entre paréntesis en cada paciente [69]

Como muestra la Figura 9, sin ninguna excepción, los pares de tumores con grandes diferencias en volumen están separados en el cluster de los pares con pequeñas diferencias, es decir, los pacientes con pares de tumores con mayores diferencias de volumen entre sí se encuentra localizados en el cluster próximos entre sí (pacientes 1, 3,2 y 4) y separados de los pacientes con pares de tumores con menor diferencia de volumen, localizados a su vez también próximos en el cluster. (pacientes 5, 9,6 y 10).

En cuanto al análisis de los microchips Affymetrix HGU133 plus 2.0, con 54.675 juegos de sondas, se utilizó un umbral de significancia de 0,01 esperando aproximadamente 550 falsos positivos. En los pares analizados cuyos ambos tumores tenían menos de 4 cm, hubo 458 genes que se expresaron

diferencialmente, menos de lo que se esperaba, sin embargo, cuando los tumores mayores de 4 cm fueron comparados con los tumores de menos de 4 cm en el mismo paciente, hubo 890 genes identificados que se expresaron diferencialmente, lo que es mayor que los falsos positivos esperados.

Cuando se examinó la ontología génica de los cambios génicos significativos, se encontró que estaban sobreexpresados los transcritos implicados en la producción de citoquinas y en la síntesis de ADN.

También fueron capaces de investigar la naturaleza de los cambios genéticos a nivel de cromosoma entre el gran y el pequeño tumor. Se utilizaron CGMA (Comparative Genomic Microarray Analysis) para examinar los cambios en la expresión génica entre las muestras de tumor y un pool de parénquima renal normal. En todos los pacientes que confirmaron una mutación VHL (von Hippel-Lindau), el nivel de expresión disminuyó en el cromosoma 3p25 en todas las muestras cuando se comparó con parénquima renal normal. Fue de particular interés el encontrar que también hubo una disminución en la expresión génica cerca del cromosoma 3p21. Todos los análisis cromosómicos revelaron que los perfiles de expresión dentro de cada par eran aparentemente similares. No se observaron diferencias constantes en la expresión sobre el nivel cromosómico entre los tumores grandes y los pequeños.

Por lo tanto, mediante el análisis de los perfiles de expresión genética entre grandes y pequeños tumores en el mismo paciente, se encontró que diferencias en volumen en tumores grandes incrementan progresivamente los cambios globales del transcriptoma. Además, aunque fueron apreciados algunos cambios en la expresión génica a nivel cromosómico, a nivel de transcripciones individuales, hubo un número creciente de cambios entre los tumores grandes y pequeños, especialmente cuando el tumor más grande era mayor de 4cm.

## 4.2 Aplicación de los “microarrays” para el estudio de enfermedades infecciosas y generación de vacunas

La tecnología de los “microarrays” ha abierto un nuevo mundo a los investigadores, el estudio de enfermedades infecciosas ya no se limitan a la observación de un único aspecto de la relación hospedador- patógeno; en su lugar se puede explorar una vista de todo el genoma de esta compleja interacción. Este amplio método de análisis está ayudando a los científicos a descubrir y comprender la enfermedad, las vías y en última instancia, desarrollar mejores métodos de detección, tratamiento y prevención [70].

La flexibilidad del análisis, en este caso en GeneChip Affymetrix permite en un solo “array” y un solo experimento abarcar diferentes tipos de estudios de enfermedades infecciosas.

Son múltiples los aspectos en la investigación de las enfermedades infecciosas en los que estos “microarray” han sido usados, entre los más importantes destacan:

- **Patogénesis.** En la identificación de factores virulentos y sus mecanismos de evasión del sistema inmune [71].

Los “microarrays” han permitido a los microbiólogos explorar la expresión de todo el genoma, descubriendo las vías de virulencia compuestas por muchos genes que eran anteriormente desconocidas.

- **Susceptibilidad.** En la determinación de la susceptibilidad del hospedador al patógeno o susceptibilidad del hospedador a una enfermedad.

¿Por qué una persona es susceptible a una infección mientras que otra no?

Seguramente, la exposición al medio ambiente juega un papel crucial, pero la genética del individuo contribuye también. Usando “microarrays” para el análisis de secuencias de ADN, se podría rápidamente identificar mutaciones en “genes de susceptibilidad” que influyen en el riesgo de un individuo a adquirir la enfermedad [72].

- **Respuesta a fármacos.** En la determinación de la respuesta a fármacos tanto del hospedador como del patógeno.

La inhibición de un determinado proceso celular por un fármaco puede dar como resultado la activación de mecanismos reguladores en la célula, lo que conlleva cambios en el perfil de expresión celular. El estudio de estos cambios en el perfil de expresión mediante “microarrays” puede revelar información sobre el modo de acción de fármacos, inhibidores o compuestos tóxicos. Los “microarrays” se han usado, por ejemplo, en un estudio de Wilson *et al* (1999) [73] para explorar los cambios en la expresión génica de *Mycobacterium tuberculosis* inducidos por el fármaco antituberculoso isoniazida. Este experimento mostró que la isoniazida inducía la expresión de varios genes que codificaban proteínas fisiológicamente relevantes para el modo de acción del fármaco. Con este tipo de estudios se pueden definir nuevas diana terapéuticas y así sintetizar nuevos compuestos que actúen sobre ellas.

- **Descubrimiento de vacunas.** En el descubrimiento de vacunas mediante el descubrimiento de antígenos.

Usando “arrays” GeneChip, los investigadores pueden examinar la actividad transcripcional de todos los genes de un microorganismo patógeno en condiciones “in vivo”, permitiendo la identificación de genes importantes aunque raramente expresados. Los “microarrays” no sólo proporcionan a los científicos una lista cuantitativamente larga de candidatos potenciales a vacunas, sino que los perfiles de expresión también ofrecen un tipo de información que permite identificar más rápidamente candidatos a vacunas más exitosas [74].

En un estudio llevado a cabo en el año 2002 por Kurz *y col.* [75] se analizó la expresión del genoma de *Neisseria meningitidis* para encontrar proteínas de superficie que fueran inducidas en condiciones in vivo. Se analizó la expresión génica durante diferentes etapas de infección de *N. meningitidis* del serogrupo B, uno de los serogrupos responsable de la mayoría de las enfermedades producidas por bacterias meningocócicas en países industrializados. Mediante la utilización de

“microarrays” fueron capaces de identificar grupos específicos de determinantes antigénicos y posibles candidatos para vacunas contra la meningitis bacteriana basada en esas proteínas de superficie.

La cepa meningocócica bacteriana usada en este estudio fue la *Neisseria meningitidis* serogrupo B cepa MC58 y el análisis de su expresión génica se llevó a cabo siguiendo el protocolo habitual utilizado en “microarrays”. Anterior al aislamiento de ARN, bacterias MC58 WT (“wild type”) fueron cultivadas en medio suplementado con proteasa peptona (PPM+) hasta la fase de crecimiento logarítmica, lavadas una vez en PBS, y separadas en dos grupos; una mitad fue incubada en suero humano mientras que la otra mitad fue incubada en PBS, actuando esta como grupo control. Posteriormente se procedió al aislamiento del ARN, asegurando su calidad y eliminación de todo posible resto de ADN humano.

Para la construcción del “microarray” de oligonucleótidos de ADN basado en el genoma completo de *N. meningitidis* serogrupo B cepa MC58, se utilizaron como sondas cada uno de los 2158 ORFs (“Open Reading Frame”) presentes en su genoma. Se procedió a la síntesis de ADNc, a su posterior marcaje, utilizando para ello cantidades iguales de ARN de ambos grupos (grupo de ARN aislado de las bacterias tratadas con suero humano y grupo de bacterias control tratadas con PBS), que fueron marcados diferencialmente con los flurocromos Cy3- dCTP y Cy5- dCTP, respectivamente, durante la primera reacción de reverso transcripción.

Los dos ADNc diferencialmente marcados fueron combinados e hibridados en el “microarrays” de genoma completo durante 16 horas a una temperatura de 50 °C, lavados y se procedió a su escaneado mediante la utilización del escanner Affymetrix 418™. El promedio de intensidad de la señal y las medidas de fondo focal fueron obtenidos para cada uno de los puntos usando el software ImaGene 4.0, los dos canales de color fueron normalizados y los experimentos biológicos y las hibridaciones fueron repetidas al menos 3 veces y combinados los datos de los experimentos independientes.

El análisis de los datos obtenidos después de la hibridación condujo a la identificación de 279 genes regulados diferencialmente. De estos:

- 134 genes fueron inducidos
- 145 genes fueron reprimidos

El número de genes diferencialmente regulados con un 12.9% del genoma del meningococo (Tabla 2). Los ORFs pudieron ser agrupados en las siguientes categorías funcionales: metabolismo, transcripción, traducción, síntesis de ADN y modificación, proteínas de membrana y transportadores, así como una amplia gama de ORF hipotéticos. 34 de los ORFs diferencialmente transcritos fueron previamente demostrados que estaban asociados a la virulencia o que eran esenciales en condiciones in vivo, sugiriendo que son necesarios para la supervivencia del meningococo y su difusión en el torrente sanguíneo.

Tabla 2. Genes (ORFs) de *N. meningitidis* serogrupo B diferencialmente regulados en suero humano tratado con este meningococo [75].

| Categoría Funcional      | Genes regulados Diferencialmente | Mínimo y Máximo | Inducidos | Reprimidos | Genes de virulencia |
|--------------------------|----------------------------------|-----------------|-----------|------------|---------------------|
| Cápsula                  | 3                                | -1.9 a +2.3     | 1         | 2          | 3                   |
| Ciclo celular            | 2                                | -2.9 a +4.2     | 1         | 1          | 1                   |
| Síntesis de ADN          | 12                               | -2.4 a +2.8     | 5         | 7          | 1                   |
| Membrana                 | 35                               | -2.4 a +3.6     | 14        | 21         | 16                  |
| Transporte               | 20                               | -3.9 a +3.1     | 9         | 11         | 4                   |
| Metabolismo              | 63                               | -5.2 a +5.7     | 14        | 49         | 8                   |
| Relacionados con el fago | 1                                | -1.9            | 0         | 1          | 0                   |
| Destino de la proteína   | 5                                | -3.1 a +2.4     | 2         | 3          | 0                   |
| Transcripción            | 10                               | -3.0 a +2.6     | 3         | 7          | 1                   |
| Traducción               | 40                               | -2.8 a +7.6     | 27        | 13         | 0                   |
| Hipotéticos              | 90                               | -3.6 a +11.9    | 60        | 30         | 0                   |
| Total                    | 279                              | -5.2 a +11.9    | 134       | 145        | 34                  |

Por lo tanto, un candidato ideal para la vacuna debería estar expuesto en la superficie de bacterias intactas con el fin de ser altamente accesible para el sistema inmune. Es probable que las proteínas de membrana se evalúen en un futuro como posibles antígenos en el desarrollo de nuevas vacunas contra *N. meningitidis*.

- **Identificación de patógenos.** En el análisis de la secuencia del patógeno.

Quizás la aplicación más directa de la tecnología de los “microarrays” para las enfermedades infecciosas es la habilidad de identificar rápidamente un patógeno

infeccioso. Debido a que los patógenos tienen distintas composiciones genéticas y los “microarrays” son capaces de examinar todas las secuencias génicas, los “arrays” son una herramienta ideal para esta aplicación [76].

#### **4.3 Aplicación de los “microarrays” en estudios llevados a cabo en *Arabidopsis thaliana* y plantas**

En un estudio realizado en 2008 por Matsui y col. [77] se aplicó la técnica del “tiling array” de Affymetrix en *Arabidopsis thaliana* para estudiar el transcriptoma completo bajo condiciones de sequía, frío, alta salinidad y tratamiento con ácido abscísico (ABA). Las plantas responden y se adaptan a la sequía, frío y estrés por alta salinidad para sobrevivir, pero el estrés producido en las plantas en estas condiciones induce varias respuestas bioquímicas y fisiológicas en las plantas, siendo varios cientos de genes identificados como los genes que responden a este estrés a nivel transcripcional.

En este estudio se describe el perfil de expresión genómico de *Arabidopsis thaliana* ecotipo Columbia usando “tiling array” en respuesta a la sequía, frío, alta salinidad y ABA.

Se procedió a la síntesis de sondas para el análisis con “tiling array” del genoma completo y a su hibridación en micromatrices, tinción y lavado siguiendo el protocolo de muestras eucariotas. El escaneado fue realizado a 0.7  $\mu\text{m}$  de resolución usando el escáner GeneChip 30000 7G. Además, para cada condición medida se tomaron 3 réplicas independientes para cada “array”, siendo identificados, tras el proceso de hibridación en el correspondiente microarray, un total de 16.876, 17.920, 17.458, 17.388 y 18.126 genes como los genes expresados en condiciones de sequía, frío, alta salinidad y plantas tratadas con ABA y no tratadas, respectivamente, en dos medidas realizadas, en un tiempo de 2 horas y 10 horas (Tabla 3).

Tabla 3. Número de genes expresados en condiciones de sequía, frío, alta-salinidad, tratamiento con ABA y plantas no tratadas medidos mediante “tiling array” [77].

| ELEMENTO       | Tiempo | Nº de genes up-regulated | Nº de genes down-regulated | Nº de genes totales expresados |
|----------------|--------|--------------------------|----------------------------|--------------------------------|
| SEQUÍA         | 2 h    | 1.188                    | 217                        | 16.744                         |
|                | 10 h   | 2.059                    | 2.075                      | 15.815                         |
| FRÍO           | 2 h    | 427                      | 14                         | 17.270                         |
|                | 10 h   | 740                      | 396                        | 16.349                         |
| ALTA-SALINIDAD | 2 h    | 1.471                    | 495                        | 16.697                         |
|                | 10 h   | 2.693                    | 2.038                      | 15.920                         |
| ABA            | 2 h    | 1678                     | 1.581                      | 16.241                         |
|                | 10 h   | 3.179                    | 2.220                      | 16.181                         |
| NO TRATADAS    |        |                          |                            | 18.126                         |
| TOTAL          |        | 5.303                    | 3.974                      | 20.264                         |

Del total de genes expresados, 2.421, 996, 2.903 y 3.623 genes se identificaron como sobre-expresados por los tratamientos, respectivamente.

Estos genes incluían algunos de los ya registrados cómo que se veían alterados en sequía, frío o alta salinidad, o que se sobre-expresaban en tratamientos con ABA, en concreto era el caso de los genes RD29 A/COR78/LT178, RD29 B, RD20 DREB2A Y AtMYC2. Estos resultados indican que el análisis llevado a cabo de “tiling array” funcionó correctamente para encontrar genes sobre-expresados que están implicados en estas cuatro condiciones.

#### 4.4 Aplicación de los “microarrays” en estudios de *Drosophila melanogaster*

*Drosophila melanogaster*, al igual que otras especies, es una especie que vive en sociedades de diferentes niveles. Las interacciones sociales de unos individuos con otros son conocidas por alterar la psicología individual [78], pero la medida en que los mecanismos transcripcionales y postranscripcionales están implicados en la inducción de estos cambios es incierta.

Como hipótesis se propone que las interacciones sociales, tales como aquellas involucradas en comportamientos reproductivos, conducen a cambios inmediatos y analizables en la expresión génica. Tales cambios pueden tener efectos importantes

en el éxito reproductivo individual y aptitud a través de alteraciones en la fisiología o a través de cambios a corto o largo plazo en función del sistema nervioso.

En el trabajo llevado a cabo en el año 2007 por Carney y col. [86] se analizaron mediante la tecnología de “microarray” los efectos que tenían las interacciones sociales en la expresión génica, como la exposición con hembras y el cortejo, en *Drosophila melanogaster* macho. Para llevar a cabo el análisis se procedió a la recolección de moscas vírgenes machos de tipo salvaje Canton-S y hembras y a su clasificación por edades. Los machos fueron aislados individualmente en frascos y en un periodo de 4 días se les introdujo a cada uno una hembra de esta especie. Se observó el comportamiento de los machos ante la introducción de la hembra durante 5 minutos. Esta observación permitió establecer una clasificación, en la cual se observó que; 1) el 35% de los machos mostraron un gran cortejo pero sin apareamiento, 2) un 36% se acopló con la hembra dentro del periodo de observación y 3) el 29% restante realizó un leve comportamiento de cortejo pero sin tratar de copular. De todos ellos, el 35%, los machos que demostraron un gran cortejo hacía la hembra (extensión de alas y pretensión de cópula) fueron recolectados y almacenados para la extracción de ARN. Dentro de este grupo se realizó un agrupamiento al azar de aproximadamente 12 individuos y el ARN total fue extraído con TRIzol siguiendo el protocolo estándar. Las muestras control consistieron en machos que fueron tratados idénticamente salvo que no estuvieron expuestos a un objeto de cortejo. Todas las colecciones y los ensayos se realizaron en el mismo tiempo cada día para evitar efectos circadianos en la expresión génica.

Tres muestras de ARN de machos que cortejaron a hembras y tres muestras de ARN de machos que no fueron expuestos a hembras fueron usadas como sondas en el “Array Drosophila Genome” (Affymetrix) para un total de 6 “array”. Este “array” usado en el experimento contenía un grupo de sondas para más de 13.500 transcritos basados en el genoma de *Drosophila*. Se realizó el marcaje con fluorocromos y la hibridación usando los protocolos estándar de Affymetrix. Para validar los resultados de los “microarrays”, se utilizó el kit “Superscript 1st Strand Synthesis” (Invitrogen) para preparar el ADNc de muestras de ARN obtenidas independientemente (3 muestras experimentales y 3 control) así como de dos conjuntos de muestras utilizado para el análisis de “microarrays” y se procedió a la validación por PCR en tiempo real.

Los resultados finales obtenidos mostraron que 43 genes presentaban niveles de expresión alterados debido a las interacciones del cortejo: 10 genes aparecieron sobre-expresados y 33 genes por debajo de su expresión (Tablas 4 y 5).

Tabla 4. Genes sobre-expresados en machos que cortejaron a las hembras [78].

| Gene identifier | Gene name   | Avg. fold change | GCOS p-value | PM+MM p-value | PM only p-value |
|-----------------|-------------|------------------|--------------|---------------|-----------------|
| CG6128          |             | 2.58             | 3.48E-04     | 1.27E-04      | 1.93E-05        |
| CG13155         |             | 2.26             |              | 1.46E-04      | 6.27E-05        |
| CG7738          |             | 2                | 7.57E-04     | 1.09E-06      | 3.22E-06        |
| CG1732          |             | 1.89             | 7.29E-04     | 3.57E-05      | 1.82E-04        |
| CG33060         |             | 1.85             | 1.41E-04     | 1.35E-05      | 4.65E-05        |
| CG30042         |             | 1.8              | 9.40E-04     | 1.63E-05      | 5.21E-04        |
| CG2958          | lectin-24Db | 1.71             |              | 5.29E-04      | 6.01E-04        |
| CG17820         | fit         | 1.6              | 5.70E-04     | 7.20E-05      | 1.10E-04        |
| CG7592          | Obp99b      | 1.52             | 2.44E-04     |               | 7.17E-04        |
| CG7106          | lectin-28C  | 1.48             | 6.77E-04     | 3.46E-04      | 2.11E-04        |

Tabla 5. Genes por debajo de su expresión en machos que cortejaron a las hembras [78].

| Gene identifier | Gene name | Avg. fold change | GCOS p-value | PM+MM p-value | PM only p-value |
|-----------------|-----------|------------------|--------------|---------------|-----------------|
| CG6639          |           | -21.2            | 1.12E-10     | 3.17E-10      | 1.24E-11        |
| CG4740          | AttC      | -11.79           | 3.00E-07     | 2.82E-05      | 4.02E-04        |
| CG18372         | AttB      | -9.62            |              | 3.19E-04      | 4.94E-04        |
| CG10146         | AttA      | -8.51            | 7.23E-04     | 3.80E-04      | 6.72E-04        |
| CG4757          |           | -7.69            | 1.26E-07     | 1.17E-06      | 1.65E-06        |
| CG15066         | IM23      | -5.17            | 1.11E-08     | 1.13E-10      | 1.61E-10        |
| CG8175          | Mtk       | -4.37            | 4.01E-06     | 1.88E-05      | 4.26E-05        |
| CG1367          | CecA2     | -4.31            | 8.34E-04     | 1.06E-05      | 9.88E-04        |
| CG13422         |           | -4               | 3.16E-08     | 2.34E-07      | 3.71E-07        |
| CG10810         | Drs       | -3.99            | 8.60E-07     | 7.77E-09      | 6.18E-09        |
| CG10816         | Dro       | -3.3             | 2.73E-05     | 1.50E-08      | 3.13E-05        |
| CG18563         |           | -3.28            | 8.55E-06     | 1.57E-04      | 8.55E-06        |
| CG1365          | CecA1     | -3.12            | 8.19E-04     | 4.65E-04      |                 |
| CG6687          |           | -2.84            | 2.39E-04     | 4.02E-05      | 6.84E-05        |
| CG18108         | IMI       | -2.46            | 8.67E-08     | 3.41E-07      | 1.58E-07        |
| CG9989          |           | -2.18            |              | 1.00E-04      | 7.14E-04        |
| CG5550          |           | -1.93            | 1.12E-04     | 1.26E-04      | 2.18E-05        |
| CG2217          |           | -1.92            |              | 1.87E-05      | 2.42E-05        |
| CG15065         |           | -1.9             | 4.43E-05     | 9.75E-08      | 7.54E-08        |
| CG14745         | PGRP-SC2  | -1.89            |              | 4.30E-06      | 6.96E-05        |
| CG9434          | Fst       | -1.76            |              | 3.46E-04      | 1.29E-04        |
| CG2042          |           | -1.67            |              | 1.40E-04      | 8.27E-05        |
| CG16772         |           | -1.66            | 5.99E-04     | 3.71E-05      | 2.53E-05        |
| CG13947         |           | -1.57            | 2.79E-04     | 9.57E-05      | 7.27E-05        |
| CG18279         | IMI0      | -1.55            | 1.00E-03     | 2.06E-05      | 3.48E-05        |
| CG13482         |           | -1.48            |              | 3.75E-04      | 2.94E-04        |
| CG18106         | IM2       | -1.42            | 5.99E-04     |               | 9.62E-04        |
| CG11992         | Rel       | -1.42            |              | 3.12E-04      | 4.54E-04        |
| CG16844         | IM3       | -1.4             |              | 3.07E-06      | 7.87E-07        |
| CG15231         | IM4       | -1.28            |              | 9.52E-05      | 4.92E-06        |
| CG18067         |           | -1.27            |              | 1.89E-05      | 9.24E-06        |
| CG10332/IMI8    |           | -1.05            |              | 3.66E-04      | 3.51E-04        |
| CG10947         |           | -1.05            |              | 2.74E-06      | 4.20E-05        |

Estos descubrimientos indican que las interacciones sociales a corto plazo entre las moscas pueden conducir a respuestas genéticas rápidas. Este estudio, junto con otro más reciente [79] propone la posibilidad de cambios rápidos en la abundancia de ARNm, dentro incluso de un intervalo de solo 5 minutos

Además, la mayoría de los genes identificados corresponden a genes relacionados con el comportamiento reproductivo masculino y la producción de espermatozoides ante un inminente apareamiento, genes que se expresan en testículos y genes que aumentan su expresión génica en la grasa corporal de la cabeza, pudiendo tener estos importantes funciones en comportamientos reproductivos específicos del sexo y en la psicología de la reproducción [80].

#### **4.5 Aplicación de los “microarrays” en estudios realizados en *Caenorhabditis elegans***

La utilización de *Caenorhabditis elegans* como organismo modelo se debe a su gran homología genómica con el sistema de los mamíferos, y aunque es un organismo menos complejo que estos, proporciona un excelente modelo para dilucidar los mecanismos que se producen por la intoxicación por metales pesados. Por lo tanto, el sistema de modelo de *C. elegans* es válido para la investigación de la toxicidad por metales, como el arsénico, y puede ser particularmente útil para examinar las interacciones gen-ambiente.

El metal arsénico, un conocido carcinógeno en humanos, está altamente distribuido alrededor del mundo y el envenenamiento crónico por este que afecta a millones de personas en el mundo está asociado con un incremento en el riesgo de muchas enfermedades incluyendo arterioesclerosis, diabetes y cáncer [81].

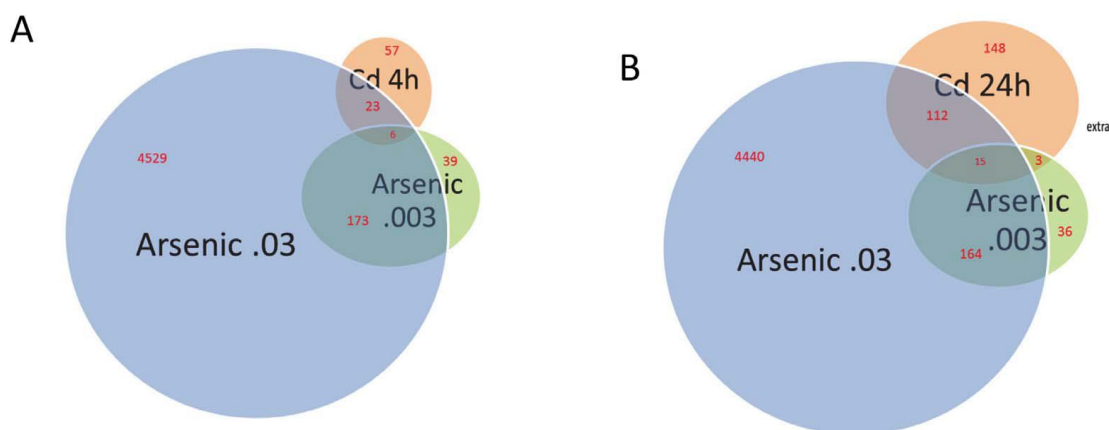
En un estudio realizado en 2013 Sahu y col. [82] se exploró la respuesta a nivel de genoma global a altos y bajos niveles de exposición a arsénico en *C. elegans* usando “microarrays” de expresión Affymetrix.

Para estudiar el patrón de expresión génica global después de la exposición a arsénico, se realizó un estudio donde *C. elegans* de tipo salvaje de la cepa Bristol N2 en la etapa de crecimiento L4 fueron expuestos a arsenito de sodio en dos concentraciones diferentes, una de 0.03% w/v y otra de 0.003% w/v, considerando esta primera como una dosis alta y la segunda como dosis baja.

Después del tratamiento con arsénico los animales fueron recolectados y lavados en un buffer M9, el ARN fue extraído usando el reactivo TRIzol y el tratamiento de purificación de este ARN se realizó siguiendo el protocolo. Se realizaron 3 aislamientos de ARN independientes con cada condición para el análisis de “microarrays”. Fue sintetizado el ADNc y las muestras fueron hibridadas en el GeneChip *C.elegans* de Affymetrix. Este “chip” comprende 22.500 sondas del genoma de *C.elegans*. Los datos fueron procesados usando el software Partek Genomics Suite, version 6.5., normalizados y corregidos según las instrucciones de Affymetrix. Se utilizó ANOVA para determinar qué genes se expresaban diferencialmente entre los grupos con tratamiento experimental y los controles.

Los resultados del “microarray” revelaron que la exposición a altas dosis (0.03% w/v) causó grandes cambios globales de expresión en comparación con la exposición a bajas dosis (0.003%). De los genes que se expresaron diferencialmente, 204 genes se sobre-expresaban 4 veces por encima en la exposición a altas dosis de arsénico, y solo 49 genes mostraban esta sobre-expresión en condiciones de baja exposición. 46 de estos genes eran comunes entre las dos listas y 43 de estos mostraron relación dosis-respuesta donde los altos niveles de arsénito de sodio condujeron a mayores niveles de expresión génica en *C.elegans* (Figura 4).

Las principales funciones en las que se encuentran implicados estos genes comunes expresados diferencialmente son: unión al ion zinc, actividad transportadora, transporte y plegamiento de proteínas, actividad hidrolasa, unión a iones metálicos, etc. (Figura 10.C).



C

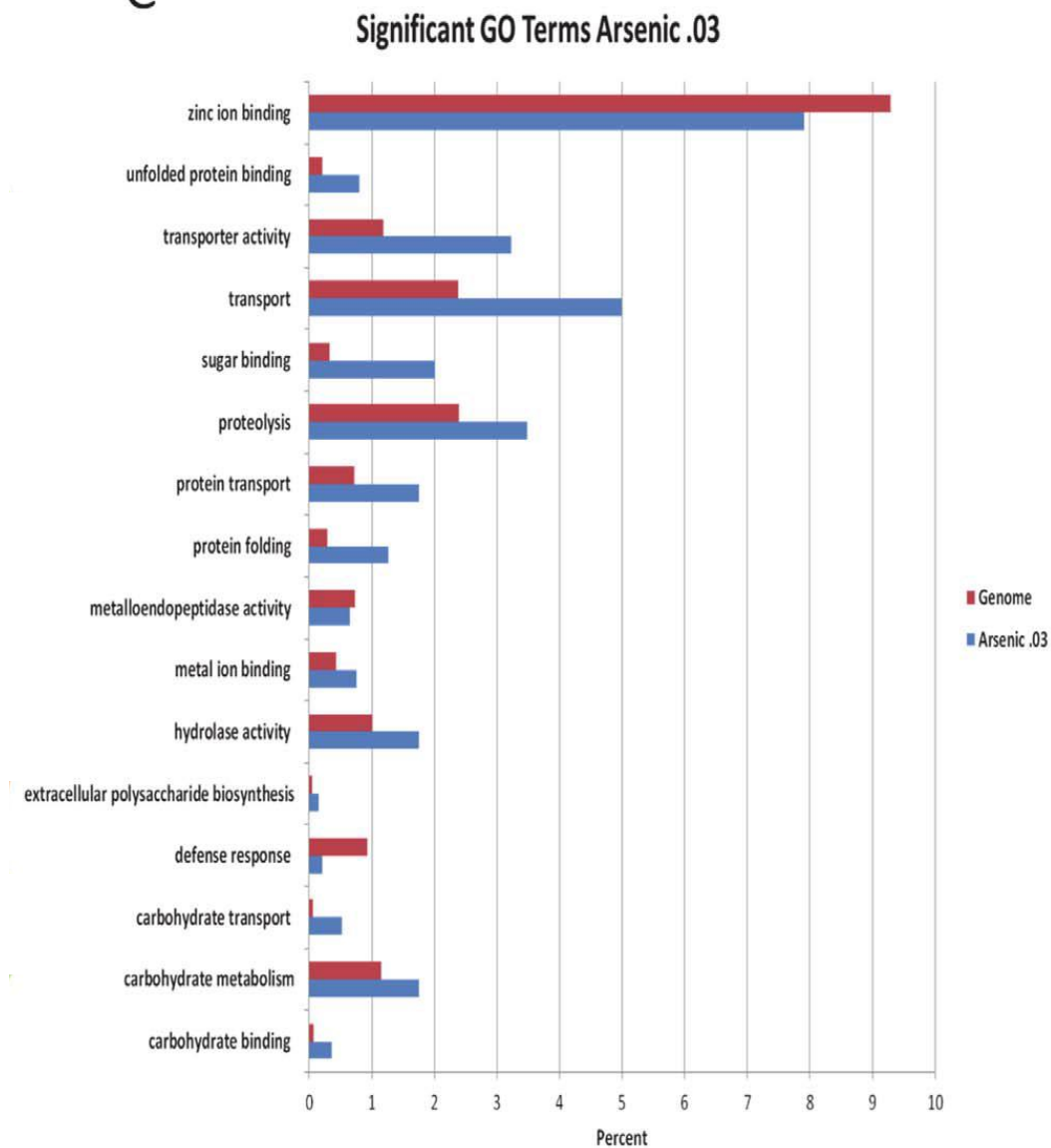


Figura 10. Perfil de expresion del genoma completo de *C. elegans* de los genes regulados por arsénico y cadmio. A: Diagrama de Venn ilustrando el número de genes expresados en exposiciones a altas y bajas dosis de arsénico y 4 horas de cadmio. B: Diagrama de Venn ilustrando el número de genes expresados en exposiciones a altas y bajas dosis de arsénico y 24 horas de cadmio. C: Enriquecimiento de categorías de ontología génica de genes expresados diferencialmente en la exposición de arsénico de dosis alta [82].

## 5. CONCLUSIONES

Los “microarrays” son una herramienta genómica muy poderosa, diseñada para dilucidar las diferencias en la expresión génica. A pesar de ser una tecnología relativamente nueva, la comunidad científica ha adoptado rápidamente su uso en una variedad de campos, incluyendo el desarrollo de fármacos, biología evolutiva y caracterización de enfermedades [1,83].

La fuerza de la tecnología se basa en varios factores, incluyendo 1) su facilidad de uso, 2) disponibilidad de plataformas y menor coste en comparación con otros métodos exploratorios como “Northern-blot” o ensayo de protección de ribonucleasa, 3) aplicación de métodos estadísticos para el análisis detallado y 4) lo más importante, una visión global de la expresión génica que abarque al genoma completo.

Las limitaciones tecnológicas asociadas con “microarrays” se manifiestan en términos de variabilidad que típicamente corresponden a errores sistemáticos. Mejoras en robótica, en su fabricación y en secuenciación del genoma continuo pueden abordar estas cuestiones, pero no eliminarlas. Estas cuestiones ponen límites a los logros que la tecnología puede conseguir, aunque una amplia comprensión acerca de los “microarrays” puede ayudar a establecer datos realmente significativos y reproducibles. Un esfuerzo para establecer un correcto diseño experimental, establecer medidas de control de calidad como la comprobación de la pureza del ARN, análisis de datos y verificación de los resultados puede combatir estos desafíos tecnológicos [8, 10, 84]. Además, la mejora y el incremento en las bases de datos y archivos son consideraciones que a menudo se pasan por alto aunque bastante importantes para ser capaces de dar con nuevos datos y para la investigación posterior o simplemente para comparar con nuevos datos obtenidos.

En cuanto a la complejidad del análisis de “microarrays”, resulta sumamente difícil determinar datos objetivos y con significación biológica sin definir claramente metas. Un aspecto muy relacionado del análisis genómico es la determinación de la interacción entre genes o grupos de genes. Son estos mecanismos que difícilmente se descifran utilizando “microarrays”. Y finalmente, la funcionalidad de un gen no puede determinarse únicamente utilizando “microarrays”, se necesitan otros métodos y herramientas experimentales para comprender las diferentes

interacciones entre genes o entre proteínas y desarrollar una visión más completa del comportamiento celular [2, 3,6].

En última instancia, en este trabajo se muestra la gran variedad de aplicaciones que tiene esta tecnología actualmente, en el futuro, los “microarrays” continuarán siendo utilizados en una variedad de áreas de investigación. Con seguridad, sus aplicaciones aumentarán cuando se disponga de más opciones en el diseño de matrices personalizadas junto con un aumento en el surtido de “microarrays” específicos. Los avances tecnológicos pueden ayudar a reducir el costo así como mejorar la reproductibilidad y fiabilidad promoviendo la aplicación de los “microarrays” en nuevos y diversos campos, no obstante, en la actualidad se ha convertido en una tecnología de vanguardia que está permitiendo determinar los niveles de expresión génica global.

## **6. BIBLIOGRAFÍA**

- [1] Brown PO, Botstein D (1999) Exploring the new world of the genoma with DNA “microarrays”. *Nat Genet.* **21**:33-37.
- [2] Nakanishi T, Oka T, Akagi T (2001) Recent Advances in DNA “microarrays”. *Acta Med Okayama.* **55**:319-328.
- [3] Li X, Gu W, Mohan S, Baylink DJ (2002) DNA “microarrays”: their use and misuse. *Microcirculation.* **9**:13-22.
- [4] Butte A (2002) The use and analysis of microarray data. *Nat Rev Drug Discov..* **1**:951-960.
- [5] Conway T, Schoolnik GK (2003) Microarray expression profiling: capturing a genome-wide portrait of the transcriptome. *Mol Microbiol.* **47**: 879-889.
- [6] Xiang Z, Yang Y, Ma X, Ding W (2003) Microarray expression profiling: analysis and applications. *Curr Opin Drug Discov Devel.* **6**:384-395.
- [7] Slonim DK (2002) From patterns to pathways: gene expression data analysis comes of age. *Nat Genet.* **32**:502-508
- [8] Quackenbush J (2001) Computational analysis of microarray data. *Nat Re Genet.* **2**:418-427.
- [9] Gadgil M, Lian W, Gadgil C, Kapur V, Hu WS (2005) An analysis of the use of genomic DNA as a universal reference in two channel DNA “microarrays”. *BMC Genomics.* **6**:66.

- [10] Simon R, Radmacher MD, Dobbin K (2006) Design of studies using DNA “microarrays”. *Genet Epidemiol.* **23**: 21-26.
- [11] Crowther DJ (2002) Applications of “microarrays” in the pharmaceutical industry. *Curr Opin Pharmacol.* **2**: 551-554.
- [12] Kerr MK, Churchill GA (2001) Statistical design and the analysis of gene expression microarray data. *Genet Res.* **77**:123-128.
- [13] Simon RM, Dobbin K (2003) Experimental Design of DNA Microarray Experiments. *Bio Techniques.* **34**: S16-S21
- [14] Chomczynski P (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques.* **15**: 532-537.
- [15] Chomczynski P, Sacchi N (1987) Single Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction. *Anal. Biochem.* **162**: 156-159.
- [16] RNeasy ® Mini Handbook (2010) Fourth Edition. Boletín técnico de la casa comercial que describe el método RNeasy para la purificación de ARN total.
- [17] RNAlater® Solutions for RNA Stabilization and Storage (2010) Boletín técnico de la casa comercial que describe el método RNAlater para la estabilización del ARN total.
- [18] Kirchner B, Paul V, Riedmaier I, Pfaffl MW (2014) mRNA and microRNA purity and integrity: the key to success in expression profiling. *Methods Mol Biol.* **1160**:43-53. doi: 10.1007/978-1-4939-0733-5\_5.
- [19] Schroeder A, Mueller O, Stocker S y col. (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology.* **7**:3 doi:10.1186/1471-2199-7-3.
- [20] Cox ML, Ph D (2006) Working with Fixatives, Fixation, and Tissue Processing to Improve Morphology and RNA Quality. *PGRD.*

- [21] 2100 Bioanalyzer ® Boletín técnico de la casa comercial que describe el método 2100 Bioanalyzer para la medida del número RIN.
- [22] Vasudevan DM, Sreekumari. (2011) Texto de Bioquímica para estudiantes de medicina. Ed. Cuéllar Ayala. Sexta Edición. **647**: 618-619.
- [23] Watson JD (2006) Biología molecular del gen. Ed. Médica Panamericana. Quinta Edición. **800**: 123-132
- [24] Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J (2000). A concise guide to cDNA Microarray Analysis. BioTechniques. **29**: 548-562
- [25] Forozan F, Karhu R, Kononen J, Kallioniemi A, Kallioniemi OP (1997) Genome screening by comparative genomic hybridization. Trends Genet. **13**:405-409
- [26] Larraga V (2010) Genómica funcional frente a las enfermedades parasitarias. Centro de Investigaciones Biológicas. CSIC. Instituto Roche.
- [27] Böttinger EP, Novetsky A, Zavadil J (2003) RNA Labeling and Hybridization of DNA “microarrays”. Methods in Molecular Medicine: **86**: 10.1385/1-59259-392-5:275.
- [28] Barrero PR. (2005). Aplicaciones de la técnica de *microarrays* en ciencias biomédicas: presente y futuro. Revista QuímicaViva. ISSN 1666-7948.
- [29] Rivas-López M.J., Sánchez-Santos J.M y De las Rivas J. (2007) Estructura y análisis de microarrays. Universidad de Salamanca
- [30] Ayala G (2014) Bioinformática estadística. Análisis de expresión de genes. Departamento de Estadística e Investigación Operativa. Universidad de Valencia.
- [31] Dalma-Weiszhausz DD, Warrington J, Tanimoto EY, Garrett Miyada C. (2006). The Affymetrix GeneChip ® Platform: An Overview. **410**: 3-28
- [32] Agilent (2008) Agilent's DNA Microarray Scanner with Sure Scan High-Resolution Technology
- [33] Busquets X and Agusti AGN (2001). Chip genético (AND array). Arch Bronconeumol. **37**:394-6
- [34] López-Campos GH, Spiteri I, García-Hernández O, Martín-Sánchez F (2007). Gestión, procesamiento y análisis de imágenes biomédicas. Ediciones de la Universidad de Castilla-La Mancha, Colección Ciencia y Técnica 52, Cuenca
- [35] Angulo J, Serra J. (2003) Automatic analysis of DNA microarray images using mathematical morphology. Bioinformatics. **19**: 553–562

- [36] Chen Y, Dougherty ER, Bittner ML (1997). Ratio based decisions and the quantitative analysis of cDNA microarray images. *J. of Biomedical Optics*. **2**: 364-374.
- [37] González J (2010) Análisis de datos de “microarrays”. Universidad Politécnica de Valencia. Ibime. Informática Biomédica.
- [38] Kelmansky M (2007) Aspectos Estadísticos de “microarrays”. Dpto. de Matemáticas. Instituto de Cálculo.
- [39] Moreno V, Solé X (2004) Uso de chips de ADN (“microarrays”) en medicina: fundamentos técnicos y procedimientos básicos para el análisis estadístico de resultados. Unidad de Bioestadística y Bioinformática. Servicio de Epidemiología y Registro del Cáncer. Instituto Catalán de Oncología. *Med Clin* **122**: 73-9
- [40] Do JH, Choi DK. (2006). Normalization of microarray data: single-labeled and dual-labeled arrays. *Mol Cells*. **22**: 254-261.
- [41] Quackenbush J. (2002) Microarray data normalization and transformation. *Nat Genet*. **32**: 496-501.
- [42] Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, *et al.* (2001) Missing value estimation methods for DNA microarrays. *Bioinformatics*;17:520-5
- [43] Tsai CA, Chen YJ, Chen JJ. Testing for differentially expressed genes with microarrays (2003) *Nucleic Acids Res*. 31:e52.
- [44] Good P. (2000) *Permutation tests*. 2nd ed. New York: Springer.
- [45] Westfal PHY. (1993) *Resampling-based multiple testing: examples and methods for p-value adjustment*. New York: John Wiley & Sons.
- [46] Gentleman RC, Carey VJ, Bates DJ, Bolstad BM, Dettling M, Dudoit S, *et al.* (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Bioconductor Project Working Papers*. Working Paper 1.
- [47] Lucas A, Gautier L. (2006) Cluster and Tree Conversion. Bioconductor. Available from: URL: <http://bioconductor.org/packages/2.2/bioc/html/ctc.html>
- [48] Pollard KS, Ge Y, Taylor S, Dudoit S. (2005) Resampling-based multiple hypothesis testing. Bioconductor’s multtest Package. Available from: <http://bioconductor.org/packages/2.2/bioc/html/multtest.html>
- [49] Wu H, Yang H, Churchill GA. (2008) R/MAANOVA: An extensive R environment for the Analysis of Microarray Experiments. Available from: <http://bioconductor.org/packages/2.2/bioc/html/maanova.html>

- [50] Yang YH (2008). Exploratory analysis for two-color spotted microarray data. Available from: <http://bioconductor.org/packages/2.2/bioc/html/marray.html>
- [51] Tibshirani R, Chu G, Hastie T. (2005) The samr Package. Available from: <http://www-stat.stanford.edu/~tibs/SAM/Rdist/index.html>
- [52] Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, *et al.* (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*; **34** (2): 374-8.
- [53] Soukas A, Cohen P, Socci ND, Friedman JM. (2000) Leptin-specific patterns of gene expression in White adipose tissue. *Genes Dev*; **14**:963-80.
- [54] Eisen MB, Spellman PT, Brown PO, Botslein D (1998). Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci, Usa*; **95**:14863-8
- [55] Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, *et al.* (1999). Interpreting patterns of gene expression with self-organizing maps. *Proc Natl Acad Sci, USA*; **96**:2907-12.
- [56] Dopazo J, Carazo JM (1997). Phylogenetic reconstruction using an unsupervised growing neural network that adopts the topology of a phylogenetic tree. *J Mol Evol*; **44**:226-33
- [57] Zar JH (1999). *Biostatistical analysis*. 4th ed., New Jersey: Prentice Hall; p.663
- [58] Lander E (2002) The importance of RNA quality in Real Time Quantitative RT-PCR\_- Eric Lander, Ambion, ABRF .
- [59] Soinov LA, Krestyaninova MA, Brazma A (2003). Towards reconstruction of gene networks from expression data by supervised learning. *Genome Biol*; **4**(1):R6
- [60] Eisen MB, Spellman PT, Brown PO, Botstein D. (1998). Cluster analysis and display of genome-wide expression patterns. **95**: 14863-14868.
- [61] Miranda J, Bringas R (2008) Análisis de datos de microarrays de ADN. Parte I: Antecedentes de la tecnología y diseño experimental. Centro de Ingeniería Genética y Biotecnología, CIGB, **25**:82-89.
- [62] Hastie TT, Friedman J. (2001) *The elements of statistical learning: data mining, inference, and prediction*. New York: Springer.
- [63] Hoopes, L. (2008) Genetic diagnosis: DNA microarrays and cancer. *Nature Education* **1**:3
- [64] Vallin C (2007) Microarrays de ADN y sus aplicaciones en investigaciones biomédicas. *Revista CENIC*. 0253-5688. ISSN electrónico: 2221-2450

- [65] Bucc G., Carruba G., Saetta A., Muti P., Castagnetta L. and Smith C.P (2004) Gene Expression Profiling of Human Cancers. Ann. N.Y. Acad. Sci., 1028, 1-10, 2004. New York Academy of Science
- [66] Ramaswamy S. and Golub T.R. (2002) DNA microarrays in clinical oncology. J. Clin. Oncol., 20, 1932-41
- [67] Xu J, Stolk J.A, Zhang X, Silva S.J, Houghton R.L and Matsumura M, et al. (2000). Identification of differentially expressed genes in human prostate cancer using subtraction and microarray. Cancer Res., 60, 1677-82
- [68] Brem R., Hildebrandt T., Jarsch M., Van Muijen G.N.and Weidle U.H (2001) Identification of metastasis-associated genes by transcriptional profiling of a metastasizing versus a non- metastasizing human melanoma cell line. Anticancer Res., 21(3B), 1731-40
- [69] Bratslavsky G, Sanford T, Srinivasan R, Liu J, Linehan WM *et al* (2011) Differential Genetic Expression in Large Versus Small Clear Cell Renal Cell Carcinoma: Results from Microarray Analysis. **2**:271-279
- [70] Aguado M (2007) Microarrays de AND en microbiología. RCCV. **1**:2.
- [71] Ehrenreich A (2006) DNA microarray technology for the microbiologist: an overview. Microbiol. Biotechnol. **73**: 255-273.
- [72] Garaizar, J., Rementería, A, Porwollic, S (2006) DNA Microarray technology: a new tool for the epidemiological typing of bacterial pathogens? Immunol. Med. Microbiol. **47**: 178-189.
- [73] Wilson M, DeRisi J, Kristensen HH *et al* (1999) Exploring drug-induced alterations in gene expression in Mycobacterium tuberculosis by microarray hybridization. Proc Natl Acaad Sci U S A: **96**:12833-12838.
- [74] Lucchini S, Thompson A, Hinton JCD (2001) Microarrays for microbiologists. Microbiology. **147**: 1403-1414.
- [75] Kurz S, Hübner C, Aepinus C, Theiss S *et al* (2002) Transcriptome-based antigen identification for *Neisseria meningitidis*. Elsevier Science. **21**:768–775
- [76] Ochman H. y Moran, NA (2001) Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science. **292**: 1096-1098.

- [77] Matsui A, Ishida J, Morosawa T *et al* (2008) *Arabidopsis* transcriptome analysis under drought, cold, high- salinity and ABA treatment conditions using a tiling rray. *Plan Cell Physiol*: **49**: 1135–1149.
- [78] Carney GE (2007) A rapid genome-wide response to *Drosophila melanogaster* social interactions. *BCM Genomics*, **8**:288 doi:10.1186/1471-2164-8-288.
- [79] Bradley KC, Boulware MB, Jiang H, Doerge RW, Meisel RL, Memelstein PG. (2005) Changes in gene expression in the nucleus accumbens and striatum following sexual experience. *Genes Brain Behav.* **4**:31-44
- [80] Drosophila 12 Genomes Consortium (2007) Evolution of genes and genomes on the Drosophilaphylogeny. *Nature*. **450**: 203-218
- [81] Kurppa K, Hietanen E, Klockars M, Partinen M, Rantanen J, Rönnemaa T, Viikari J. (1984). Chemical exposures at work and cardiovascular morbidity. Atherosclerosis, ischemic heart disease, hypertension, cardiomyopathy and arrhythmias. **10**: 381-388.
- [82] Sahu SN, Lewis J, Patel I, Bozdog S, Lee JH y cool. (2013) Genomic Analysis of Stress Response against Arsenic in *Caenorhabditis elegans*. *PLoS ONE* **8**: e66431. doi:10.1371/journal.pone.0066431.
- [83] Budowle B, Johnson MD, Fraser CM, Leighton TJ, Murch RS, Charabarty R (2005) Genetic analysis and attribution of microbial forensic evidence. *Crit Rev Microbiol*: **31**: 233-254.
- [84] Willse A, Straub TM, Wunschel SC, Small JA, Call DR, Daly DS, Chandler DP (2004). Quantitative oligonucleotide microarray fingerprinting of *Salmonella enterica* isolates. *Nucleic Acids Res*: **32**: 1848-1856.