

# ¿QUÉ SON LOS MICROARRAYS? APLICACIÓN AL DIAGNÓSTICO DE ANOMALÍAS CONGÉNITAS

M.D. Sánchez-Izquierdo<sup>1</sup>, M.L. Martínez-Fernández<sup>1,2</sup>, M.L. Martínez-Frías<sup>1,2,3</sup>

<sup>1</sup>ECEMC. Centro de Investigación sobre Anomalías Congénitas (CIAC), Instituto de Salud Carlos III, Ministerio de Ciencia e Innovación. Madrid.

<sup>2</sup>Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER). Madrid.

<sup>3</sup>Profa. Dpto. De Farmacología de la Facultad de Medicina de la Universidad Complutense de Madrid

## Summary

### Title: What are microarrays? Applications for the diagnosis of birth defects

Arrays are made up of small fragments of DNA from known locations within each chromosome, and labelled probes that bind covalently to a silica or glass surface in a specific position (hence the name *in silico*). Both the sequence and genomic position of each probe are recorded in a database associated with a computer analysis program. These sensors detect changes in gene sequence, particularly in the number of copies, but also in the methylation status or heterozygosity. There are many different types of arrays available and it is possible to differentiate between them in terms of density, distance and the number of probes that they contain, as well as their distribution throughout the genome. In this paper we review the different types of arrays and the current situation in diagnosing patients with birth defects.

## Introducción histórica. El inicio de los *arrays*

El objetivo de los primeros *arrays* fue medir de forma simultánea la expresión de muchos genes en un tejido o tipo celular concreto. Estos primeros 'microchips', o ***arrays de expresión***, contenían pequeños trozos de ADN complementario a los ARN mensajeros de los genes conocidos<sup>1</sup> y detectaban los genes que estaban siendo expresados en un momento y en un grupo celular concreto (tejido, cultivo celular, muestra tumoral...). Esto permitía realizar todas las mediciones en un mismo experimento, evitando los sesgos entre ensayos, y ahorrar material procedente de los pacientes. La comparación entre muestras de pacientes e individuos sin enfermedad determinaba qué genes variaban su expresión, relacionándose con la enfermedad.

Por otra parte, los primeros *arrays* utilizados (Cuadro 1) para detectar cambios en el número de copias de ADN se basaron en la hibridación genómica comparada denominada CGH (por sus siglas en inglés *Comparative Genomic Hybridization*), una técnica que utilizaba el marcaje fluorescente de forma análoga a la hibridación *in situ* con fluorescencia (FISH por *Fluorescence In Situ Hybridization*). La CGH consiste en hibridar juntos, sobre un portaobjetos que contiene cromosomas sin alteraciones, un ADN de referencia marcado en rojo,

y el ADN del paciente, marcado en verde. El perfil de fluorescencia obtenido determina si existen cambios en la cantidad de material genético del paciente porque, si el resultado sobre cada cromosoma es normal el color será amarillo, pero si hay ganancias se incrementa la señal verde, y si hay pérdidas la señal será roja (Figura 1). Al sustituir los cromosomas del portaobjetos por unos fragmentos de ADN de localización conocida, de forma análoga al *array de expresión*<sup>2,3</sup> se obtiene un ***array genómico*** (Figura 2). El tamaño de los *arrays* genómicos se ha ido reduciendo con los años al disminuir también el tamaño de las sondas<sup>4</sup>, y es por ello que hoy en día reciben nombres como 'microchips de ADN' o *microarrays*.

En la actualidad, los *arrays* están constituidos por pequeños fragmentos de ADN de localización conocida dentro de cada cromosoma y denominados sondas, que se unen covalentemente a una superficie de sílice o cristal en unas posiciones concretas (de ahí la denominación *in silico*). Tanto la secuencia como la posición genómica de cada sonda quedan registradas en una base de datos asociada a un programa de análisis informático. Estas sondas pueden detectar cambios en la secuencia génica, principalmente en el número de copias, pero también en el estado de

metilación o de heterozigosis.

Aunque los primeros *arrays* genómicos fueron desarrollados en laboratorios de citogenética molecular y orientados a la detección de ganancias y pérdidas

de ADN en muestras tumorales, ya se aplican a una gran variedad de patologías, incluyendo también los reordenamientos asociados a defectos congénitos. Si bien es cierto que se han convertido en poderosas

FIGURA 1.

A) Cromosomas clasificados en un experimento de hibridación genómica comparada CGH (los cromosomas cruzados no son analizables porque en los cruces se distorsiona el perfil de fluorescencia). Los cromosomas 1p, 6p, entre otros, muestran incrementos en la señal verde correspondiente a ganancias de ADN en el paciente. B) Imagen obtenida después del análisis de varias metafases hibridadas con CGH: por convención, a la derecha de cada cromosoma se muestran las ganancias en verde, por ejemplo en 6p o en el cromosoma 7 completo, y a la izquierda en rojo las pérdidas, como 11q. Debajo de cada perfil se identifica el cromosoma y el número de perfiles analizados.

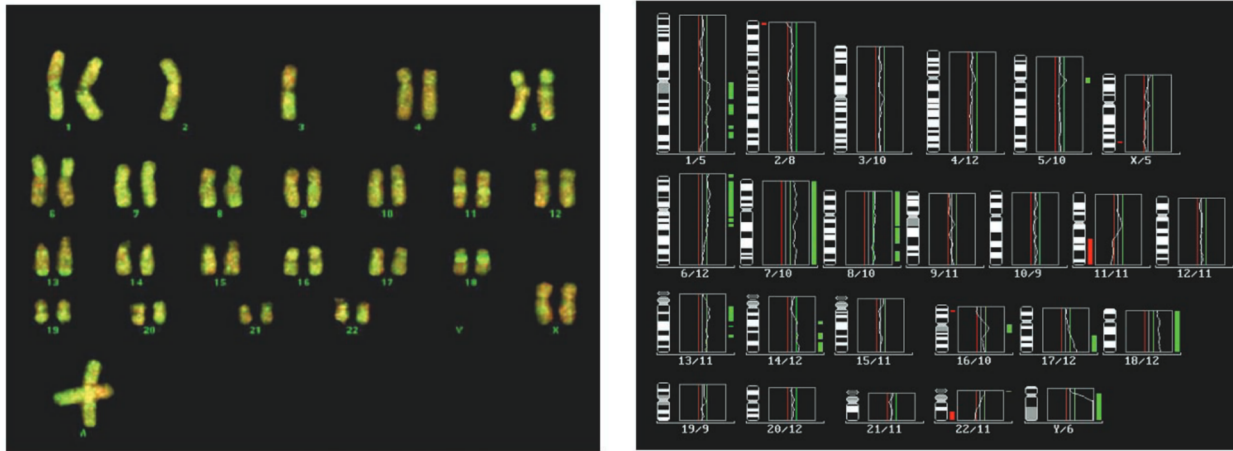
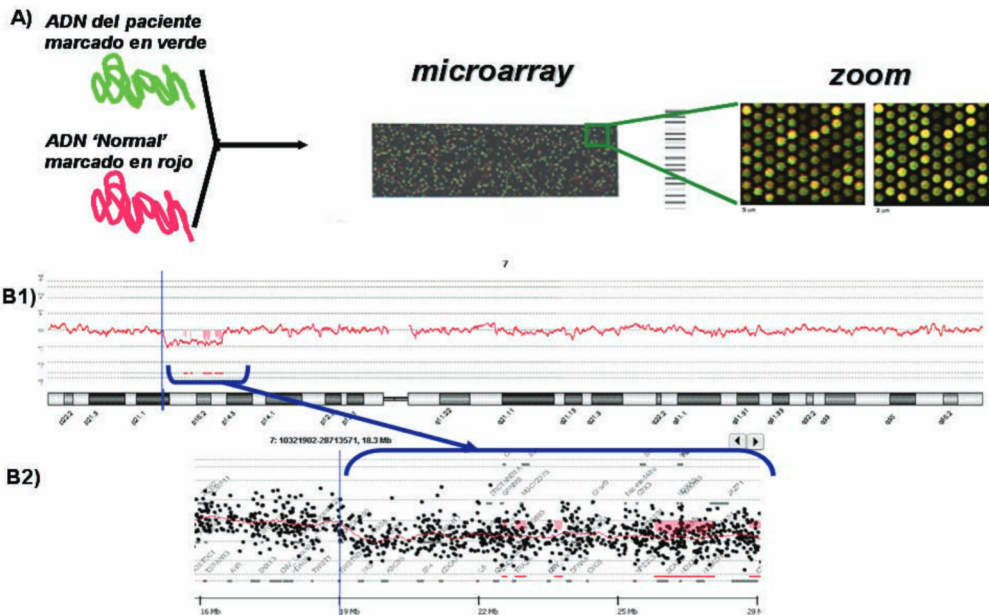


FIGURA 2.

A) Esquema de cómo se realiza un CGH array mostrando la imagen de un microarray hibridado y un zoom sobre una región de la matriz de puntos fluorescentes. B1) Resultado del análisis de un CGH array en un paciente que presenta una pérdida heterocigota (una sola copia) de material genético en 7p. B2) Al ampliar la zona perdida en 7p se ven los puntos negros correspondientes a las sondas de ADN de localización conocida en el cromosoma que contiene el microchip.



herramientas diagnósticas, su uso e interpretación presenta aún ciertas limitaciones que deben ser tenidas en cuenta a la hora de utilizar la información que de ellos se obtiene en el diagnóstico y manejo clínico del paciente.

Existen muchos *arrays* en el mercado y es posible diferenciarlos tanto en función de la densidad, distancia o número total de sondas que contienen, como de la distribución de éstas a lo largo del genoma. Por ejemplo, los **arrays de genoma completo**, denominado WGA (del inglés *Whole Genome Array*), analizan todos los cromosomas (son los más frecuentemente utilizados), mientras que los **arrays específicos** detectan alteraciones en zonas concretas del genoma como pueden ser telómeros, regiones que se expresan (exones), islas de metilación CpG, zonas implicadas en microdeleciones o translocaciones citogenéticas, entre otras. Por otra parte, los **arrays a la carta** o 'customizados', permiten al investigador realizar su propio diseño del *array*, ubicando las sondas en áreas genómicas concretas de su interés: una banda citogenética concreta, regiones promotoras o genes relacionados con una patología determinada.

El progreso en la tecnología de fabricación permite

que hoy en día los *arrays* puedan incluir más de dos millones de sondas, llegando a detectar alteraciones de unas pocas kilobases [Kb] (Cuadro 2). Sin embargo, y tal y como se expondrá más adelante, el incremento en la resolución no va a ser siempre sinónimo de una mayor capacidad de diagnóstico. Esto se debe en gran medida a la gran complejidad estructural mostrada por el genoma humano. También se verá que, aunque todos los *arrays* localizan ganancias y pérdidas de material genético, sólo algunos diseños son capaces de definir específicamente las regiones con disomía uniparental (UPD del inglés *UniParental Disomy*) o el grado de metilación de determinadas zonas del genoma. Por otro lado, cada vez es más frecuente encontrar en la literatura estudios de asociación genómica o GWAS (del inglés *Genome-Wide Association Study*), que utilizan los *arrays* para determinar regiones candidatas o que generan susceptibilidad a una u otra patología. En estos estudios se analizan amplios grupos de pacientes y de población control dentro de una misma área geográfica y, mediante potentes herramientas estadísticas, se identifican variantes poblacionales, tanto de número de copias (CNV de *Copy Number Variation*) como de polimorfismos de una sola base (SNP de *Single Nucleotide Polymorphism*), que se

#### CUADRO 1 DEFINICIONES

**Sonda:** Fragmento de ADN de localización conocida dentro del genoma

**FISH (Fluorescence In Situ Hybridization):** Hibridación de sondas de ADN marcadas

**BAC (Bacterial Artificial Chromosome):** Contiene secuencias genómicas de unas 100 Kb que son utilizadas como sondas tanto en técnicas de FISH como de arrays

**SNP (Single Nucleotide Polymorphism):** Pequeñas secuencias del genoma que varían en un único nucleótido (una base) de un individuo a otro

**CNV (Copy Number Variation):** Ganancias o pérdidas de material genético, en general de pequeño tamaño, aunque pueden llegar a tener varias Mb

**UPD (UniParental Disomy):** Un fragmento o una pareja de cromosomas que se reciben del mismo progenitor

**CGH (Comparative Genomic Hybridization):** Hibridación simultánea de material genético

**GWAS (Genome-Wide Association Study):** Estudio amplio y comparativo que establece asociaciones con variantes genéticas y ciertos fenotipos, comparando un grupo de individuos que tienen una determinada característica con otro grupo sin esa característica (grupo control)

## CUADRO 2

**UNIDADES DE MEDIDA DE LA SECUENCIA DEL GENOMA**

**Pb (par de bases):** Se refiere a un par de bases de la cadena de ADN (un nucleótido de cada una de las dos moléculas del ADN). Es la unidad básica de la secuencia del ADN.

**Kb (kilobase):** Se refiere a mil pares de bases.

**Mb (megabase):** Se refiere a un millón de pares de bases.

encuentran con mayor frecuencia en el grupo de pacientes que en el de la población control. Este tipo de estrategia ha permitido identificar muy recientemente un posible gen candidato para el síndrome de Kabuki<sup>5</sup>. Actualmente hay numerosos estudios GWAS en marcha cuyo objetivo es caracterizar nuevas variantes génicas asociadas a patologías tan diferentes como pueden ser el autismo o la susceptibilidad a sufrir un infarto de miocardio o a padecer diabetes (<http://www.ncbi.nlm.nih.gov/sites/entrez?Db=gap>).

## Características de los diferentes tipos de arrays

### 1. Arrays genómicos

En general, se puede decir que estos arrays utilizan principalmente dos tipos de sondas (Cuadro 1): cromosomas artificiales bacterianos (BAC) y oligonucleótidos. Los arrays de oligonucleótidos pueden estar basados en la técnica de CGH o bien utilizar secuencias de SNP sobre las que se realiza una hibridación directa sin ADN de referencia.

El **BAC array**, constituido por BACs, se basa en la CGH y contiene sondas de un tamaño de unas 100 Kb que son iguales a las utilizadas en la técnica de FISH. Debido al gran tamaño de las sondas, la resolución es algo menor que en los arrays basados en oligonucleótidos, por lo que detectan alteraciones de un tamaño algo mayor. En compensación a esta limitación, la hibridación es muy fiable y altamente específica, y los resultados obtenidos son relativamente sencillos de interpretar. Además, las alteraciones detectadas son sencillas de validar, ya que es posible obtener de la casa comercial la sonda o sondas que se encuentran ganadas o perdidas y aplicarlas en un ensayo de FISH sobre la muestra estudiada.

El **CGH array** hibrida una mezcla de ADN control

(procedente de donantes sanos) junto a la muestra del paciente. Está compuesto de pequeñas sondas de entre 20 y 60 nucleótidos que determinan con gran precisión qué genes se encuentran afectados por una pérdida o ganancia. Además, la distribución de las mismas puede ser homogénea ya que no depende de la presencia de secuencias génicas específicas, como es el caso del **SNP array**. Esto consigue que el nivel de resolución del array no varíe demasiado de unas posiciones genómicas a otras.

En el **SNP array** el ADN objeto de estudio se hibrida directamente sobre el array. Los valores de fluorescencia obtenidos en cada ensayo son normalizados primero de forma interna y, a continuación, comparados frente a un gran número de muestras de población control previamente analizadas. La comparación estadística de las señales obtenidas frente a la media de las señales de las muestras de la población control determina qué sondas muestran ganancias o pérdidas. Los resultados obtenidos en los arrays realizados a la población utilizada como control se encuentran accesibles en bases de datos públicas dentro del proyecto HapMap: (<http://hapmap.ncbi.nlm.nih.gov>). Las sondas en el **SNP array** corresponden a secuencias SNP y, además de detectar ganancias y pérdidas, permiten conocer el estado de homocigosis o heterocigosis de una región genómica. Esto es posible porque estas sondas discriminan si las dos copias de la secuencia génica que se une a ellas en la hibridación poseen el mismo nucleótido (homocigosis) o uno distinto (heterocigosis). Con esta estrategia se detectan aquellas regiones con la misma secuencia génica en las dos copias heredadas, sospechando la existencia de UPD.

### 2. Arrays de expresión

Hasta ahora, se han descrito principalmente las características de los arrays utilizados con mayor frecuencia en el estudio del ADN genómico. Como

se ha comentado en la introducción, los *arrays* de expresión fueron los primeros aplicados a la clínica con el fin de evaluar los niveles de expresión de los ARN mensajeros en una situación y grupo celular concreto. Los ensayos de este tipo se han aplicado fundamentalmente a tumores de diferente origen, con diferentes características clínico-patológicas, o a células tratadas con un determinado fármaco en comparación con células no tratadas. En el caso concreto de la aplicación al estudio de los defectos congénitos, su utilización se encuentra por el momento limitada al campo de la investigación. Así, este tipo de *arrays* se ha utilizado, por ejemplo, en modelos murinos y en muestras de cerebros de personas con síndrome de Down<sup>6,7</sup>, para determinar qué genes del tejido cerebral cambian sus niveles de expresión debido a la trisomía 21. El objetivo consiste en establecer qué moléculas deben contribuir mayoritariamente al fenotipo asociado, así como buscar dianas terapéuticas útiles en el tratamiento de algunas de sus manifestaciones clínicas.

## Situación actual de las técnicas de *array* para el diagnóstico de los pacientes con defectos congénitos: ventajas e inconvenientes

### 1. Capacidad resolutive

Como ya se ha comentado, la mayor ventaja de las técnicas de *array* es la resolución. Ésta viene determinada fundamentalmente por el número de sondas que contiene, pero además, y paradójicamente, por el uso al que se destinen los resultados obtenidos. Resulta obvio que el *array* genómico permite el estudio de los cromosomas que se observan en el cariotipo a un nivel de resolución muy superior al de la citogenética convencional, porque pone de manifiesto el perfil génico del paciente. Dicho perfil, en algunos casos, permitiría establecer el diagnóstico de la patología del paciente e, incluso, de aspectos que aún no se han manifestado por tratarse de efectos clínicos de aparición tardía. Una información anticipatoria que propicia un mejor manejo del paciente, así como paliar y, en ciertos casos, prevenir las patologías asociadas a la evolución de la enfermedad.

Sin embargo, en la práctica diaria, los criterios para determinar si una alteración concreta puede estar relacionada con una patología son mucho más

restrictivos que si se trata de un estudio en el área de la investigación. Estos criterios incluyen parámetros como el número de sondas ganadas o perdidas y el tamaño de la alteración detectada. Como el valor de esos parámetros depende de la densidad con la que se encuentra cubierta una determinada región genómica, la distribución de las sondas en relación con el genoma es crucial a la hora de validar la existencia de una pérdida o ganancia concreta. Por tanto, una alteración en una zona del genoma representada con pocas sondas en el *array* será más difícil de validar porque poseerá una resolución inferior a la de otra zona con mayor densidad de sondas. Es fácil entender que la resolución dentro de un mismo *array* es un valor relativo y variable a lo largo del genoma. En la Tabla 1 se muestran diseños de *arrays* con diferente densidad de sondas (1 millón, 400.000, 180.000, etc.) y se observa como varía la distancia media entre sondas en las zonas intragénicas e intergénicas. Por otra parte, al aumentar el número de sondas, es posible detectar un gran número de variantes consistentes en ganancias o pérdidas de material genético de tamaño variable, que no tienen trascendencia diagnóstica. Esto es así, porque se ha demostrado que en el genoma humano existe una enorme cantidad de esos tipos de variantes sin repercusión clínica, que se consideran polimórficas y se denominan CNV (Cuadro 1). Por esto, ante resultados de pérdidas o ganancias en muestras de pacientes, aún resulta difícil interpretar clínicamente los resultados que se obtienen mediante *arrays* de muy alta densidad.

### 2. ¿Qué detectan y no detectan los arrays?

#### A. Síndromes de microdelección y UPD

Tanto el *CGH array* como el *SNP array* detectan desbalances en el ADN, pero existen diferencias entre ambos que vienen dadas por las características del diseño de cada uno y de la metodología empleada en el análisis de los datos obtenidos. Por tanto:

Con el *CGH array* y el de *BAC array*, basados en la técnica de CGH, se obtienen resultados más fiables por dos motivos:

a. Son comparativos y se eliminan errores en la señal de fluorescencia que producen las regiones que hibridan mal (heterocromatina, subteloméros y subcentroméros) y que dificultan la normalización de los datos.

b. Las sondas son de mayor tamaño, lo que implica una mayor especificidad y una mejor señal de fluorescencia (especialmente los *arrays* de BACs),

TABLA 1. Diseños de arrays de diferente densidad o número de sondas. (1 millón, 400.000, 180.000 y 60.000). El array se identifica con relación al número de sondas que incluye. Los datos se expresan en pares de bases. Todos los diseños incluyen un pequeño grupo de sondas repetidas (en este caso 5 veces) para comprobar que se hibridan de forma homogénea. Los controles negativos permiten establecer el ruido de fondo o fluorescencia basal ya que no hibridan con el ADN y no deberían dar señal. Las sondas en exones corresponden a la secuencia del gen que se expresa y las regiones intragénicas e intergénicas a las zonas dentro y entre genes, respectivamente. La distancia media en general en el array es mayor que la distancia en y entre genes porque en las zonas del genoma 'vacías' de genes hay menos sondas.

Denominación	1M	400K	180K	60K
nº de sondas	974.016	420.288	180.880	62.976
Sondas repetidas (5x)	1.000	1.000	1.000	1.000
Controles negativos	302	106	7	13
Sondas en exones	51.297 (5,3%)	35.945 (8,7%)	24.011 (14,1%)	14.259 (25,9%)
Sondas intragénicas	514.337 (53,4%)	218.058 (53,0%)	93.675 (55,0%)	36.995 (67,2%)
Sondas intergénicas	448.692 (46,6%)	192.998 (47,0%)	76.659 (45,0%)	18.082 (32,8%)
Distancia media sondas				
intragénicas	1.814	4.606	11.190	33.307
intergénicas	2.752	6.876	17.655	78.946
CNV	2.092	5.246	12.690	26.688
Distancia media general	3.118	7.304	17.627	54.455

aunque su resolución es menor.

Con el SNP *array*, se obtienen los siguientes resultados:

a. El proceso de normalización en cada experimento se basa en valores de referencia de sondas específicas dentro del array (Tabla 1), por lo que la calidad de la hibridación es crítica a la hora de realizar el análisis.

b. Se detectan, además de las pérdidas y ganancias de ADN, las regiones UPD por lo que si se sospecha un síndrome que pueda tener su origen en un suceso de UPD, se debe aplicar este tipo de *array* preferentemente.

c. Se comparan los resultados obtenidos en los pacientes con los obtenidos en un grupo de personas de la población general (grupo control), por lo que algunas variantes polimórficas sin repercusión clínica pueden no encontrarse en el grupo control y serán difíciles de interpretar si se detectan en la muestra de un paciente.

Esta última limitación se está resolviendo mediante la aplicación del SNP *array* a un gran número de genomas de individuos de poblaciones diferentes en todo el mundo<sup>8</sup>. Además, se ha creado una base de datos pública que recoge todas las variantes, tanto de CNV como de SNP, que pueden ser consultadas para identificar aquellas alteraciones detectadas en los pacientes que corresponden a polimorfismos (<http://projects.tcag.ca/variation/>).

## B. Análisis citogenético y translocaciones

En el contexto de la citogenética clínica, los *arrays* detectan con gran precisión los puntos de rotura únicamente en las translocaciones no balanceadas, en las cuales se produce la pérdida de parte del ADN adyacente o próximo al punto de rotura de uno o ambos cromosomas afectados por la translocación. Por otro lado, y a diferencia del cariotipo, no detectan translocaciones balanceadas o aquellas con un desbalance cuyo tamaño sea inferior a la capacidad de resolución del *array* empleado. Por tanto, el estudio del cariotipo resulta en la actualidad necesario y complementario, especialmente en los casos de translocaciones heredadas que precisan de un buen asesoramiento genético a los padres. Además, porque es conocido desde hace tiempo que las translocaciones balanceadas pueden alterar la expresión de los genes localizados en las proximidades del punto de rotura, debido a efectos posicionales, cambios en los promotores, alteración de los sitios de procesamiento alternativo de exones (*splicing*), interrupción de genes o de secuencias reguladoras de la expresión (*enhancers*), entre otros. Por el momento, los *arrays* tampoco detectan mutaciones puntuales, las cuales deben ser caracterizadas por secuenciación genómica u otras técnicas moleculares.

Otra limitación de las técnicas de *arrays* viene determinada por la calidad de la muestra. Al ser un análisis estadístico y de normalización el que finalmente

determina si una alteración está o no presente, la fiabilidad del ensayo depende de la dispersión de los datos. Esta dispersión viene dada por la intensidad y variabilidad de la señal fluorescente que, a su vez, se relaciona directamente con la calidad de partida del ADN hibridado. Como el resultado final es un dato estadístico, en caso de detectarse una alteración que se relacione claramente con el fenotipo, es muy recomendable comprobar mediante otras técnicas más objetivas la presencia real de dicha alteración. Estas técnicas pueden ser la de FISH (si la alteración es lo suficientemente grande y se dispone de la sonda adecuada), la PCR cuantitativa con los oligonucleótidos correspondientes a las sondas detectadas en la región alterada y la secuenciación génica o el análisis de microsatélites en los casos de UPD. Si el gen o genes afectados implican cambios en determinados parámetros fisiológicos, se puede plantear otro tipo de análisis indirectos como, por ejemplo, un estudio bioquímico de las proteínas relacionadas con la alteración.

En resumen, de estas técnicas se pueden destacar los siguientes aspectos:

a) Los estudios con *arrays* incrementan muy sensiblemente la resolución de los estudios citogenéticos.

b) Permiten analizar desde secuencias específicas de ADN hasta el genoma completo.

c) No detectan translocaciones balanceadas pero sí la existencia de ganancias y pérdidas asociadas a translocaciones no balanceadas, siempre y cuando el desbalance tenga un tamaño mayor al límite de resolución de la técnica utilizada.

d) Sólo los *arrays* basados en SNPs son capaces de detectar las regiones con UPD por lo que son los más indicados cuando existe una sospecha de síndrome cuya causa pueda ser una UPD.

e) La alta resolución de esta metodología ha revelado la existencia de una gran complejidad en nuestro genoma, consistente en la existencia de una gran cantidad de CNVs y de SNPs (para más

información consultar referencia 9). El significado de estas variantes, consideradas polimórficas, es aún incierto y complican, además, la interpretación de los resultados en las muestras de los pacientes.

## Referencias

1. Lipshutz RJ, Morris D, Chee M, Hubbell E, Kozal MJ, Shah N, Shen N, Yang R, Fodor SP. Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques*. 1995 Sep;19(3):442-7.
2. Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Döhner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer*. 1997 Dec;20(4):399-407.
3. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet*. 1998 Oct;20(2):207-11.
4. Lucito R, West J, Reiner A, Alexander J, Esposito D, Mishra B, Powers S, Norton L, Wigler M. Detecting gene copy number fluctuations in tumor cells by microarray analysis of genomic representations. *Genome Res*. 2000 Nov;10(11):1726-36.
5. Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet*. 2010 Sep;42(9):790-3.
6. Sultan M, I Piccini, Balzereit D, Herwig R, Saran NG, Lehrach H, Reeves RH, Yaspo ML. Gene expression variation in Down's syndrome mice allows prioritization of candidate genes. *Genome Biology* 2007, 8:R91
7. H.E. Lockstonea, 1, L.W. HARRISA, 1, J.E. Swattonb, M.T. WaylandA, A.J. Hollandc and S. Bahna. Gene expression profiling in the adult Down syndrome brain. *Genomics* 2007 Dec; 90(6):647-660.
8. The International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations. *Nature* 2010 Sep;467:52-58.
9. Shaikh TH. Oligonucleotide arrays for high-resolution analysis of copy number alteration in mental retardation/multiple congenital anomalies. *Genet Med*. 2007 Sep;9(9):617-25. Review.