



**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO**

**FACULTAD DE MEDICINA
DIVISION DE ESTUDIOS DE POSGRADO**

**APLICACION CLÍNICA DE LOS
MICROARREGLOS EN GENÉTICA**

**ESTANDARIZACIÓN DEL ANÁLISIS DE NÚMERO DE COPIAS EN
PACIENTES CON REVERSIÓN SEXUAL**

TESIS

QUE PARA OBTENER EL TÍTULO DE

MÉDICO GENETISTA

PRESENTA

LUIS MANUEL ZAPEDA INCLÁN

DIRECTOR DE TESIS: DRA. GLORIA EUGENIA QUEIPO GARCIA



MÉXICO, D.F .

2008



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

1. INTRODUCCION

1.2.1. SNPs

1.3.1. GENOMICA Y USO CLINICO DE LOS MICROARREGLOS

1.4.1. PROYECTO HapMap

1.5.1. DIAGNOSTICO

1.6.1. CANCER

2. PLANTEAMIENTO

3. JUSTIFICACION

4. OBJETIVO

4.1.1 OBJETIVO PRIMARIO

4.2.1 OBJETIVO SECUNDARIO

5. MATERIAL Y METODOS

6. RESULTADOS

7. DISCUSION

8. CONCLUSIONES

9. REFERENCIAS

ABSTRACT

Los cambios en el número de copias de segmentos de DNA contribuyen a la patogénesis de diversas patologías. Los microarreglos de alta densidad de SNPs pueden ser usados para detectar estas alteraciones en el número de copias reforzando nuestras habilidades para descubrir genes causantes de patología humana en el desarrollo de algunos órganos.

En el presente estudio se demuestra la utilidad de los microarreglos de 500 000 SNPs para caracterizar cambios en el número de copias en DNA genómico, ya sean amplificaciones o deleciones, en pacientes con reversión sexual. Aquí se reportan las ganancias y pérdidas de segmentos de DNA en un panel de análisis conformado por 30 pacientes con reversión sexual y 270 controles obtenidos de la primera fase del HapMap, utilizando la misma plataforma de microarreglos Affymetrix de 500 000 SNPs.

1. INTRODUCCION

La morfogénesis en el humano comprende una serie de procesos complejos, aún no bien entendidos, en los que participan activamente factores genéticos y ambientales. Los defectos al nacimiento (DN) son un grupo heterogéneo de enfermedades que afectan la estructura, función o el metabolismo del individuo y se observan en alrededor de 1 de cada 40 RNV (Spranger J et al., 1988). Estos padecimientos son causas importantes de discapacidad y de muerte infantil durante el primer año de vida, en donde los factores etiológicos en la gran mayoría de los casos no se encuentran bien dilucidados. La estandarización de estudios genómicos en el abordaje de estos padecimientos resulta esencial como un primer paso hacia el uso rutinario de métodos genómicos que tienen aplicación clínica.

POLIMORFISMOS DE UN SOLO NUCLEOTIDO (SNPs)

Los SNPs (del inglés single nucleotide polymorphism) son cambios en la secuencia de DNA que ocurren por lo menos 1% de la población y cada SNP se considera como un polimorfismo bialélico en donde se cambia un solo nucleótido por otro. Se estima que existe un SNP por cada 1000 pares de bases en el genoma humano y hay por lo menos 9 millones de SNPs en común en la población general, lo que constituye 90% de la variación genética en la población humana. Los SNPs pueden ser sinónimos, lo que significa que aún cuando cambio un nucleótido, no da lugar a un cambio en el aminoácido codificado; por el contrario los SNPs no sinónimos si originan un cambio de aminoácido en la proteína resultante cuando el SNP se encuentra en una región codificante. Ejemplo: El codón CAU codifica para histidina que es un aminoácido polar cargado, al igual que el codón CAC, sin embargo CAA codifica para glutamina que es un aminoácido polar sin carga y esto daría lugar a una mutación no conservativa por cambiar un aminoácido por uno con diferentes propiedades. El estudio de estos cambios ha demostrado que existen polimorfismos funcionales que pueden afectar genes de manera adversa cambiando aminoácidos (SNPs no sinónimos), introduciendo codones de paro prematuros, o generando cambios en sitios de splicing o regiones reguladoras. Sin embargo, el significado biológico de estos polimorfismos aún está en estudio. Debido a su abundancia, estabilidad, y distribución homogénea a lo largo del genoma, los SNPs ofrecen gran potencial para diagnosticar

enfermedades humanas mediante análisis automatizado de microarreglos (Joblin M.A. et al, 2004)

GENOMICA Y USO CLINICO DE LOS MICROARREGLOS

El descubrimiento de la estructura del DNA y de las técnicas para manipularlo, dio lugar a una revolución en la segunda mitad del siglo pasado. Sin embargo, las técnicas limitantes hicieron que el enfoque fuera el análisis de cambios en genes y proteínas en forma individual. De esta manera, la medicina y la biología en general han cursado un camino entre sí, esencialmente reduccionista. Gracias a este reduccionismo, se cuenta con mucha información sobre el funcionamiento de un gran número de genes y proteínas a nivel molecular. En la actualidad se está llevando a cabo una segunda revolución en la biología y la medicina. Esto se debe a la capacidad cada vez mayor de analizar los ácidos nucleídos en forma global, tanto a nivel del DNA como del RNA. Los estudios masivos de ácidos nucleicos han desencadenado cambios en la medicina que conducen a un enfoque más global y menos reduccionista. Con el advenimiento de la tecnología del arreglo se lograrán analizar múltiples anomalías y/o variaciones submicroscópicas en el genoma humano en un solo experimento. Un arreglo (*del inglés array*) está conformado por una serie de elementos microscópicos ordenados específicamente en una superficie plana. Los arreglos están estructurados por una superficie sólida (vidrio, plástico o silicio), a la cual se unen una serie de fragmentos de DNA o RNA y constituyen una de las

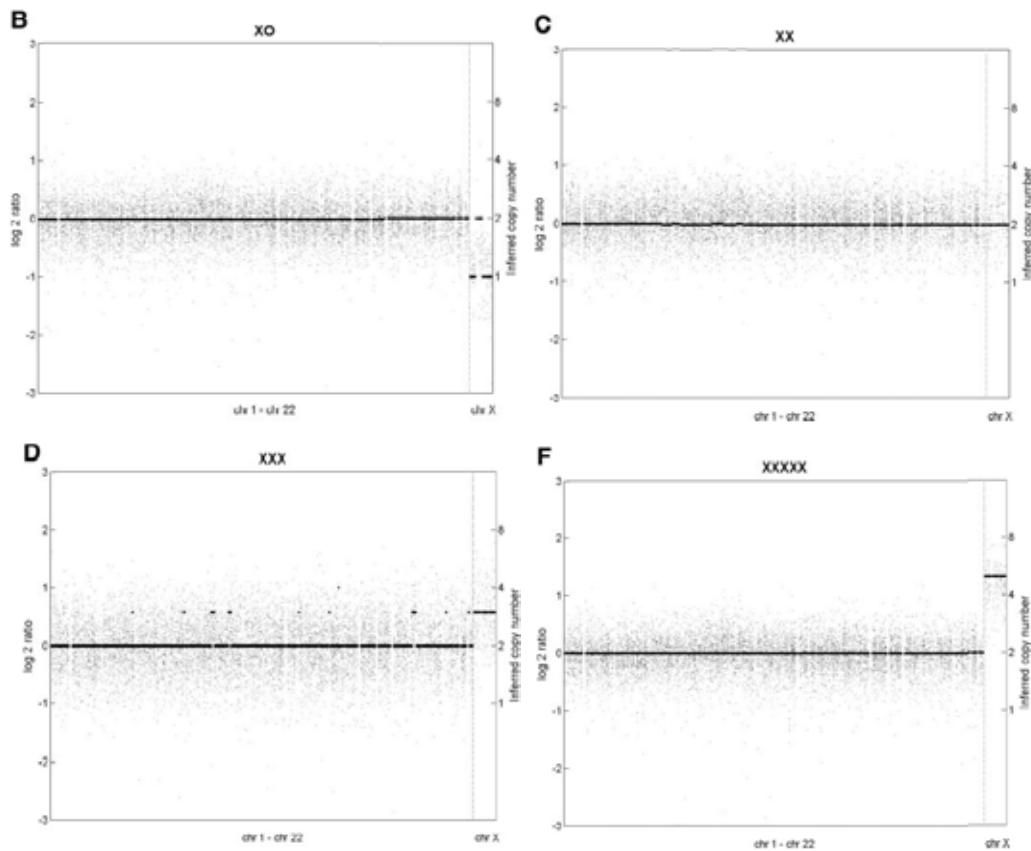
herramientas más importantes en la genómica porque permiten analizar varios genes al mismo tiempo.

Existen diferentes tipos de arreglos de RNA y DNA. Los primeros se utilizan para análisis cuantitativo de un transcrito. Los arreglos de RNA son utilizados para conocer los niveles de expresión de genes y monitorizar miles de transcritos de forma simultánea. Con esta tecnología se puede observar de forma casi instantánea la expresión de todos los genes del genoma de un organismo y por esta razón pueden ser utilizados para identificar genes que producen ciertas enfermedades. Los diferentes niveles de expresión entre células sanas y células patológicas pueden ser responsables del fenotipo en los individuos afectados. Los estudios con microarreglos de expresión han tenido una gran utilidad en el estudio de diferentes tipos de cáncer, ya que al analizar patrones de expresión se puede predecir metástasis. Hasta ahora la mayor atención ha sido para el pronóstico de supervivencia de los pacientes.

También se pueden utilizar arreglos para observar genotipos y frecuencias alélicas o para cuantificar el número de copias de un SNP determinado en una posición física conocida dentro del genoma humano, esto último se puede hacer de manera comparativa simple (casos y controles), o pareada con muestras del tumor y muestras de sangre periférica u otro tejido libre de tumor del mismo paciente. Los microarreglos de DNA para valorar número de copia son un método para detectar pérdidas o ganancias con gran resolución a lo largo del

genoma. Los arreglos de DNA para análisis número de copia son muy prometedores para el uso clínico ya que el DNA puede extraerse de tejidos preservados en parafina y es biológicamente más estable que el RNA. Se cree que los arreglos para el análisis de número de copias comenzarán a desplazar a los de expresión como herramienta clínica en el manejo del paciente con defectos congénitos o cáncer, gracias al potencial diagnóstico y pronóstico que ofrecen. Recientemente se han descrito segmentos de DNA con variación en número de copias (CNVs) en individuos fenotípicamente normales, éstas pueden ser deleciones, inserciones, o duplicaciones. Esto dificulta el análisis de polimorfismos de este tipo cuando no se conocen los rangos de variación normal en el número de copia de individuos sanos de una población. El estudio de CNVs ha tenido un impacto profundo en el conocimiento de las enfermedades relacionando a diversas patologías con el efecto directo de dosis génica (microdeleciones y microduplicaciones) o bien en forma indirecta a través de efecto de posición. En el último caso el defecto puede estar distante al gen que causa la enfermedad, sin embargo, puede tener un efecto deletéreo afectando la expresión génica. También se puede alterar la estructura del genoma haciéndolo más susceptible a cambios que pueden ocurrir en las mitosis o meiosis subsecuentes. Estas regiones también se han asociado con riesgo a enfermedades complejas y a diferentes locus relacionados con el desarrollo. Los CNVs pueden predisponer a cambios genéticos funcionales o pueden servir como sustrato para cambios evolutivos. Hace unos años, las herramientas metodológicas para poder detectar pequeñas ganancias o pérdidas en el

genoma de menos de 50kb no eran accesibles y su costo era elevado, sin embargo, con las nuevas plataformas de arreglos de polimorfismos de un solo nucleótido (SNPs) que abarcan más del 80% del genoma actualmente se puede realizar este tipo de estudios. Estos cambios son posibles causas de síndromes de genes contiguos o de susceptibilidad a enfermedades complejas. También pueden influenciar mediante efectos de posición, o ser blanco de cambios en la evolución de los cromosomas. Las regiones de número de copia variable tienen un promedio de 250Kb y cerca del 12% del genoma humano son CNVs. Estas regiones variables contienen más de 3000 genes que pueden estar relacionados con enfermedades (Redon R. et al 2006). Se cree que estos cambios en el número de copias son causados por eventos de recombinación. Para una muestra en particular, mientras más grande sea la distancia genética entre dos marcadores, es más probable que ocurran cambios en el número de copia dentro de esta región (Sebat J. et al 2004). Para validar la técnica de los microarreglos, Zhao X. et al en un estudio clásico analizaron líneas celulares con número de copias bien definidos. El DNA genómico aislado de células con 1 a 5 copias del cromosoma X, fue digerido, amplificado e hibridado a los microarreglos. Después de normalizar las señales de cada muestra, se digitalizaron las señales que representaban a cada SNP. La relación de la hibridación de cada SNP fue computarizada para las líneas celulares con X, XX, XXX, XXXX, XXXXX con respecto al promedio de muestras normales 46,XX (Gráfico1).



➤ **Gráfico 1: Número de copias del X de acuerdo a la dosis génica**

Los números de copia observados en todos los SNPs del cromosoma X correlacionaron con el número conocido de X. Este estudio motivó a muchos investigadores a valorar el análisis de número de copias en DNA de individuos con rasgos fenotípicos de interés.

Otros tipos de arreglos, como los de re-secuenciación, están surgiendo como una estrategia que podría reemplazar los métodos de rutina actuales para conocer secuencias de DNA dentro del genoma humano y facilitar la detección de mutaciones. Recientemente se han comercializado arreglos que facilitan el

análisis de los patrones de metilación en patologías relacionadas con impronta genómica. Aunque los arreglos se han utilizado principalmente en cáncer, esta tecnología se usa en estudios neurológicos, de autoinmunidad, trasplantes, endocrinología de la reproducción, trauma, sépsis y enfermedades cardiovasculares (Pfeifer J.D. et al, 2006). Entre otras aplicaciones, los arreglos de SNPs pueden ser utilizados para identificar variaciones entre individuos o diferentes poblaciones. Estos SNPs podrían ser los responsables de la mayoría de la variación genotípica dentro de una población, la fuente de susceptibilidad a distintas enfermedades. En general, la aplicación de estas técnicas de genotipificación son usadas para conocer la predisposición a enfermedades o incluso permitir el uso de medicamentos específicos para tratar enfermedades bien caracterizadas de acuerdo a cada individuo. Los arreglos de SNPs también son utilizados para identificar mutaciones somáticas, como la pérdida de heterocigocidad, la amplificación o la delección de regiones de DNA en el genoma individual de pacientes afectados.

Las principales desventajas de los microarreglos en la actualidad son el costo, la complejidad del diseño del análisis y el tiempo requerido para procesarlos. Pueden dar resultados falsos positivos cuando no hay suficiente muestra de casos o no hay controles adecuados o arrojar falsos negativos cuando las muestras no son de alta calidad. Sin embargo lo más difícil, sigue siendo la validación clínica de estos resultados. El uso clínico puede verse limitado por la dificultad de conseguir el espécimen blanco. A diferencia los tumores que se extirpan rutinariamente, otros tejidos son difíciles de obtener y en algunas enfermedades sistémicas no están bien definidos los tejidos involucrados.

PROYECTO HapMap

Este proyecto internacional se puso en marcha en octubre 2002 para crear una base de datos pública sobre la variación genómica en humanos que servirá como guía en estudios genéticos y genómicos de fenotipos clínicos de los que se tiene poca información o información incompleta acerca de su fisiopatología. La utilidad y la practicidad mejoró al confluir: la disponibilidad de la secuencia del genoma humano, la creación de bases de datos útiles para diseñar los estudios de genotipificación. Los datos del HapMap fueron generados para servir como guía en el diseño y análisis de estudios médicos. Estos recursos abren una nueva era en la genética de poblaciones, con la oportunidad de investigar las fuerzas evolutivas que dieron lugar a la variación natural.

La primera fase del HapMap tuvo como objetivo al menos un SNP cada 5Kb a lo largo del genoma en cada una de las 269 muestras estudiadas, mientras que en la segunda fase el objetivo fijado fue genotipificar 4.6 millones de SNPs adicionales. Las muestras examinadas son de 90 individuos formados por 30 tríos (padre, madre e hijo) de 4 poblaciones Yoruba, Nigeria (YRI); otros 30 tríos de Utah, USA del Centre d' Etude du Polymorphisme Humain collection (CEU); 45 de Beijing, China (CHB); 44 de Tokio, Japón (JPT). Dado que la frecuencia alélica de los Chinos y los Japoneses es muy similar, algunos análisis combinan datos en 3 paneles de análisis (YRI, CEU, CHB+JPT). La primera fase del HapMap contiene 1,007,329 SNPs que pasaron los filtros de control de

calidad en los 3 paneles de análisis, y dentro de estos hay 11,500 cSNPs (SNPs no sinónimos) en regiones codificantes. En la segunda fase se pretende caracterizar todos los cSNPs (Monpetit A 2005). Desde que se inició el HapMap se han identificado más de 10 000 000 de SNPs confirmados en el genoma humano en estas 4 poblaciones. Muchos de estos polimorfismos contribuyen a la heredabilidad en enfermedades multifactoriales. Es difícil examinar cada SNP para su asociación con la enfermedad. Por fortuna los haplotipos pueden ayudar, ya que varios SNPs adyacentes se heredan como una unidad. Estos SNPs no se separan durante la recombinación. A esta combinación de alelos se le conoce como haplotipo y a la secuencia de DNA que los contiene se le denomina bloque de haplotipo. Los haplotipos permiten usar solo pocos SNPs representativos para determinar el genotipo de una gran región genómica. El HapMap busca regiones con desequilibrio de ligamiento en el genoma por asociación preferencial entre alelos específicos de los SNPs cercanos. Si uno ocurre con mayor frecuencia de la esperada se dice que los marcadores están en desequilibrio de ligamiento entre sí (Watson J.D. 2007).

DIAGNOSTICO

La mayoría de las enfermedades crónico degenerativas tienen influencia multifactorial en las que se reconoce una carga genética en estrecha relación con el ambiente. Además, muchas otras enfermedades son de origen poligénico, por lo que determinar sus bases genéticas es difícil. Una vez identificado y caracterizado el gen se hace posible desarrollar pruebas moleculares para la enfermedad (DeSalle R. et al, 2005). Algunos de estos diagnósticos pueden ser predictivos y realizarse años antes de que se manifieste la enfermedad (Watson J.D. et al, 2007). Conforme adquieran validez clínica los estudios genómicos en diferentes tejidos serán usados como auxiliar diagnóstico rutinario de la práctica médica (Pfeifer J.D. et al, 2006). Limitados por el espacio de almacenamiento de datos y de la capacidad de análisis de los arreglos. La re-secuenciación, o secuenciación por hibridación se convertirá en el método de rutina para diagnóstico molecular de algunas enfermedades en las que ya se conoce el gen responsable de la enfermedad (Pfeifer J.D. et al, 2006). Los arreglos son la herramienta ideal para el estudio de neoplasias, ya que permiten evaluar número de copias y pérdida de heterocigocidad simultáneamente en todo el genoma (Watson J.D. et al, 2007). Las regiones variantes en el número de copia que son demasiado pequeñas para verse por citogenética convencional o molecular, pueden detectarse fácilmente y con mucha resolución utilizando microarreglos de ácidos nucleicos.

Las reacciones a fármacos varían enormemente y ponen en riesgo la vida de algunos individuos. Unos pacientes mejoran con un fármaco mientras otros no muestran mejoría pero sí efectos adversos serios. Las enzimas que metabolizan estos fármacos pueden conocerse mediante el uso de arreglos para poder ofrecer tratamientos personalizados. (Watson J.D. et al, 2007). La farmacogenómica intenta eliminar las reacciones adversas en medicamentos con un rango muy discreto entre toxicidad y dosis terapéutica (DeSalle R. et al, 2005).

2. PLANTEAMIENTO

Existen muchas enfermedades para las cuales se desconoce gran parte, o se ignora por completo las bases moleculares de su fisiopatología. Los pacientes con patologías representativas de dichas enfermedades pueden ser utilizados como modelos biológicos para estudiar y entender las diferentes vías de señalización implicadas en el desarrollo normal.

Los microarreglos son una herramienta ideal cuando se cuenta con escasa, o no se tiene información biológica suficiente en esas vías, ya que la simple cuantificación y análisis de los datos obtenidos, puede convertirse en nuevas propuestas o en diseños de experimentos más dirigidos o enfocados a un área en particular. Llevando a cabo la búsqueda de genes de enfermedades por medio de análisis de ligamiento y estudios de asociación se contribuye a la disección de la variación en el genoma humano y el desarrollo de métodos para analizar su variación en muestras grandes.

3. JUSTIFICACION

Para poder resolver interrogantes al respecto de alteraciones en la organogénesis y comprender mejor los procesos involucrados en el desarrollo de patologías clínicas, se necesita adquirir experiencia en el análisis de microarreglos dentro de un ambiente clínico, que tenga el suficiente respaldo académico y cuente con la infraestructura necesaria para procesar y analizar los microarreglos.

4. OBJETIVOS

OBJETIVO GENERAL

Estandarización de los Microarreglos de alta densidad GeneChip Mapping 500K Affymetrix en pacientes con alteraciones del desarrollo gonadal y cariotipo 46XX y 46,XY.

OBJETIVOS ESPECIFICOS

Estandarización de las condiciones de laboratorio óptimas para utilizar los microarreglos de alta densidad GeneChip Mapping 500K Affymetrix utilizando como modelo un grupo de pacientes con alteraciones del desarrollo gonadal XX y XY.

Estandarización de la paquetería informática necesaria para el análisis e interpretación de los datos.

Análisis de regiones variables en número de copias (CNV) de un grupo de pacientes con alteraciones del desarrollo gonadal XX y XY.

Comparación de los datos con el HapMap Internacional

Comparación de los datos con un grupo de individuos con desarrollo gonadal normal.

5. MATERIAL Y METODOS

Se utilizó el 500k Mapping Array de Affymetrix, un arreglo comercial hecho en California, diseñado con base en los resultados del análisis de desequilibrio de ligamiento de las poblaciones seleccionadas para el proyecto HapMap: Yoruba, Nigeria (YRI); Utah, USA del Centre d' Etude du Polymorphisme Humain collection (CEU); Beijing China (CHB) y Tokio, Japón (JPT).

La distancia promedio entre cada polimorfismo de un solo nucleótido (SNP) es de 5.8 Kb así, el 85% del genoma esta a 10 Kb de uno de los SNPs incluidos en el arreglo (Figura1).



Figura 1. Arreglo Affymetrix GeneChip Human Mapping 500K Array

Cada arreglo utilizó una enzima de restricción diferente, un arreglo utilizó la enzima de restricción *NspI* con 262 000 SNPs, mientras que *StyI* es la enzima para el complementario y contiene 238 000 SNPs. Los arreglos y su correspondiente ensayo deben procesarse independientemente del otro y separados por al menos una jornada laboral, de preferencia por diferentes individuos en diferentes días con áreas del laboratorio bien definidas.

Cada arreglo incluyó sondas que consisten en más de un millón de copias de un oligonucleótido de 25pb sintetizado por fotolitografía. El análisis de cada SNP es integrado por 6 o 10 oligonucleótidos que pueden hacer una hibridación tipo perfect match o mismatch para cada alelo, por lo que hay un total de 24 a 40 diferentes oligonucleótidos por cada SNP incluido en cada arreglo.

Para el procesamiento de la muestra se utilizaron el kit GeneChip Affymetrix Mapping 250K Assay que contienen reactivos validados y DNA de referencia como control positivo (Tabla 1).

En breve, la estrategia del ensayo fue reducir la complejidad del DNA. Se digirieron 250ng de DNA genómico total con una de las enzimas y se ligó a un adaptador que reconoce las 4 pares de bases de los extremos cohesivos. Para amplificar los fragmentos de DNA, se usó un primer genérico que reconoce el adaptador. Mediante una PCR se amplificaron principalmente fragmentos de 250 a 1100pb, los cuales se cortan DNasal para obtener fragmentos menores a

180pb. Posteriormente fueron marcados e hibridados en el chip, Por último se realizó una contra-tinción y se escanearon (Figura 2).

GeneChip Affymetrix Mapping 250K Assay

Adaptador	Dos oligonucleótidos específicos que se unen a los extremos cohesivos
Primer 002	Oligonucleótidos que reconocen el adaptador ligado a cada fragmento de DNA durante la PCR
DNA genómico de referencia 103	Control positivo de DNA humano
Reactivo de fragmentación	DNasa I para fragmentar los productos de PCR
Buffer de fragmentación 10x	Buffer utilizado en la reacción de fragmentación
Reactivo de marcaje de DNA	Marcaje de biotina para los amplicones.
Deoxinucleotidil transferasa terminal	Enzima utilizada para el marcaje
Buffer para la transferasa	Buffer para el marcaje
Reactivo Oligo control, 0100	Mezcla de 5 oligonucleótidos marcados que se hibridan a las regiones control de los arreglos

Tabla 1. Reactivos necesarios Affymetrix Mapping 250K Assay Nsp I

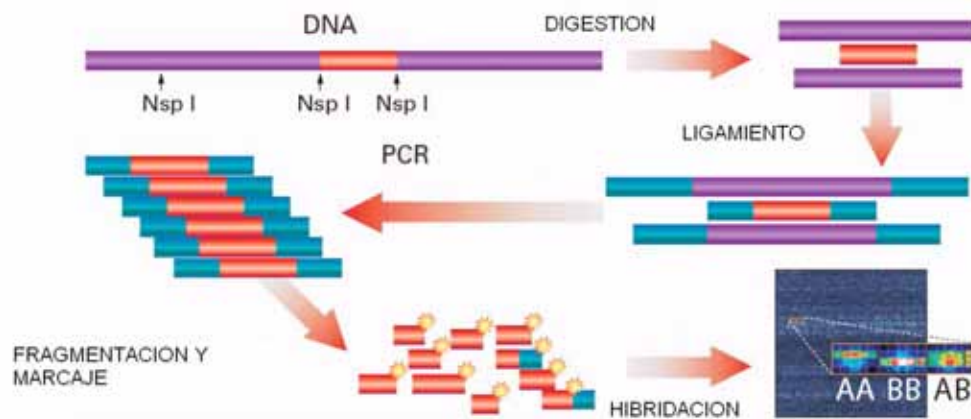


Figura 2. Esquema del proceso: inicia con digestión del DNA y ligamiento del adaptador que posteriormente se fragmentan antes de hibridar y finalizar con el escaneo.

Sujetos y métodos

Se realizó análisis genómico de pacientes con reversión sexual XX y XY, se analizaron mediante el uso de microarreglos de alta densidad Gene Chip 500K de Affymetrix, que contiene oligonucleótidos específicos que permite identificar SNP distribuidos en el genoma a una distancia promedio inter-marcadores de 5.8 Kb lo que representa el sondeo de aproximadamente de 93% del genoma humano.

Sujetos de estudio

Se incluyeron en el estudio 30 casos de pacientes con reversión sexual que pertenecen al banco de DNA del Servicio de Genética del Hospital General de México.

- Reversión sexual XX.

* 11 Pacientes con diagnóstico clínico de varón XX. Con cariotipo 46,XX y SRY negativos

* 11 Pacientes con diagnóstico clínico e histológico de Hermafrodita Verdadero. Con cariotipo 46,XX y SRY negativos

- Reversión sexual XY.

* 8 Pacientes con diagnóstico clínico de Disgenesia Gonadal Pura. Cariotipo 46, XY sin mutación en SRY.

Calidad de DNA

DNA genómico inicial debe ser de doble hebra, libre de inhibidores, sin contaminantes y no degradado. La calidad del DNA se analizó por electroforesis y espectrofotometría.

GENE CHIP HUMAN MAPPING 500K NspI

I. PREPARACIÓN DEL DNA

Se determinó la concentración del DNA genómico y se diluyó con buffer (EDTA 0.1mM, Tris HCl 10 mM): a una concentración de 50ng/μl. En total se prepararon 15μl de DNA genómico a (total 750ng)

II. DIGESTION

En el área de pre-PCR se digirieron 500ng de DNA genómico con una de las enzimas. Las concentraciones y condiciones utilizadas en la reacción se muestran a continuación (Tabla 2):

Mezcla de reacción	
Reactivo (stock)	Volumen (μl)
Agua	19.5
NE Buffer2 (10x)	4
BSA(10x)	4
<i>NspI</i> (10U/μl)	2
DNA	10
Volumen final	39.5

Termociclador	
Temperatura	Tiempo
37°C	120min
65°C	20min
4°C	hold

Tabla 2 Reactivos utilizados en la mezcla de digestión.

Para verificar que el DNA fue digerido, se corrió en un gel de agarosa al 1%, 2.0 μl de DNA sin digerir (100ng) y 7.9 μl de DNA digerido (100ng)

III. LIGAMIENTO

Los extremos cohesivos de cada fragmento de DNA se ligaron a una secuencia corta que funciona como adaptador para el primer. Las concentraciones y condiciones de la reacción se muestran a continuación (Tabla 3):

Mezcla de reacción	
Reactivo (stock)	Volumen (μl)
Adaptador Nspl (50 mM)	0.75
Buffer T4 DNA ligasa (10X)	2.5
T4 DNA ligasa (400U/ml)	2.0
DNA digerido	19.75
Volumen final	25.00

Termociclador	
Temperatura	Tiempo
16°C	180 min
70°C	20 min
4°C	hold

Tabla 3 Reactivos utilizados en la mezcla de ligamiento.

Al terminar la reacción de ligamiento se diluyeron 25 μ l de DNA ligado con 75 μ l de agua grado biología molecular fría.

IV. PCR

Durante la reacción de PCR se utilizó un oligonucleótido común que se une al adaptador ligado a cada fragmento de DNA del paso anterior. Se realizaron 3 reacciones de 100 μ l cada una por cada muestra con el fin de obtener la cantidad necesaria de producto de PCR para la fragmentación (90 μ g) e hibridación. se Agrego una mezcla de PCR sin DNA, como control blanco con el fin de identificar posibles contaminaciones. Se preparó la mezcla en hielo con 90 μ l de mezcla de PCR y 10 μ l de DNA ligado (Tabla 4).

Mezcla de reacción	Volumen		Ciclaje		
	(μ l)	3X+bco (μ l)			
Reactivo (stock)			1	94°C	3 min
Agua	39.5	158.0	30	94°C	30 seg
TITANIUM Taq PCR Buffer (10X)	10	40.0		60°C	45 seg
G-C melt (5M)	20	80.0		68°C	15 seg
dNTP (2.5mM) mix	14	56.0	1	68°C	7 min
Primer 002 (100uM)	4.5	18.0		4°C	hold
TITANIUM Taq DNA polimerasa (50X)	2	8.0			
Reacción de ligamiento	10	-			
Volumen final	100	360			

Tabla 4 Reactivos utilizados en la mezcla de PCR.

Se incubó en las condiciones indicadas y posteriormente se corrieron 3 μ l de cada PCR en un gel de agarosa al 1% para observar productos de PCR de 250 a 1100pb (Figura 3).

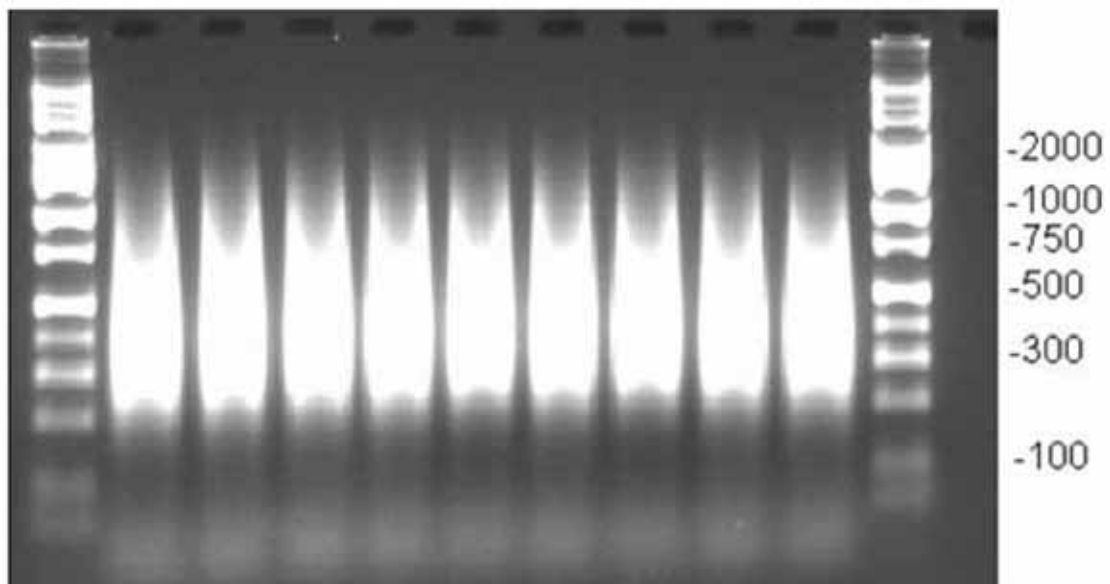


Figura 3. Gel de agarosa al 2% marcador de peso molecular en los carriles 1 y 11, el resto de los carriles muestra el DNA digerido.

Purificación de productos de PCR

La purificación se realizó mediante el Clean-up Kit. Se agregaron 8 μ l de EDTA 0.1M a cada reacción de PCR, se mezclaron las 3 reacciones y se colocaron en un solo pozo. A este pozo se le aplicó vacío a 600mbar hasta quedar completamente seco. Se llevaron a cabo 3 lavados con 50 μ l de agua grado biología molecular, aplicando 600mbar de vacío hasta secar. Se eluyó con 45 μ l de buffer RB y se agitó la placa durante 10 minutos a 1000 revoluciones por minuto para asegurar la elución y finalmente se transfirió la muestra a un tubo estéril para cuantificarla por espectrofotometría.

V. FRAGMENTACIÓN

Para lograr una hibridación más eficiente, esta reacción se realizó para obtener fragmentos menores a 180 pares de bases. El producto de PCR purificado se diluyó a una concentración de 90 μ g/45 μ l y se agregaron 5 μ l de buffer de fragmentación y 5 μ l de la siguiente mezcla de fragmentación conservada en hielo (Tabla 5):

Mezcla de fragmentación		TERMOCICLADOR	
	Volumen	Temperatura	Tiempo
Reactivo (stock)	2U/ μ l	37°C	35 min
Agua	52.5	95°C	15 min
Buffer de fragmentación 10X	6	4°C	Hold
Reactivo de fragmentación	1.5		

Tabla 5 Reactivos utilizados en la mezcla de fragmentación.

Se incubó en las condiciones adecuadas como control de calidad se corrieron 3µl de la muestra fragmentada en un gel de agarosa al 2% para observar fragmentos menores a 180 pares de bases (Figura 4). Inmediatamente después se procedió al marcaje.

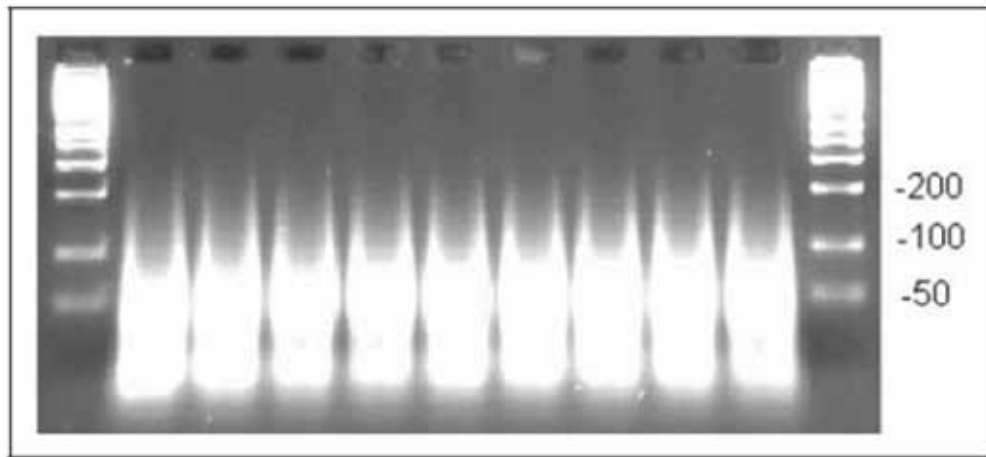


Figura 4. Gel de agarosa que muestra fragmentos de menos de 180 pb.

VI. MARCAJE

Los fragmentos obtenidos en la reacción anterior fueron marcados agregando 19.5µl de la siguiente mezcla (Tabla 6).

Mezcla de reacción	Volumen
Reactivo stock	(µl)
Buffer 5X TdT	14
Gene Chip DNA labeling reagent	2
TdT enzyme (30U/µl)	3.5
Volumen final	19.5

TERMOCICLADOR	
Temperatura	Tiempo
37°C	4 horas
95°C	15 min
4°C	hold

Tabla 6 Reactivos utilizados en la mezcla de marcaje.

VII. HIBRIDACIÓN

Se agregaron 190 μ l de la mezcla de hibridación a 70 μ l de la muestra marcada (Tabla 7).

MEZCLA DE HIBRIDACION

Reactivo	1x (μ l)
MES (12X, 1.22M)	12
Denhart's sol (50X)	13
EDTA (0.5M)	3
HSDNA (10mg/ml)	3
OCR. 0100	2
Human Cot-1 DNA (1mg/ml)	3
Tween-20 (3%)	1
DMSO (100%)	13
TMACL	140
Volumen total	190

Tabla 7 Reactivos utilizados en la mezcla de hibridación.

Se incubaron a 95°C durante 10 min, se mantuvieron a 49°C, luego se inyectaron 200 μ l al GeneChip 500K Affymetrix de la enzima correspondiente. Este último se incubo 18 horas a 40°C agitando a 60rpm en el horno de hibridación (Figura 5).



Figura 5. Horno de hibridación Affymetrix GeneChip Hybridization Oven 640

VIII. LAVADO, TEÑIDO Y ESCANEEO

Al terminar el tiempo de hibridación se quitó la mezcla de hibridación y se llenó el Gene Chip con 260 μ l de Buffer de mantenimiento (Tabla 8).

BUFFER DE MANTENIMIENTO

Reactivo	Volumen (μ l)
Agua	73.1
Buffer MES 12X	8.3
NaCl 5M	18.5
Tween-20 10%	0.1
Volumen final	100ml

Tabla 8 Reactivos utilizados en la mezcla de buffer de mantenimiento.

Se colocó el chip en la estación de fluidos GeneChip Fluidics Station 450 de Affymetrix, donde se llevaron a cabo lavados con las siguientes soluciones:

Solución SAPE en la posición 1 de la estación de fluidos.

594 μ l de Buffer de tinción + 6 μ l de estraptavidina-phycoerythrina (SAPE) = 600 μ l, solución de anticuerpo biotinilado en la posición 2 de la estación de fluidos, 594 μ l de Buffer de tinción + 6 μ l de anticuerpo biotinilado (0.5mg/ml), y por ultimo 800 μ l de buffer de mantenimiento en la posición 3 de la estación de fluidos (Figura 6).



Figura 6. Estación de fluidos Affymetrix GeneChip Fluidics Station 450

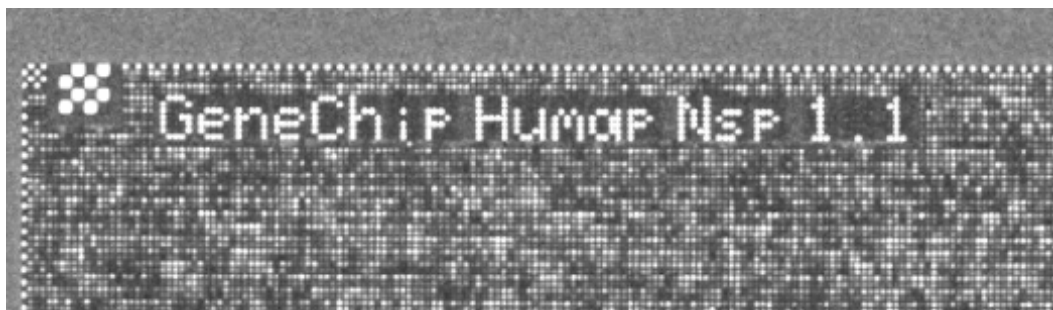
Mediante del programa GeneChip Operating Software (GCOS) versión 1.4 se seleccionó el protocolo Mapping 500Kv1_450

ESCANEEO

El proceso de escaneo se realizó en un Scanner 3000 7G (Figura 7) se aseguró que la imagen del arreglo escaneado indicara en el extremo inferior izquierdo el nombre de la enzima correspondiente al arreglo (Figura 8).



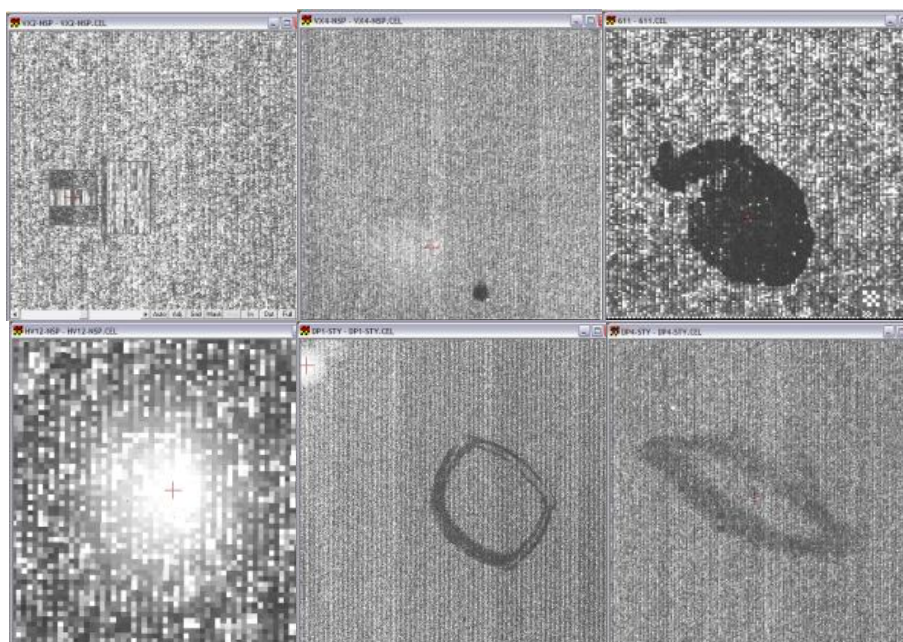
Figura 7. Scanner 3000 7G Affymetrix



A



B



C

Figura 8. (A) Imágenes obtenidas en formato .cel para ambas enzimas, Nsp I y (B) Sty I en. (C) Las imágenes resultantes del escaneo pueden verse con varios niveles de acercamiento en diferentes tonalidades como se muestra en.

6. RESULTADOS

El presente estudio se enfocó en el análisis de copias y pérdida de heterocigocidad. Para ello se siguieron de forma precisa las indicaciones del protocolo, los tiempos y programas de los termocicladores estuvieron apegados a las indicaciones de la compañía. También se hizo lo posible por trabajar en días separados el procesamiento de las dos diferentes enzimas para evitar contaminación cruzada entre muestras y obtener call rates adecuados. Una vez obtenido los resultados de la hibridación con las condiciones adecuadas, se procedió al análisis de los datos, esto incluyó implementación y utilización de los diferentes programas incluido el Scanner 3000 Affymetrix, disponibles en la página web de Affymetrix. El primero en instalar y estandarizar su uso fue el GCOS, luego el GTYPE y por último el CNAT. Se siguieron de forma precisa las indicaciones del protocolo, los tiempos y programas de los termocicladores estuvieron apegados a las indicaciones de la compañía. También se hizo lo posible por trabajar en días separados el procesamiento de las dos diferentes enzimas para evitar contaminación cruzada entre muestras y obtener call rates adecuados. Una vez obtenido los resultados de la hibridación con las condiciones adecuadas, se procedió al análisis de los datos

INSTALACION DE LOS PROGRAMAS

Los datos crudos de los pacientes y los controles obtenidos del ensayo se almacenaron en archivos electrónicos que fueron analizados en múltiples formatos compatibles con programas de GeneChip Affymetrix. Para realizar los diferentes tipos de análisis a lo largo del estudio se instalaron sucesivamente los programas GCOS 1.4 (GeneChip Operating Software), GTYPE 4.1 (GeneChip Genotyping Analysis Software) y CNAT (Copy Number Analysis Tool). Estos programas permitieron importar y exportar datos entre las diferentes aplicaciones con las que cuenta el programa. El programa GCOS 1.4 fue la base de datos cardinal para el análisis inicial y la interacción con otros sistemas de cómputo (Figura 9).

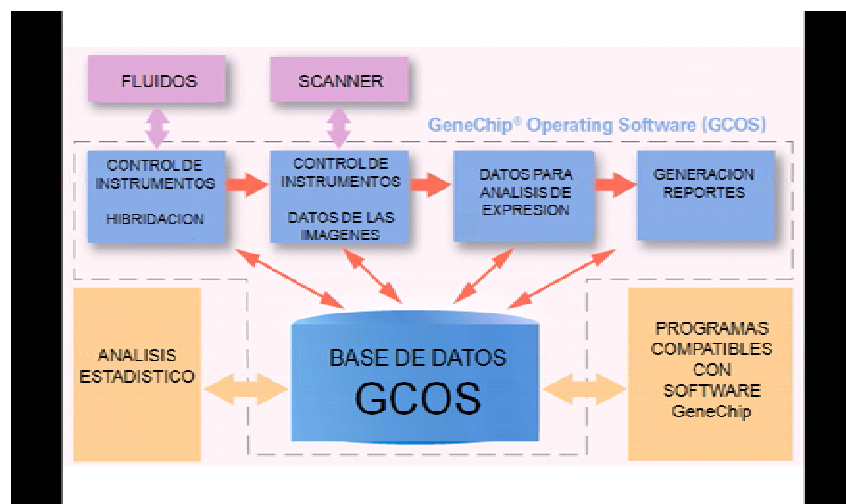


Figura 9. Diagrama de flujo de la principal base de datos GCOS.

En la figura 10 se muestran las diferentes herramientas funcionales del GTYPE después de la instalación (Figura 10).

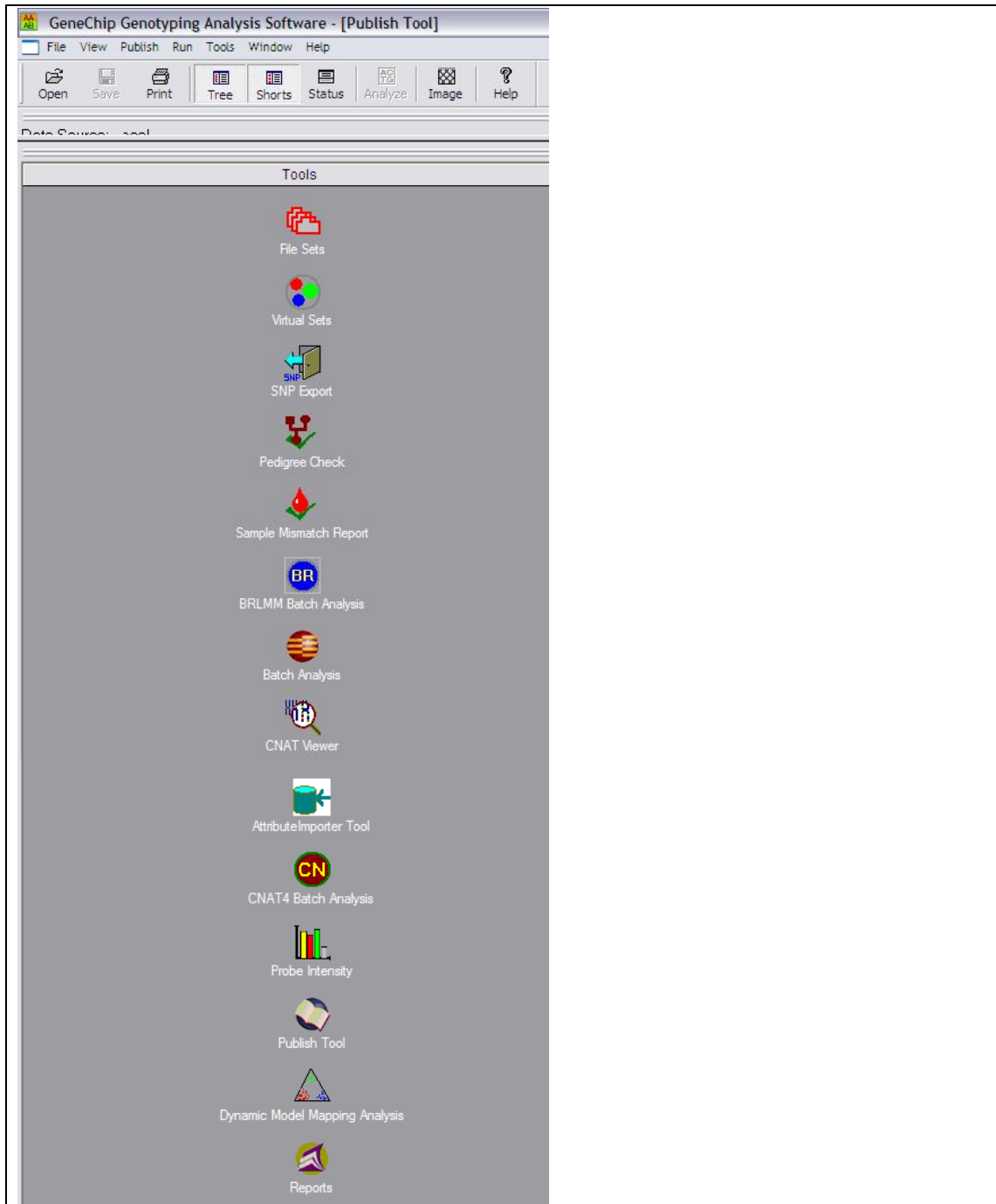


Figura 10. Múltiples herramientas disponibles en GTYPE versión 4.1.

DIGITALIZACION DE LA IMAGEN Y OBTENCION DE LOS DATOS

La configuración estándar para instrumentos es un GeneChip Scanner 3000 7G y 3 estaciones de fluidos GeneChip 450s controlados por un programa central, GeneChip Operating Software (GCOS). Es posible agregar varias estaciones de fluidos “Fluidics Station 450s” en cadena lo que permitirá en un futuro realizar de forma óptimas más ensayos. En este caso una computadora central controló los instrumentos y el análisis de datos se realizó en otro equipo. Se exportaron los datos crudos de la computadora del laboratorio central y se importaron a las computadoras destinadas exclusivamente al análisis, esto se hizo mediante la herramienta Data Transfer Tool que viene integrada en el paquete del GCOS. Los archivos que se generaron fueron *.DAT y *.CEL que son específicos del software Affymetrix, es importante mencionar que estos archivos no son reconocidos por Windows XP.



Figura 11. Los datos se pueden ser importar o exportar entre diferentes computadoras con el Data Transfer Tool (DTT).

ANALISIS DEL CHIP

Una vez que se obtiene la imagen del chip (*.CEL) se transfirió mediante la herramienta Data Transfer Tool (DTT) al GCOS ver 1.4 (Figura 11). Esto permitió eliminar aquellas muestras no deseadas o que no se incluyeron en el análisis y se procedió a la genotipificación.

GENOTIPIFICACION

Esta se realizó utilizando “GeneChip Genotyping Analysis Software (GTYPE) versión 4.1” (Figura 12). Este programa de cómputo diseñado por Affymetrix utiliza diferentes algoritmos automatizados que determinan la calidad de la muestra e hibridación.

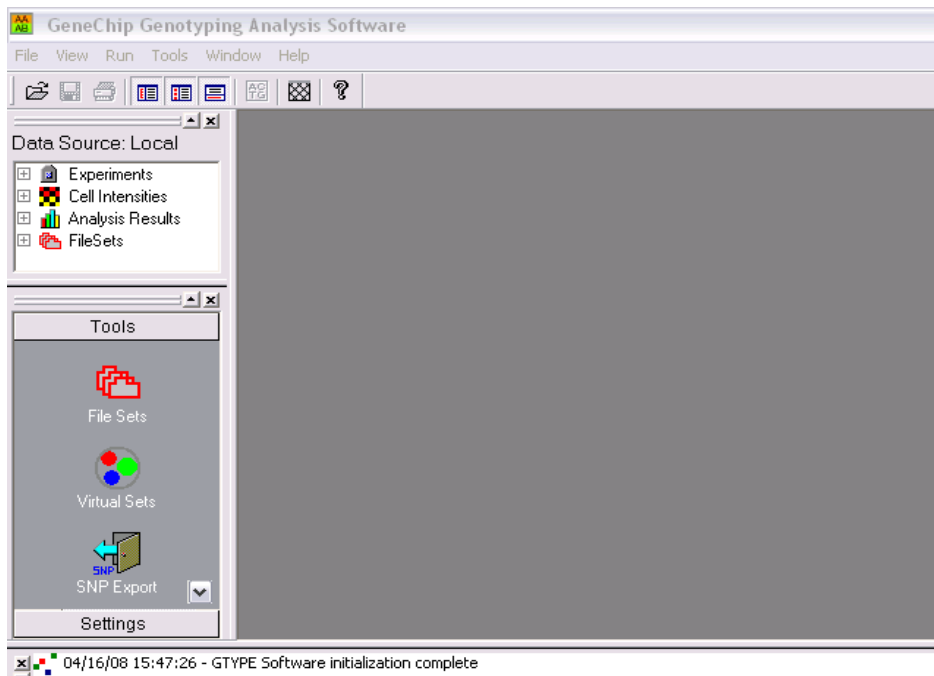


Figura 12. GTYPE versión 4.1

Existen dos algoritmos disponibles (Figura 13) el primero es un Modelo Dinámico (DM) y el segundo utiliza un algoritmo basado en análisis Bayesiano (BRLMM). Unas de las ventajas del segundo es la capacidad para poder analizar múltiples “chips” y poder funcionar independientemente de forma aislada o de manera integrada dentro de GTYPE versión 4.1 como lo hicimos en este estudio. Para llevar a cabo el análisis con el BRLMM fue necesario correr el análisis con el algoritmo DM, ya que es un requisito para continuar con el algoritmo Bayesiano del BRLMM.

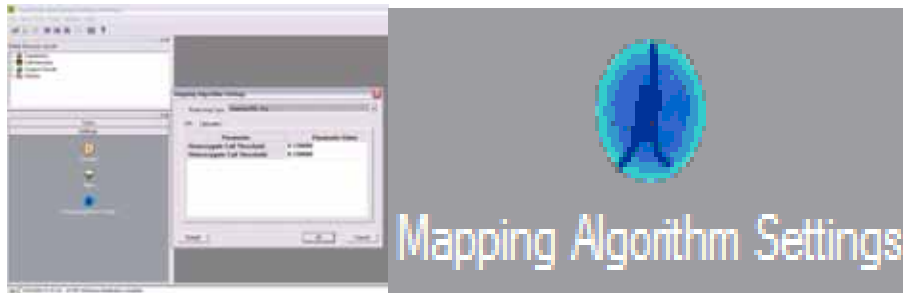


Figura 13. Establecimiento de los parámetros que se utilizados en cada algoritmo.

VIRTUAL SET

La herramienta Virtual Set (Figura 14) se encuentra disponible dentro del programa GTYPE 4.1, esta herramienta permitió integrar los datos obtenidos de ambas enzimas de una misma muestra.

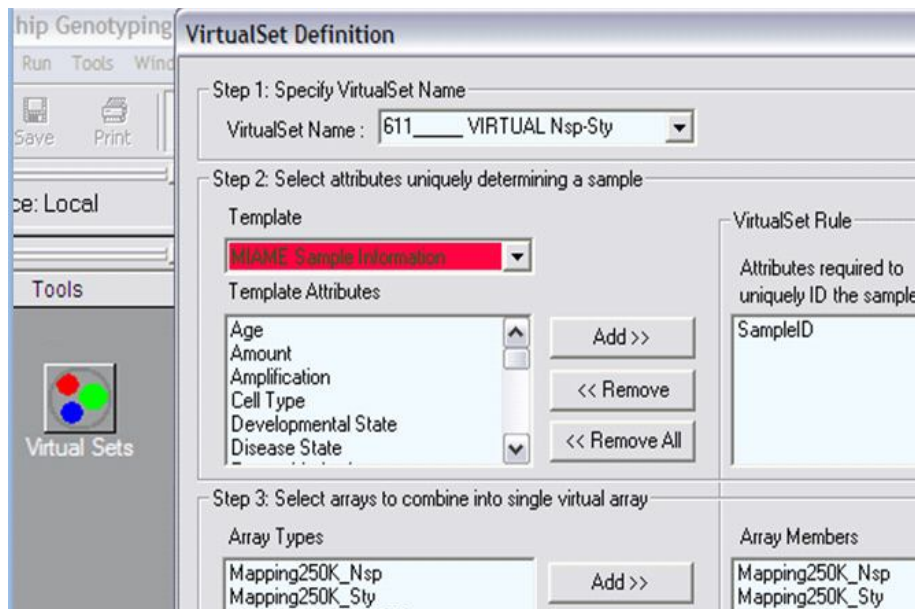


Figura 14. Creación de sets virtuales con GTYPE versión 4.1 para combinar los datos obtenidos de cada enzima por separado.

Una vez realizada la genotipificación se inició el análisis de los archivos *.CEL de cada paciente y de cada individuo usado como control, para lo cual se utilizó la herramienta Batch Análisis del GTYPE 4.1 (GeneChip Genotyping Analysis Software). Esta aplicación generó una serie de archivos *.CHP para cada paciente y cada control, que son necesarios para el análisis de número de copia con el programa Copy Number Analysis Tool (CNAT). Este programa comparó los datos individuales de cada paciente con reversión sexual contra los datos de los controles sanos como grupo. Estos individuos fueron elegidos con el propósito de estandarizar los parámetros de los programas y permitir un análisis fino en análisis subsecuentes. El DNA geonómico de referencia incluido el kit Affymetrix fue usado como control positivo durante el proceso de estandarización. Se muestra el manejo de datos dentro del CNAT (Figura 14).

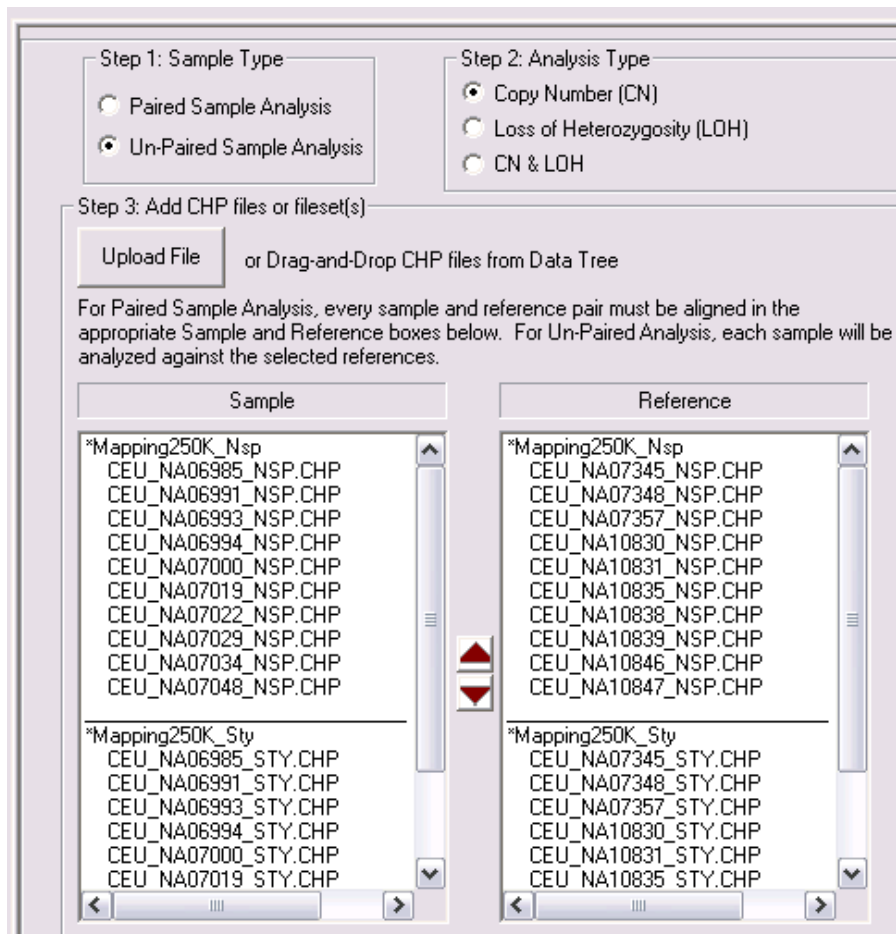


Figura 14. Se observa el manejo de datos de un experimento con el programa CNAT. La columna del lado izquierdo muestra una lista de pacientes y el listado de la derecha a los controles. La línea horizontal separa por tipo de enzima utilizada en cada ensayo.

La calidad de hibridación del microarreglo (Figura 15) se define mediante el indicador "CALL". Para un ensayo óptimo este debe ser mayor a 94%. Los factores que pueden disminuir este indicador son: errores metodológicos o desviación del protocolo, la utilización de un DNA que no cumpla con los parámetros de calidad, mala calidad de los reactivos y finalmente errores en la cuantificación del DNA. Es importante mencionar que el programa tiene un

algoritmo adicional para reconocer muestras contaminadas, sin embargo los valores de “CALL” obtenidos a lo largo de este proceso fueron ideales para continuar con el análisis.

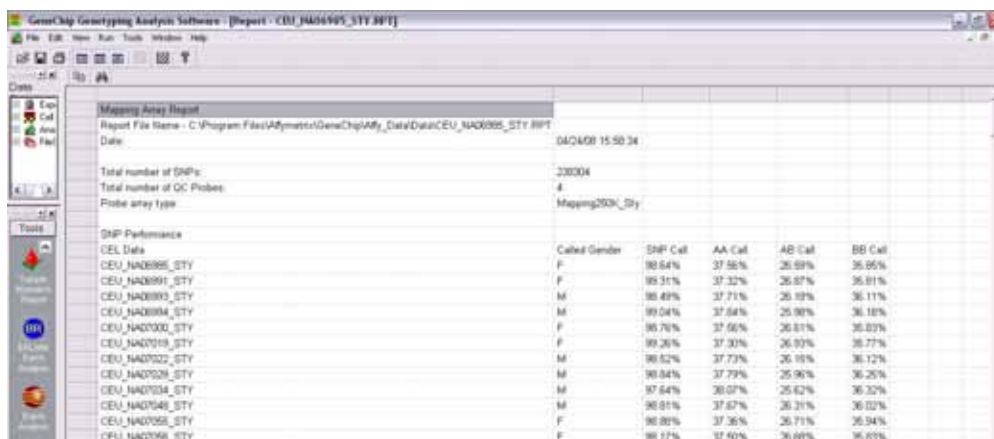
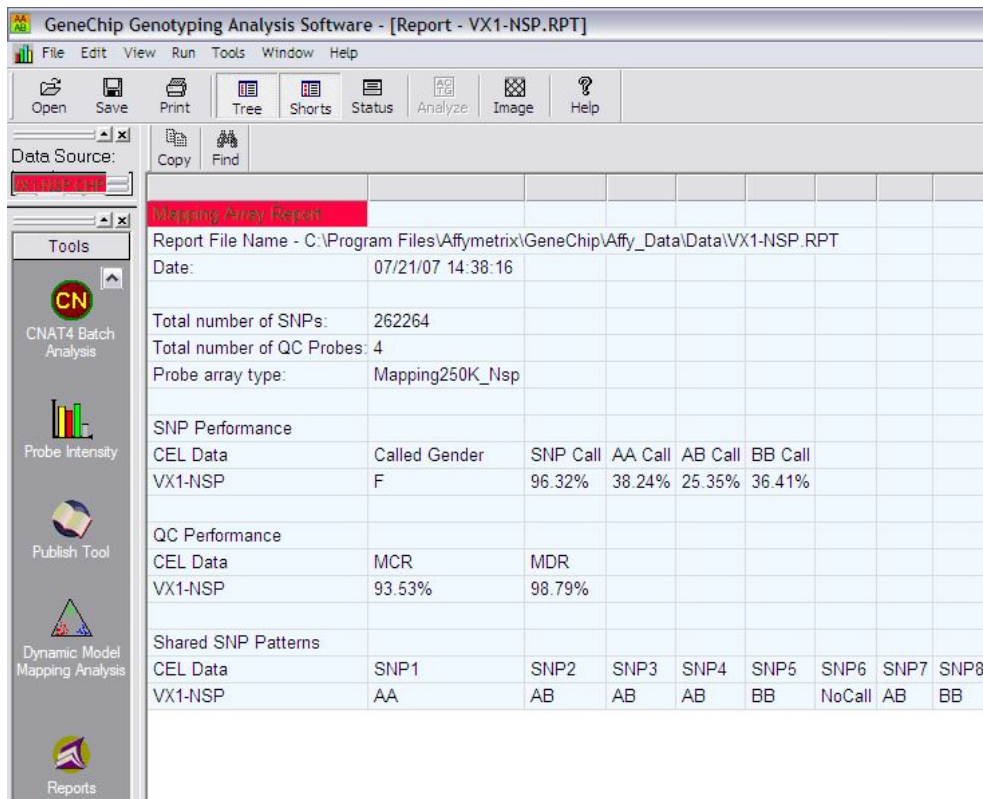


Figura 15. Reportes y control de calidad de los datos de un experimento donde se muestra el sexo de cada muestra y el contenido de SNPs para cada enzima utilizada en el ensayo.

Una vez creados los sets virtuales con la información de ambas enzimas tanto para casos como para controles, se analizaron los genotipos fusionados a manera de tablas de datos o de forma grafica con el CNAT (Copy Number Analysis Tool). El CNAT también permite analizar perdida de heterocigocidad o ambos tipos de análisis simultáneamente (Figura16).

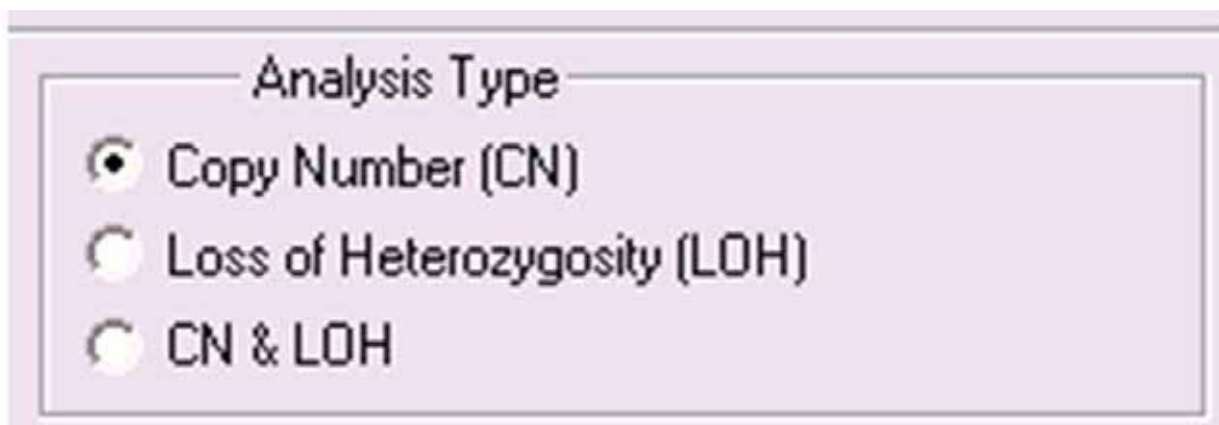


Figura 16. Se selecciono únicamente análisis de número de copias.

Dentro de este mismo programa, (Figura 17A) CNAT viewer permitió visualizar las ganancias y pérdidas con gráficos comparativos por cromosoma individual que se pueden personalizar por colores o seleccionar los tipos de gráficos mostrados (Figura 17B) eliminando así los datos que no son útiles para el usuario en ese momento. Cuando se ha detectaron deleciones (< 2 copias) o amplificaciones (3 copias o más del mismo SNP) se exportaron los datos del

individuo con la presencia de alteración hacia el IGB (Integrated Genome Browser) para manejar con facilidad a varios pacientes a la vez en un contexto de la secuencia de genómica que puede incluir los genes relacionados a dichos SNPs, el nucleótido en que se encuentran y la citobanda del cromosoma que se está visualizando. Esta práctica herramienta puede ocuparse con diferentes niveles de zoom para ver todo el cromosoma o bien solamente una región delimitada con gran detalle y así se puede ver el límite entre intrones y exones así como discernir entre promotores y regiones intergenicas, lo que permite precisar si el polimorfismo de un solo nucleótido cae en región codificante o no. Además de esto se muestran los diferentes transcritos alternativos y los exones involucrados en cada uno de ellos. Es recomendable actualizar la información del IGB mensualmente para contar con la más reciente información acerca del genoma humano y el HapMap. Todo esto al igual que las otras herramientas de los programas Affymetrix ocurre bajo un ambiente automatizado compatible con usuarios con conocimientos limitados de programación y bioinformática.

Se realizó un análisis mediante el programa CNAT v 4.0 para identificar variaciones en el número de copia. En dicho análisis se comparó la intensidad de fluorescencia de cada grupo de pacientes contra el grupo control. El grupo control utilizado consistió en 48 casos del HapMap Internacional.

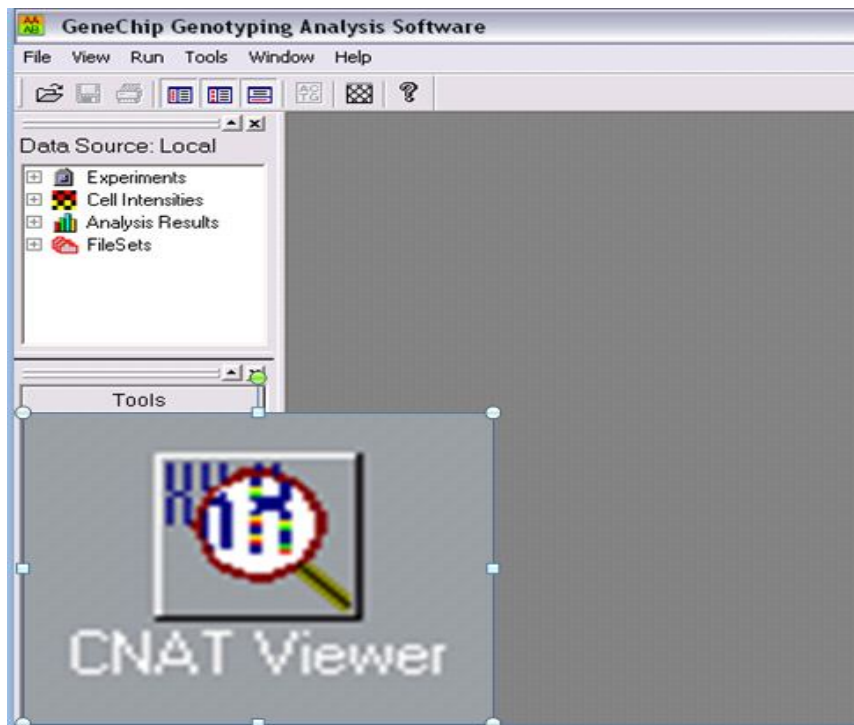


Figura 17A El CNAT es una herramienta para análisis de número de copias y se encuentra dentro de la barra de herramientas en la parte lateral izquierda inferior de la pantalla principal del GTYPE versión 4.1.

Se identificaron todos aquellos SNPs alterados que se compartían en al menos el 40% de los casos por grupo patológico. Así, se obtuvieron 43 regiones en VXX, 28 en MXY y 35 en HVXX. Algunas de estas regiones alteradas en su número de copia, se presentaban en los tres grupos de pacientes lo que hace suponer que estas variaciones son CNVR (por las siglas en inglés de Copy Number Variant Regions) de nuestra población, por lo que es necesario compararlos con un grupo control mexicano y discriminar las regiones verdaderamente asociadas a las patologías de estudio.

En el siguiente paso, identificamos alteraciones exclusivas de cada grupo de pacientes, CNVRs previamente reportadas en la población internacional y regiones codificantes, es decir, identificamos SNPs alterados localizados en genes. Es así, como se aislaron 3 regiones (2 deletadas y una duplicada) asociadas a VXX, 3 regiones asociadas a HV (todas duplicadas), 1 región duplicada asociada a MXY y 3 regiones de pérdidas comunes en VXX y HV (Tabla 9).

Pacientes	Total de CNVR	Ganancias		Perdidas	
		Regiones	CNVR	Regiones	CNVR
MXY	28	8	11	2	7
VXX	43	4	21	7	11
HVXX	35	5	15	8	7

Tabla 9 Cariotipo de los pacientes y la subdivisión de las regiones con variación en el número de copia (CNVR) encontradas.

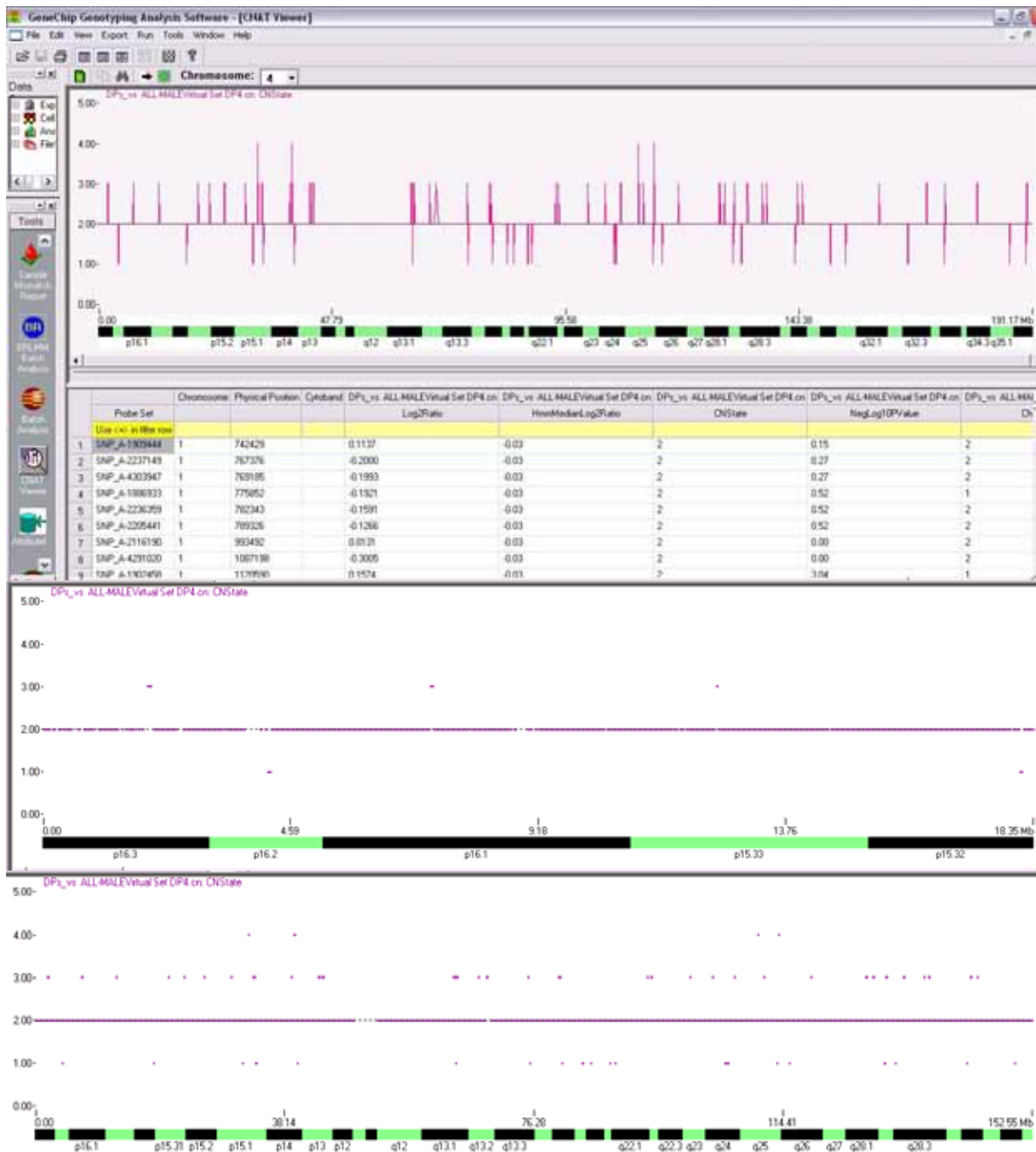


Figura 17B Muestra una grafica de barras en color rosa con barrad gruesas indicando 2 copias y delgadas en más de 2 copias.17C y D Son gráficos de puntos morados en donde se ve que hay pocas ganancias y pérdidas con relación a la mayoría de los SNPs con 2 copias. En la grafica de 17C se aprecia mayor resolución por un zoom más poderoso que en 17D.

Cuando se ha detectaron deleciones (< 2 copias) o amplificaciones (3 copias o más del mismo SNP) se exportaron los datos del individuo con la presencia de alteración hacia el IGB (Integrated Genome Browser) para manejar con facilidad a varios pacientes a la vez en un contexto de la secuencia de genómica que puede incluir los genes relacionados a dichos SNPs, el nucleótido en que se encuentran y la citobanda del cromosoma que se está visualizando. Esta práctica herramienta puede ocuparse con diferentes niveles de zoom para ver todo el cromosoma o bien solamente una región delimitada con gran detalle y así se puede ver el límite entre intrones y exones así como discernir entre promotores y regiones intergénicas, lo que permite precisar si el polimorfismo de un solo nucleótido cae en región codificante o no. Además de esto se muestran los diferentes transcritos alternativos y los exones involucrados en cada uno de ellos. Es recomendable actualizar la información del IGB mensualmente para contar con la más reciente información acerca del genoma humano y el HapMap. Todo esto al igual que las otras herramientas de los programas Affymetrix ocurre bajo un ambiente automatizado compatible con usuarios con conocimientos limitados de programación y bioinformática.

7. DISCUSION

Los estados patológicos en el humano tienen, en su gran mayoría importante influencia multifactorial, en las que se reconoce un componente genético en estrecha relación con el ambiente. Aún en aquellos padecimientos monogénicos, se observa con mayor frecuencia la participación de eventos epigenéticos, genes modificadores así como la relación y la influencia del individuo con su genoma. Por lo que el advenimiento de la medicina genómica y la utilización de herramientas masivas, para estudiar de forma integral el genoma, ha modificado drásticamente la visión de la medicina moderna. Estos estudios permiten conocer mecanismos fisiopatológicos, eventos del desarrollo, asociación y predisposición con todas aquellas enfermedades y condiciones multifactoriales crónico degenerativas. Con la tecnología del arreglo se logran analizar múltiples anomalías y/o variaciones submicroscópicas en el genoma humano en un solo experimento. Un arreglo (*del inglés array*) está conformado por una serie de elementos microscópicos ordenados específicamente en una superficie plana. Los arreglos están estructurados por una superficie sólida (vidrio, plástico o silicio), a la cual se unen una serie de fragmentos de DNA o RNA. Con la finalidad de aplicar este tipo de conocimiento a diversos estados patológicos se realizó la estandarización de microarreglos de DNA para SNPS 500K de Affymetrix en el Servicio de Genética del Hospital General de México. Este proceso se enfocó en dos etapas, la primera en el trabajo experimental realizado en el laboratorio y la segunda en el trabajo de análisis de datos mediante la instalación y aplicación de los diferentes programas de cómputo que existen en

el mercado. El presente estudio se enfocó en el análisis de copias y pérdida de heterocigocidad. La estandarización de análisis se dividió en 2 partes la primera fue la genotipificación de individuos y principal mente en el estudio de la variación del número de copias con el fin de asignar y delimitar regiones cromosómicas de interés.

Múltiples estudios en el genoma de humanos (Feuk L. et al. 2006) y otros mamíferos (Freeman, J. L. et al. 2006) han descrito una gran cantidad de regiones variables en el número de copias, esta variación puede ser desde algunas kilobases (Kb) hasta megabases (Mb) (Sebat, J. et al. 2004). Estas regiones pueden ser deleciones, inserciones o duplicaciones (Fredman, D. et al. 2004) y son denominadas de manera conjunta como CNV, por las siglas en inglés de Copy Number Variation o CNPs por las siglas en inglés de Copy Number Polymorphism. Los CNV alteran la dosis génica y con ello, el fenotipo del individuo (Repping, S. et al. 2006), esto se ve reflejado en desordenes de microdelección y microduplicación. (Inoue, K. & Lupski, J. R. 2002) Estas regiones también se han asociado con riesgo a enfermedades complejas como la infección por HIV-1 (Aitman, T. J. et al. 2006). Los CNV pueden influir en la expresión génica indirectamente por efecto de posición, predisponer a cambios genéticos deletéreos o puede funcionar como sustrato para cambios evolutivos. (Feuk L. et al. 2006).

Con los resultados del HapMap se ha facilitado considerablemente el análisis de los datos obtenidos en este tipo de estudios. Sin embargo, es importante mencionar las desventajas del HapMap como por ejemplo los sujetos utilizados no incluyen muestras de individuos mexicanos, esto es primordial para hacer un correcto análisis de los datos. En esta primera parte del análisis se precisó la importancia de tener un grupo control con las características étnicas de nuestra población que tome en cuenta la variabilidad polimórfica intrínseca de cada región. Primero para hacer las comparaciones correspondientes entre la población Mestizo-mexicana con las poblaciones del HapMap y después para comprar los datos de nuestra población de estudio. El uso del CNAT mostró tanto en los tantos casos que se incluyeron como en el DNA de controles Mestizo-Mexicanos sanos, un gran número de regiones variables en número de copias, entre los diferentes cromosomas y se observaron de igual forma microdeleciones y microduplicaciones. Estos datos confirman la gran variabilidad polimórfica en el humano y nos permiten tener una nueva visión hacia la complejidad del genoma. Se adquirió la experiencia necesaria para poder discriminar entre variantes normales y variantes patológicas. La cantidad de variantes en número de copia es similar a la presente en controles, esto puede reflejar que el procesamiento de datos y el análisis están bien configurados. Consideramos que al menos una mínima parte de estas regiones podría estar relacionada a los fenotipos estudiados y que la mayoría no tendrán significancia clínica. Comenzamos con un gran número de CNV de interés y de ahí fueron consideradas únicamente las regiones que incluían genes.

Este estudio, representa la primera aplicación genómica de la plataforma comercial más utilizada para genotipificación y cuantificación de la variación en número de copias. Estos resultados proveen al Servicio de Genética una plataforma de alta resolución para análisis de número de copias y pérdida de heterocigocidad simultáneamente en pacientes con defectos al nacimiento.

De esta manera se pueden identificar regiones y genes asociadas a diferentes estados patológicos, se realizan diagnósticos precisos, y con esto se facilita el manejo integral de pacientes.

Es importante tener en cuenta que la presencia de cambios en el número de copias no es prueba absoluta de que los genes presentes en esa región estén involucrados en procesos patológicos. Por ello se requiere de una caracterización detallada para elucidar el significado clínico.

8. CONCLUSIONES

Se estandarizó la técnica de GeneChip Mapping 500K Affymetrix en el laboratorio de Genética del Hospital General de México. Se cuenta con toda la infraestructura necesaria para continuar con este tipo de estudios. Es una buena herramienta para poder identificar loci asociados a patología. Permite conocer las variantes polimórficas normales de nuestra población. En nuestro grupo de pacientes se identificaron diferentes regiones con número de copias variables que pueden estar asociados a la patología en estudio. Se identificaron genes candidatos pero es necesario realizar una validación de los datos obtenidos.

9. REFERENCIAS

Aitman T J, Dong R, Vyse TJ, Norsworthy PJ. (2006) Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. Nature 439, 851–855.

DeSalle R., Yudell M. (2005) Welcome to the GENOME. Wiley

Ellis M.C. (2000) –Spot-On- SNP Genotyping. Genome Research 10:895-97

Feuk L, Carson A, Scherer S. (2006) Structural variation in the human genome. Nature Rev. Genet. 7, 85–97.

Fingerlin T.E, Boehnke M, Abecasis G.R. (2004). Increasing the Power and Efficiency of Disease-Marker Case-Control Association Studies through Use of Allele-Sharing Information. Am. J. Hum.Genet. 74:432-43

Fredman D, White SJ, Potter S, Eichler EE, Den Nunnen JT, Brookes AJ. (2004) Complex SNP-related sequence variation in segmental genome duplications. Nature Genet. 36, 861–866.

Freeman J, Perry G, Feuk L, Redon R (2006) Copy number variation: New insights in genome diversity. *Genome Res.* 16, 949–961.

Gram. R.B. (2004) High-Resolution Analysis of DNA Copy Nuber Using Oligonucleotide Microarrays. *Genome Res* 14: 287-95

Inoue K, Lupski JR. (2002) Molecular mechanisms for genomic disorders. *Annu. Rev. Genomics Hum. Genet.* 3, 199–242.

Jobling M.A., Hurles M.E, Tyler-Smith C. (2004) HUMAN EVOLUTIONARY GENETICS: Origins, Peoples & Disease. Garland Science

Jonathan S, Lakshmi B. (2004) Large-Scale Copy Number Polymorphism in the Human Genome. *Science* 305 525-28

Knudsen S. (2004) A GUIDE TO ANALYSIS OF DNA MICROARRAY DATA. Wiley-Liss

Nyholt D.R. (2004) A Simple Correction for Multiple Testing for SNPs in Linkage Disequilibrium with Each Other. *Am. J. Hum.Genet.* 74:765-69

Monpetit A. (2005) A haplotype map of the human genome. *Nature* 437 1299-320.

Ramensky V, Bork P, Sunyaev S. (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Research* 30 17: 3894-900.

Redon R, Ishikawa S, Fitch K.R. (2006) Global variation in copy number in the human genome. *Nature* 444:23 444-53

Repping S, van Daalen SK, Brown LG, Korver CM, Lange J, Marszalek. (2006) High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nature Genet.* 38, 463–467.

Sebat J. (2004) Large-Scale Copy Number Polymorphism in the Human Genome. *Science* 305 525-8

Shena M. (2003) MICROARRAY ANALYSIS. Wiley-Liss

Watson J.D, Caudy A.A, Myers R.M,Witkowski J.A. (2007) Recombinant DNA: GENES AND GENOMES. 3 Edition. CSHL Press

Zhao X, Li C, Paez G. (2004) An Integrated View of Copy Number and Allelic Alterations in the Cancer Genome using SNP Arrays. *Cancer Research* 64:3060-

Zhang K. (2004) Haplotype Block Partitioning and Tag SNP Selection Using Genotype Data and Their Applications to Association Studies. *Genome Res.* 14:908-16

7. DISCUSION

Los estados patológicos en el humano tienen, en su gran mayoría importante influencia multifactorial, en las que se reconoce un componente genético en estrecha relación con el ambiente. Aún en aquellos padecimientos monogénicos, se observa con mayor frecuencia la participación de eventos epigenéticos, genes modificadores así como la relación y la influencia del individuo con su genoma. Por lo que el advenimiento de la medicina genómica y la utilización de herramientas masivas, para estudiar de forma integral el genoma, ha modificado drásticamente la visión de la medicina moderna. Estos estudios permiten conocer mecanismos fisiopatológicos, eventos del desarrollo, asociación y predisposición con todas aquellas enfermedades y condiciones multifactoriales crónico degenerativas. Con la tecnología del arreglo se logran analizar múltiples anomalías y/o variaciones submicroscópicas en el genoma humano en un solo experimento. Un arreglo (*del inglés array*) está conformado por una serie de elementos microscópicos ordenados específicamente en una superficie plana. Los arreglos están estructurados por una superficie sólida (vidrio, plástico o silicio), a la cual se unen una serie de fragmentos de DNA o RNA. Con la finalidad de aplicar este tipo de conocimiento a diversos estados patológicos se realizó la estandarización de microarreglos de DNA para SNPS 500K de Affymetrix en el Servicio de Genética del Hospital General de México. Este proceso se enfocó en dos etapas, la primera en el trabajo experimental realizado en el laboratorio y la segunda en el trabajo de análisis de datos mediante la

instalación y aplicación de los diferentes programas de cómputo que existen en el mercado. El presente estudio se enfocó en el análisis de copias y pérdida de heterocigocidad. La estandarización de análisis se dividió en 2 partes la primera fue la genotipificación de individuos y principal mente en el estudio de la variación del número de copias con el fin de asignar y delimitar regiones cromosómicas de interés.

Múltiples estudios en el genoma de humanos (Feuk L. et al. 2006) y otros mamíferos (Freeman, J. L. et al. 2006) han descrito una gran cantidad de regiones variables en el número de copias, esta variación puede ser desde algunas kilobases (Kb) hasta megabases (Mb) (Sebat, J. et al. 2004). Estas regiones pueden ser deleciones, inserciones o duplicaciones (Fredman, D. et al. 2004) y son denominadas de manera conjunta como CNV, por las siglas en inglés de Copy Number Variation o CNPs por las siglas en inglés de Copy Number Polymorphism. Los CNV alteran la dosis génica y con ello, el fenotipo del individuo (Repping, S. et al. 2006), esto se ve reflejado en desordenes de microdelección y microduplicación. (Inoue, K. & Lupski, J. R. 2002) Estas regiones también se han asociado con riesgo a enfermedades complejas como la infección por HIV-1 (Aitman, T. J. et al. 2006). Los CNV pueden influir en la expresión génica indirectamente por efecto de posición, predisponer a cambios genéticos deletéreos o puede funcionar como sustrato para cambios evolutivos. (Feuk L. et al. 2006).

Con los resultados del HapMap se ha facilitado considerablemente el análisis de los datos obtenidos en este tipo de estudios. Sin embargo, es importante mencionar las desventajas del HapMap como por ejemplo los sujetos utilizados no incluyen muestras de individuos mexicanos, esto es primordial para hacer un correcto análisis de los datos. En esta primera parte del análisis se precisó la importancia de tener un grupo control con las características étnicas de nuestra población que tome en cuenta la variabilidad polimórfica intrínseca de cada región. Primero para hacer las comparaciones correspondientes entre la población Mestizo-mexicana con las poblaciones del HapMap y después para comprar los datos de nuestra población de estudio. El uso del CNAT mostró tanto en los tantos casos que se incluyeron como en el DNA de controles Mestizo-Mexicanos sanos, un gran número de regiones variables en número de copias, entre los diferentes cromosomas y se observaron de igual forma microdeleciones y microduplicaciones. Estos datos confirman la gran variabilidad polimórfica en el humano y nos permiten tener una nueva visión hacia la complejidad del genoma. Se adquirió la experiencia necesaria para poder discriminar entre variantes normales y variantes patológicas. La cantidad de variantes en número de copia es similar a la presente en controles, esto puede reflejar que el procesamiento de datos y el análisis están bien configurados. Consideramos que al menos una mínima parte de estas regiones podría estar relacionada a los fenotipos estudiados y que la mayoría no tendrán significancia clínica. Comenzamos con un gran número de CNV de interés y de ahí fueron consideradas únicamente las regiones que incluían genes.

Este estudio, representa la primera aplicación genómica de la plataforma comercial más utilizada para genotipificación y cuantificación de la variación en número de copias. Estos resultados proveen al Servicio de Genética una plataforma de alta resolución para análisis de número de copias y pérdida de heterocigocidad simultáneamente en pacientes con defectos al nacimiento.

De esta manera se pueden identificar regiones y genes asociadas a diferentes estados patológicos, se realizan diagnósticos precisos, y con esto se facilita el manejo integral de pacientes.

Es importante tener en cuenta que la presencia de cambios en el número de copias no es prueba absoluta de que los genes presentes en esa región estén involucrados en procesos patológicos. Por ello se requiere de una caracterización detallada para elucidar el significado clínico.

8. CONCLUSIONES

Se estandarizó la técnica de GeneChip Mapping 500K Affymetrix en el laboratorio de Genética del Hospital General de México. Se cuenta con toda la infraestructura necesaria para continuar con este tipo de estudios. Es una buena herramienta para poder identificar loci asociados a patología. Permite conocer las variantes polimórficas normales de nuestra población. En nuestro grupo de pacientes se identificaron diferentes regiones con número de copias variables que pueden estar asociados a la patología en estudio. Se identificaron genes candidatos pero es necesario realizar una validación de los datos obtenidos.

9. REFERENCIAS

Aitman T J, Dong R, Vyse TJ, Norsworthy PJ. (2006) Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. Nature 439, 851–855.

DeSalle R., Yudell M. (2005) Welcome to the GENOME. Wiley

Ellis M.C. (2000) –Spot-On- SNP Genotyping. Genome Research 10:895-97

Feuk L, Carson A, Scherer S. (2006) Structural variation in the human genome. Nature Rev. Genet. 7, 85–97.

Fingerlin T.E, Boehnke M, Abecasis G.R. (2004). Increasing the Power and Efficiency of Disease-Marker Case-Control Association Studies through Use of Allele-Sharing Information. Am. J. Hum.Genet. 74:432-43

Fredman D, White SJ, Potter S, Eichler EE, Den Nunnen JT, Brookes AJ. (2004) Complex SNP-related sequence variation in segmental genome duplications. Nature Genet. 36, 861–866.

Freeman J, Perry G, Feuk L, Redon R (2006) Copy number variation: New insights in genome diversity. *Genome Res.* 16, 949–961.

Gram. R.B. (2004) High-Resolution Analysis of DNA Copy Nuber Using Oligonucleotide Microarrays. *Genome Res* 14: 287-95

Inoue K, Lupski JR. (2002) Molecular mechanisms for genomic disorders. *Annu. Rev. Genomics Hum. Genet.* 3, 199–242.

Jobling M.A., Hurles M.E, Tyler-Smith C. (2004) HUMAN EVOLUTIONARY GENETICS: Origins, Peoples & Disease. Garland Science

Jonathan S, Lakshmi B. (2004) Large-Scale Copy Number Polymorphism in the Human Genome. *Science* 305 525-28

Knudsen S. (2004) A GUIDE TO ANALYSIS OF DNA MICROARRAY DATA. Wiley-Liss

Nyholt D.R. (2004) A Simple Correction for Multiple Testing for SNPs in Linkage Disequilibrium with Each Other. *Am. J. Hum.Genet.* 74:765-69

Monpetit A. (2005) A haplotype map of the human genome. *Nature* 437 1299-320.

Ramensky V, Bork P, Sunyaev S. (2002) Human non-synonymous SNPs: server and survey. *Nucleic Acids Research* 30 17: 3894-900.

Redon R, Ishikawa S, Fitch K.R. (2006) Global variation in copy number in the human genome. *Nature* 444:23 444-53

Repping S, van Daalen SK, Brown LG, Korver CM, Lange J, Marszalek. (2006) High mutation rates have driven extensive structural polymorphism among human Y chromosomes. *Nature Genet.* 38, 463–467.

Sebat J. (2004) Large-Scale Copy Number Polymorphism in the Human Genome. *Science* 305 525-8

Shena M. (2003) MICROARRAY ANALYSIS. Wiley-Liss

Watson J.D, Caudy A.A, Myers R.M,Witkowski J.A. (2007) Recombinant DNA: GENES AND GENOMES. 3 Edition. CSHL Press

Zhao X, Li C, Paez G. (2004) An Integrated View of Copy Number and Allelic Alterations in the Cancer Genome using SNP Arrays. *Cancer Research* 64:3060-

Zhang K. (2004) Haplotype Block Partitioning and Tag SNP Selection Using Genotype Data and Their Applications to Association Studies. *Genome Res.* 14:908-16