

ORIGINALES

Artículo bilingüe inglés/español

## Desarrollo y validación de un panel de secuenciación masiva en paralelo para farmacogenética clínica

### Development and validation of a next-generation sequencing panel for clinical pharmacogenetics

Luis Ramudo-Cela<sup>1,2</sup>, Jesús María López-Martí<sup>1</sup>, Daniel Colmeiro-Echeverría<sup>1</sup>, David de-Uña-Iglesias<sup>1,2</sup>, José Luis Santomé-Collazo<sup>1</sup>, Lorenzo Monserrat-Iglesias<sup>1,2,3</sup>.

<sup>1</sup>Health in Code S. L., Departamento científico, A Coruña. España. <sup>2</sup>Instituto de Investigación Biomédica, A Coruña. España. <sup>3</sup>Universidade da Coruña, GRINCAR (Cardiovascular Research Group), A Coruña. España.

#### Autor para correspondencia

Luis Ramudo Cela  
Edificio o Fortín  
Hospital marítimo de Oza,  
As Xuvias s/n  
15006 A Coruña, España.

Correo electrónico:  
[luis.ramudo@healthincode.com](mailto:luis.ramudo@healthincode.com)

DOI: 10.7399/fh.11353

#### Resumen

La rápida implantación clínica de las técnicas de secuenciación masiva en paralelo se debe a su capacidad para secuenciar un gran número de regiones genéticas con un coste menor a las técnicas convencionales. Sin embargo, su uso en el ámbito de la farmacogenética es, todavía, muy escaso.

**Objetivo:** Diseño, desarrollo, implementación y validación de un panel de secuenciación masiva en paralelo de farmacogenética orientado a la práctica clínica.

**Método:** Se desarrolló un panel de sondas de captura híbrida (SureSelect®) para el análisis de las regiones genéticas de interés clínico recopiladas mediante búsqueda bibliográfica. Se empleó la plataforma de secuenciación Illumina HiSeq 1500®. Se desarrolló un algoritmo de análisis bioinformático para la anotación de variantes puntuales, inferencia de haplotipos y determinación de variantes estructurales en los genes de interés. Los resultados obtenidos se validaron con materiales de referencia Coriell® de los repositorios de farmacogenética.

**Resultados:** El panel desarrollado permite el estudio de un total de 12.794 regiones comprendidas en 389 genes. Los resultados de validación mostraron una sensibilidad superior al 99% para variantes puntuales e inserciones y delecciones pequeñas. La imputación de haplotipos fue coherente con los resultados consenso de los materiales de referencia caracte-

#### Abstract

The rapid clinical implementation of next generation sequencing techniques is due to its ability to sequence a large number of genetic regions at lower costs than conventional techniques. However, its use in the field of pharmacogenetics is still very limited.

**Objective:** Design, development, implementation and validation of a clinical pharmacogenetics next-generation sequencing panel.

**Method:** We developed a panel of hybrid capture probes (SureSelect®) for the analysis of the genetic regions of clinical interest collected by literature search and using Illumina HiSeq 1500® sequencing platform. We developed a bioinformatic algorithm for variant annotation, haplotype inference and determination of structural variants in the genes of interest. The results obtained were validated with Coriell® reference material from the pharmacogenetic repositories.

**Results:** The developed panel allows the study of a total of 12,794 regions comprised in 389 genes. Validation results showed a sensitivity greater than 99% for single nucleotide variants and small INDELS. Haplotype imputation was consistent with the consensus results in the characterized reference materials. Furthermore, the developed tool was able to correctly identify different types of CYP2D6 copy number variations as well as a wide variety of HLA-B alleles.

#### PALABRAS CLAVE

Secuenciación de alto rendimiento; Farmacogenética;  
Medicina de precisión; Biología computacional; CYP2D6;  
HLA-B; Variaciones del número de copias de ADN.

#### KEYWORDS

High-throughput nucleotide sequencing; Pharmacogenetics;  
Precision medicine; Computational biology; CYP2D6; HLA-B;  
DNA copy number variations.



rizados. Además, la herramienta desarrollada pudo identificar correctamente diferentes tipos de variaciones de número de copias de CYP2D6, así como una gran variedad de alelos de HLA-B.

**Conclusiones:** Esta tecnología representa una alternativa adecuada para su empleo asistencial con ventajas frente a las técnicas convencionales en su rendimiento de producción y sus capacidades de estudio de genes complejos (CYP2D6, HLA-B).

## Introducción

La farmacogenética es la disciplina que estudia la influencia del genotipo en la respuesta a los medicamentos. La farmacogenética clínica propone estrategias individualizadas para el manejo de los medicamentos en base al genotipo de cada paciente. El fin último de la farmacogenética es la mejora de los resultados en salud de los tratamientos farmacológicos, eficacia y efectos adversos, así como la mejora en relación coste-beneficio; por tanto, se trata de una estrategia para el uso racional del medicamento<sup>1</sup>.

La incorporación de los datos genéticos en los procesos asistenciales debe sustentarse en una buena tecnología de secuenciación de alta calidad, así como en procesos de análisis e interpretación de resultados estandarizados. El desarrollo en los últimos años de las técnicas de secuenciación masiva en paralelo o NGS (del inglés: *next-generation sequencing*) ha supuesto un salto tecnológico disruptivo. Se ha pasado de la capacidad de analizar decenas de variantes a la posibilidad de estudiar decenas de miles de variantes en un único estudio. Su incorporación a los procesos de diagnóstico médico es cada vez más frecuente; sin embargo, en el ámbito de la farmacogenética se continúa empleando tecnologías convencionales como reacción en cadena de la polimerasa cuantitativa, métodos de extensión de primer tales como TaqMan® (Thermofisher®) o MassARRAY® (Sequenom®)<sup>2,3</sup>.

El empleo de estas tecnologías convencionales no permite un estudio en profundidad de todas las regiones genéticas de interés y obliga a estudiar solamente un subconjunto de ellas. Diversas revisiones y estudios comparativos enfocados a las técnicas de farmacogenética han puesto de manifiesto que existen discrepancias en el análisis e interpretación de los resultados entre los diferentes laboratorios y tecnologías. Estas discrepancias se han atribuido a que los diversos laboratorios analizan regiones genéticas diferentes sin incluir todo el espectro de variantes de interés<sup>4,5</sup>. Todo ello, des-

graciadamente, reduce el valor clínico de los estudios, motiva la desconfianza en los resultados, dificultando y retrasando la incorporación de una farmacogenética de más valor en la práctica clínica asistencial.

En este artículo se expone el diseño, desarrollo, implementación y validación de un panel de secuenciación NGS de farmacogenética orientado a la práctica clínica.

## Métodos

### Definición de las regiones genéticas de interés clínico

La selección de las regiones de interés se realizó mediante búsqueda bibliográfica. Una región se consideró clínicamente relevante si su análisis aporta información que pueda modificar la estrategia terapéutica asistencial para un fármaco (por ejemplo, selección del tratamiento, dosificación, seguimiento del paciente). Las fuentes bibliográficas principales empleadas comprenden guías de práctica clínica de sociedades científicas farmacogenéticas, fichas técnicas de agencias reguladoras del medicamento y bases de datos relacionadas con la genética que se detallan en la tabla 1. Las principales plataformas de farmacogenética convencional analizadas fueron Affymetrix DMET® y Agena Bioscience iPLEX®. Ciertas regiones genéticas como CYP2D6 o HLA-B requirieron de un diseño específico. Para CYP2D6 se diseñaron sondas para la captura de todas las regiones codificantes y para las regiones de alta homología (CYP2D7 y CYP2D8) con el objeto de estudiar variaciones en el número de copias o CNV (del inglés: *copy number variation*) y variantes estructurales. Para HLA-B se realizó un diseño de sondas frente a las secuencias de referencia de la base de datos IMGT/HLA<sup>7</sup> similar a la aproximación de Wittig *et al.*<sup>8</sup>.

**Tabla 1.** Fuentes de información

Clasificación	Fuente	Url
<b>Sociedad científica elaboradora de guías de práctica clínica</b>	CPIC	Clinical Pharmacogenetics Implementation Consortium <a href="https://cpicpgx.org/">https://cpicpgx.org/</a>
	DPWG	Dutch Pharmacogenetics Working Group <a href="https://www.knmp.nl/patientenzorg/medicatiebewaking/farmacogenetica/pharmacogenetics-1">https://www.knmp.nl/patientenzorg/medicatiebewaking/farmacogenetica/pharmacogenetics-1</a>
	CPNDS	Canadian Pharmacogenetics Group for Drug Safety <a href="http://cpnds.ubc.ca/">http://cpnds.ubc.ca/</a>
<b>Agencias reguladoras del medicamento</b>	FDA	U. S. Food and Drug Administration <a href="https://www.fda.gov/">https://www.fda.gov/</a>
	EMA	European Medicines Agency <a href="https://www.ema.europa.eu/en">https://www.ema.europa.eu/en</a>
	AEMPS	Agencia Española de Medicamentos y Productos Sanitarios <a href="https://www.aemps.gob.es/">https://www.aemps.gob.es/</a>
<b>Bases de datos relacionadas con farmacogenética, genética médica y genética de poblaciones</b>	PharmGKB	Pharmacogenomics Knowledge Base <a href="https://www.pharmgkb.org/">https://www.pharmgkb.org/</a>
	PharmVar	Pharmacogene Variation Consortium <a href="https://www.pharmvar.org/">https://www.pharmvar.org/</a>
	PharmaADME	<a href="http://www.pharmaadme.org/joomla/">http://www.pharmaadme.org/joomla/</a>
<b>Predictores bioinformáticos de los efectos de una mutación sobre una proteína</b>	ClinVar	<a href="https://www.ncbi.nlm.nih.gov/clinvar/">https://www.ncbi.nlm.nih.gov/clinvar/</a>
	dbSNP	Single Nucleotide Polymorphism Database <a href="https://www.ncbi.nlm.nih.gov/snp/">https://www.ncbi.nlm.nih.gov/snp/</a>
	gnomAD	Genome Aggregation Database <a href="https://gnomad.broadinstitute.org/">https://gnomad.broadinstitute.org/</a>
	EVS	Exome Variant Server <a href="http://evs.gs.washington.edu/EVS/">http://evs.gs.washington.edu/EVS/</a>
	SIFT	Sorting Intolerant From Tolerant <a href="https://sift.bii.a-star.edu.sg/">https://sift.bii.a-star.edu.sg/</a>
	PMut	Pathogenic Mutation Prediction <a href="http://mmbpcb.ub.es/PMut/">http://mmbpcb.ub.es/PMut/</a>
	CADD	Combined Annotation Dependent Depletion <a href="https://cadd.gs.washington.edu/">https://cadd.gs.washington.edu/</a>

## Diseño de las sondas de captura y proceso de secuenciación

Se diseñaron sondas de captura híbrida para dichas regiones genéticas con software SureSelect® Design Tool (Agilent). Las muestras se han sometido a un proceso de extracción de ADN genómico automatizada (Qiasymphony SP®, Qiagen®). La preparación del ADN genómico se ha realizado mediante el protocolo SureSelect XT® (Agilent Technologies®), en combinación con un panel de sondas SureSelect Custom Target Enrichment®, que captura selectivamente las regiones genómicas de interés. La secuencia se realizó con equipos Illumina HiSeq 1500®. Todos los procedimientos se han llevado a cabo siguiendo las especificaciones de los fabricantes<sup>9,11</sup>.

## Análisis e interpretación de datos

Se empleó un algoritmo de análisis bioinformático de desarrollo propio que permite la anotación de variantes de tipo SNP (del inglés: *single nucleotide polymorphism*), INDEL (inserciones-delecciones) y variantes estructurales o CNV, estas últimas mediante el análisis comparativo de la profundidad de lecturas<sup>12,14</sup>. Las variantes identificadas se combinan con el sistema de gestión del conocimiento de desarrollo propio que asocia a cada variante valores específicos de calidad e información, datos estructurales y de SNP como dbSNP, 1000 genomas<sup>15</sup> EVS, así como datos de frecuencia de los alelos en distintas poblaciones, o valores de predictores bioinformáticos como SIFT, PMut, CADD (Tabla 1)<sup>12,14</sup>. A partir de las fuentes de información (Tabla 1), especialmente PharmVar y PharmGKB, se desarrollaron tablas en el sistema de gestión del conocimiento para realizar una inferencia automatizada de haplotipos basadas en datos de variantes tipo SNP y pequeñas INDEL. La inferencia de haplotipos de CYP2D6 se realiza con un módulo que filtra los alineamientos procedentes de CYP2D7 y CYP2D8; este filtrado tiene el objeto de evitar artefactos en las gráficas de cobertura procedentes de las lecturas en regiones homólogas<sup>16</sup>. Además, este módulo normaliza los datos de profundidad de lecturas con una muestra control con un número de copias CYP2D6 conocido y secuenciada en la misma placa de análisis. Esta normalización tiene el objeto de evitar artefactos debido a la presencia de CNV en otras muestras secuenciadas en la misma placa<sup>17</sup>. La inferencia de alelos HLA-B se realiza con un módulo específico basado en el desarrollo de Wittig *et al.*<sup>8</sup>.

## Validación analítica de los resultados con materiales de referencia

La determinación de SNP y pequeñas INDEL (< 20 pb) se validó con la muestra Coriell® NA12878. La muestra se procesó por triplicado y los resultados obtenidos se compararon con los datos de referencia disponibles para este material resultantes de la integración de varios conjuntos de datos procedentes de la secuenciación del genoma completo en diferentes plataformas de secuenciación masiva<sup>18</sup>. Mediante este análisis se obtuvieron los valores medianos de sensibilidad (Se) y especificidad (Sp) analíticas, así como los valores predictivos positivo (VPP) y negativo (VPN). Se realizó este análisis con tres filtros diferentes de calidad usando dos parámetros de calidad NGS "Qual" (del inglés: *quality score*) y "DP" (del inglés: *sequencing depth* o profundidad de lecturas): alta [Qual > 49 & DP > 29], intermedia [Qual > 49 & DP > 14], baja [Qual > 0 & DP > 0]<sup>9</sup>.

La determinación de haplotipos farmacogenéticos basados en SNP y pequeñas INDEL se validó con la muestra Coriell® NA12878. Los resultados obtenidos se compararon con los datos de referencia para este material resultantes de la integración de varios conjuntos de datos generados con distintas plataformas de farmacogenética (proyecto GeT-RM)<sup>20,21</sup>. Estos datos incluyen los siguientes genes: CYP1A1, CYP1A2, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2E1, CYP3A4, CYP3A5, CYP4F2, DPYD, GSTM1, GSTP1, GSTT1, NAT1, NAT2, SLC15A2, SLC22A2, SLCO1B1, SLCO2B1, TPMT, UGT1A1, UGT2B7, UGT2B15, UGT2B17, VKORC1. HLA-B y CYP2D6, por su complejidad, se incluyeron un mayor número de probandos, las muestras Coriell® NA12878, NA02016, NA17254 y NA17281, que incluyen distintos tipos de variantes estructurales de CYP2D6 y alelos HLA-B de interés clínico (HLA-B\*58:01)<sup>8,20,21</sup>. Mediante este análisis se obtuvo la concordancia entre los haplotipos determinados y los haplotipos de referencia.

## Resultados

### Regiones genéticas de interés incluidas en el panel

El panel desarrollado incluye un total de 12.794 pares de bases comprendidas en 389 genes. En la tabla suplementaria 1 se incluyen el tamaño de la región en pares de bases que se estudia para cada gen y la cobertura promedio de la región. Estas regiones se clasificaron en tres grupos: *genes principales*, *secundarios* y *candidatos*, siguiendo los criterios internos que se detallan a continuación:

- **Genes principales** (17): CYP2C19, CYP2C9, CYP2D6, CYP3A5, CYP4F2, DPYD, F5, G6PD, HLA-B, IFNL3, RARG, SLC28A3, SLC01B1, TPMT, UGT1A1, UGT1A6, VKORC1. Son genes descritos en las guías de práctica clínica elaboradas por los principales consorcios clínicos de farmacogenética (Tabla 1). Por lo tanto, cuentan con los niveles de evidencia más altos y para todos ellos se han publicado estrategias de dosificación personalizada en base al genotipo.
- **Genes secundarios** (50): ABCB1, ABCC2, ABCG2, ACE, ADH1A, ADH1B, ADH1C, ADRB1, ADRB2, AHR, ADH1A1, ALOX5, ASS1, COMT, CPS1, CYP1A1, CYP1A2, CYP2A6, CYP2C8, CYP2E1, CYP2J2, CYP3A4, DRD2, GSTM1, GSTP1, GSTT1, HMGCR, KCNJ11, MTHFR, NAT1, NAT2, NQO1, NR1I2, P2RY1, P2RY12, POLG, PTGIS, SCN1A, SLC15A2, SLC19A1, SLC22A1, SLC22A6, SLC6A4, SLCO1B3, SUIT1A1, TYMS, UGT2B15, UGT2B17, UGT2B7, VDR. Los genes secundarios, a pesar de no estar incluidos en guías de práctica clínica, consideramos que pueden tener interés clínico según las fuentes de datos consultadas (Tabla 1). En este grupo destacan los listados de biomarcadores farmacogenéticos de la Food and Drug Administration, genes PharmGKB con niveles de evidencia de 2 o superior, categorías A y B de prioridad CPIC o genes core del listado pharmaADME. Aunque no se han publicado estrategias para la individualización del tratamiento en base al genotipo, consideramos que la presencia de variantes en estos genes puede ser de utilidad para la evaluación de la respuesta al tratamiento en pacientes individuales tras un hallazgo clínico.
- **Genes candidatos:** corresponde a regiones genómicas incluidas en alguna de las principales plataformas farmacogenómicas empleadas habitualmente en la práctica clínica<sup>22,23</sup>, o regiones descritas clasificadas en las fuentes de información empleadas (Tabla 1) con niveles de evidencia inferiores a los otros grupos, por ejemplo, PharmGKB nivel 3 o 4, CPIC C/D. Su estudio se contempla en un contexto de investigación. El listado completo de genes y regiones se encuentra en el suplemento del artículo.

### Validación de la determinación de SNP y pequeñas INDEL

Los resultados de la determinación de variantes genéticas de tipo sustituciones puntuales (SNP) y pequeñas inserciones/delecciones (INDEL) ≤ 20 pb se muestran en la tabla 2. La sensibilidad y especificidad analíticas del

**Tabla 2.** Validación de la determinación de variantes genéticas de tipo sustituciones puntuales (SNP) y pequeñas inserciones/delecciones (INDEL) < 20 pb

Calidad de secuencia	Alta	Media	Baja
<b>Verdaderos positivos (VP)</b>	553	572	573
<b>Falsos negativos (FN)</b>	0	0	4
<b>Falsos positivos (FP)</b>	0	0	0
<b>Verdaderos negativos (VN)</b>	12.241	12.222	12.217
<b>Valor predictivo positivo (VPP)</b>	> 99,99%	> 99,99%	> 99,99%
<b>Valor predictivo negativo (VPN)</b>	> 99,99%	> 99,99%	99,97%
<b>Sensibilidad (Se)</b>	> 99,99%	> 99,99%	99,31%
<b>Especificidad (Sp)</b>	> 99,99%	> 99,99%	> 99,99%

La muestra NA12878 se procesó por triplicado y los resultados se compararon con los datos de referencia generados en diferentes plataformas de secuenciación masiva<sup>18</sup>; se muestra la mediana de los valores para las tres ejecuciones. Los filtros de calidad de secuencia se definieron como alta [Qual > 49 & DP > 29], intermedia [Qual > 49 & DP > 14] o baja [Qual > 0 & DP > 0].

ensayo fueron superiores al 99%. Por otro lado, los VPP y VPN fueron superiores también al 99% para todos los valores de calidad analizados.

## Validación de la determinación de alelos farmacogenéticos (haplotipos)

Los resultados de la determinación de haplotipos farmacogenéticos con la muestra Coriell® NA12878 se muestra en la tabla 3. Se dispone de datos de referencia para 28 genes (cada uno con 2 haplotipos, 56 haplotipos)<sup>20,21</sup>; los resultados fueron diferentes en cuatro genes: CYP2D6, NAT2, SLCO1B1 y UGT2B17. Para los genes CYP2D6, NAT2 y SLCO1B1 las diferencias fueron debidas a que la plataforma NGS desarrollada posee un número mayor de regiones genéticas y por lo tanto supera las capacidades de detección de las plataformas empleadas en el proyecto GeT-RM<sup>20,21</sup>. Para CYP2D6, la plataforma NGS desarrollada detectó la presencia de un reordenamiento de tipo híbrido D6/D7 en tandem, descrito en publicaciones anteriores<sup>24,25</sup>. Para NAT2 la técnica fue capaz de detectar el alelo \*12, este alelo sólo está incluido en la plataforma Agena Bioscience iPLEX

ADME PGx Pro®. Para SLCO1B1 nuestra técnica identificó dos combinaciones posibles de haplotipos \*1A/\*15 y \*1B/\*5; este resultado es similar al que determina la plataforma Affimetrix DMET® y se debe a la inclusión de los subalelos \*1A y \*1B. Para el caso de UGT2B17 nuestra técnica no detectó la presencia del haplotipo \*2, que corresponde con una delección del gen. Este requisito no se incluyó en las especificaciones del diseño. Teniendo esto en cuenta, la concordancia general en la determinación de alelos es superior al 98%. Si el análisis se restringe a aquellos genes incluidos en las principales guías de práctica clínica (Tabla 1), la concordancia se eleva hasta el 100%.

## Validación de los resultados de la identificación de variantes estructurales en CYP2D6

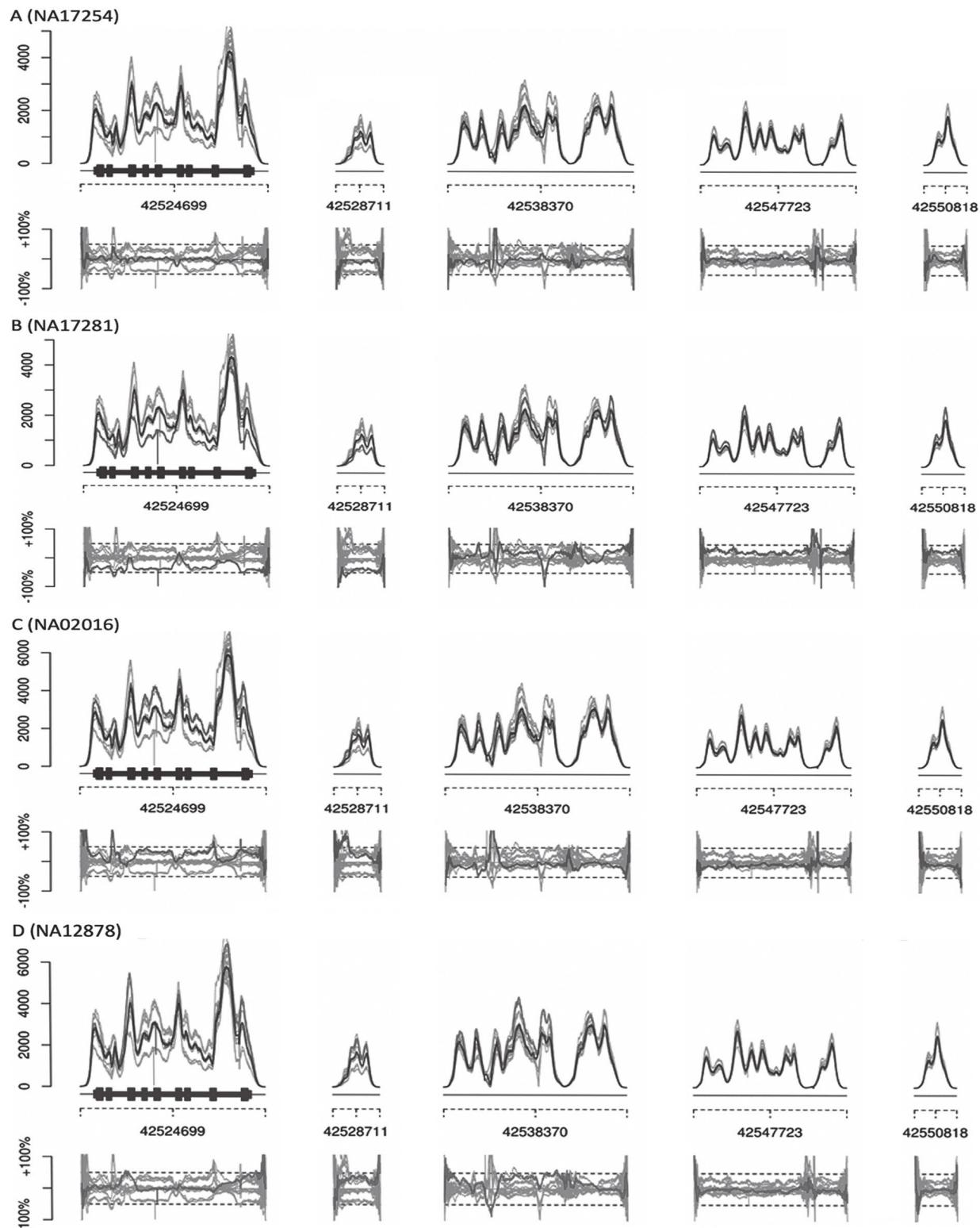
La técnica basada en el estudio comparativo de profundidad de lecturas desarrollada fue capaz de identificar adecuadamente las variantes estructurales presentes en las cuatro muestras control seleccionadas. Los resultados del análisis gráfico de coberturas para CYP2D6 se muestran en la figura 1.

**Tabla 3.** Validación de la determinación de alelos farmacogenéticos (haplotipos)

Gen	Alelos de referencia	Alelos identificados por la técnica
CYP1A1	*1/*1	*1/*1
CYP1A2	*1F/*1F	*1F/*1F
CYP2A6	*1/*1	*1/*1
CYP2B6	*1/*1	*1/*1
CYP2C8	*1/*3	*1/*3
CYP2C9	*1/*2	*1/*2
CYP2C19	*1/*2	*1/*2
CYP2D6	*3/*4	*3/*4 + híbrido D6/D7
CYP2E1 [1]	*5/*7	*5/*7
CYP3A4	*1/*1	*1/*1
CYP3A5	*3/*3	*3/*3
CYP4F2	*1/*1	*1/*1
DPYD	*1/*4	*1/*4
GSTM1	*B/*B	*B/*B
GSTP1 [2]	*A/*C,*B/*D	*A/*C,*B/*D
GSTT1	*A/*B	*A/*B
NAT1	*4/*4	*4/*4
NAT2	*4/*5	*12/*5
SLC15A2	*1/*1	*1/*1
SLC22A2	*1/*1	*1/*1
SLCO1B1	*1/*15	*1A/*15 o *1B/*5
SLCO2B1	WT / WT	*1/*1
TPMT	*1/*1	*1/*1
UGT1A1	*1/*28	*1/*28
UGT2B7	*1/*2	*1/*2
UGT2B15 [1]	*4/*4	*4/*4
UGT2B17	*2/*2	*1/*1
VKORC1 c.-1639G>A (rs9923231)	GA	GA

[1] Los resultados GeTRM para la muestra NA12878 y los genes CYP2E1 y UGT2B15 indican "no consenso". Esto se debe a las diferencias en las dos técnicas empleadas: Affimetrix DMET® y Agena Bioscience iPLEX®. El estudio iPLEX® no es capaz de detectar el alelo CYP2E1\*4 ni el UGT2B15\*5. Por lo tanto, en la tabla se incluyen los resultados de Affimetrix DMET®.

[2] Para GSTP1 el resultado consenso es ambiguo, esto es, se establecen dos combinaciones de alelos posibles para una combinación de variantes determinada.

**Figura 1.** Gráficas de profundidad de lecturas y profundidad de lecturas normalizada.

Las líneas grises intermedias representan la profundidad de lecturas de muestra analizada, las líneas grises más oscuras representan la profundidad de lecturas de la muestra control frente a la que se normaliza la profundidad de lecturas, las líneas grises más claras representan el resto de las muestras analizadas en la misma placa. La profundidad de lecturas se representa en las regiones genómicas de CYP2D6 y sus pseudogenes CYP2D7 y CYP2D8. Las muestras analizadas son (A) NA17254 con genotipo CYP2D6\*4/\*41 coherente con dos copias de CYP2D6, (B) NA17281 con genotipo CYP2D6\*5/\*9 coherente con una delección en uno de los alelos, (C) NA02016 con genotipo CYP2D6\*2x2/\*17 coherente con una duplicación en uno de los alelos y (D) NA12878\*3/\*4+68 coherente con un híbrido D6/D7 en tandem.

**Tabla 4.** Resultados de la identificación de alelos HLA-B

Muestra Coriell NA	Alelos consenso	Alelos identificados por la técnica
NA17281	*39:06/*56:01	*39:06:02/*56:01:01
NA17254	*08:01/*35:01	*08:01:01/*35:01:01
NA02016	*53:01/*58:01	*53:01:01/*58:01:01
NA12878	*08:01/*56:01	*08:01:01/*56:01:01

La muestra NA17254 con el genotipo CYP2D6\*4/\*41 presenta dos alelos de CYP2D6, se trata de la muestra control frente a la que se normaliza la profundidad de lecturas de CYP2D6. NA17281 con el genotipo CYP2D6\*5/\*9 presenta una delección de uno de los alelos CYP2D6, la profundidad de lecturas normalizada tuvo decremento de aproximadamente un 50%. NA02016 con el genotipo CYP2D6\*2x2/\*17 tiene una duplicación de uno de los alelos, tuvo un incremento del 50% en la profundidad de lecturas. La muestra NA12878 con el genotipo\*3/\*4+68 tiene un híbrido D6/D7 en tandem, tuvo un incremento de 50% en la profundidad de lecturas en una parte de la región de CYP2D6 y una parte de la región de CYP2D7.

## Validación de los resultados de la identificación de variantes en HLA

La técnica desarrollada para la imputación de los alelos en la región de HLA-B fue capaz de identificar adecuadamente los alelos presentes en las tres muestras control seleccionadas. La tabla 4 muestra los alelos consenso establecidos y los resultados obtenidos por nuestra técnica.

## Discusión

En este artículo se describe el desarrollo y la validación analítica de una plataforma de secuenciación NGS para la farmacogenética clínica, así como los resultados obtenidos con las herramientas bioinformáticas empleadas en la interpretación de los datos.

El panel desarrollado cubre un total de 12.794 pares de bases comprendidas en 389 genes. Este tamaño es muy superior al de las plataformas convencionales más exhaustivas que se emplean habitualmente: Affymetrix DMET® (1.936 marcadores), Agena Bioscience® (Sequenom®, iPLEX ADME PGx Pro® (270 marcadores), Illumina VeraCode ADME® (184 marcadores) y Roche AmpliChip® (36 alelos CYP2D6 y CYP2C19)<sup>22,23</sup>.

Los resultados obtenidos en las muestras control fueron coherentes con los resultados de referencia del proyecto GeT-RM. Las diferencias obtenidas se deben a una capacidad superior de nuestra plataforma, salvo en el caso de UGT2B7. Para este gen se han descrito variantes estructurales de tipo delección que requería un diseño específico en nuestra plataforma. Dicho diseño se descartó en las especificaciones, ya que no ha sido incluido en ninguna de las principales guías de práctica clínica farmacogenéticas.

Ciertas regiones genómicas de gran interés farmacogenético como CYP2D6 y HLA-B son complejas de estudiar. CYP2D6 es responsable del metabolismo de aproximadamente un 25% de los fármacos comercializados<sup>26</sup> y su genotipo es esencial para individualizar tratamientos antidepresivos como paroxetina<sup>27</sup>, antipsicóticos como aripiprazol<sup>28</sup>, tamoxifeno<sup>29</sup>, antieméticos como ondansetrón<sup>30</sup> o analgésico como codeína<sup>31</sup>, entre otros. La dificultad para su análisis genético radica en la existencia de dos regiones con muy alta homología, CYP2D7 y CYP2D8, y a la existencia de variantes tipo CNV y reordenamientos con los pseudogenes (híbridos de tipo conversión y en tandem con alelos funcionales). Se han publicado otras herramientas para el análisis bioinformático de CYP2D6 que obtienen buenos resultados en el análisis de variantes tipo SNP o pequeñas INDEL, pero con limitaciones en la determinación de variantes estructurales, especialmente híbridos<sup>16,17,24,32,33</sup>. Para este fin, tal y como se expone en la sección de materiales y métodos, la técnica NGS desarrollada realiza un filtrado de lecturas procedentes de las regiones homólogas y normaliza la profundidad de lecturas frente a una muestra control conocida. La plataforma NGS desarrollada ha identificado de manera adecuada diferentes

tipos de variantes estructurales y CNV analizados en materiales de referencia y proporciona unas gráficas con la profundidad de lecturas que muestran el comportamiento de la cobertura a lo largo de toda la región genómica (Figura 1).

La región HLA (antígeno leucocitario humano) es la región más densamente polimórfica del genoma. Su importancia farmacogenética radica en su asociación con las reacciones de hipersensibilidad a fármacos inmunomediatas<sup>34</sup>. Las técnicas de tipado convencionales permiten solamente el análisis de unos pocos haplotipos, mientras que los procedimientos generales para el estudio de dicha región mediante tecnologías NGS no obtienen buenos resultados debido a la variabilidad y la homología de secuencia entre los genes y los pseudogenes<sup>35</sup>. El presente desarrollo NGS incorpora un diseño de sondas específico y un software de análisis previamente validado que permite un estudio global de HLA-B.

Las principales limitaciones de la plataforma NGS en comparación con las plataformas convencionales incluyen el tiempo de estudio, la cualificación del personal técnico necesario, los costes de desarrollo y puesta a punto, así como los costes de procesamiento informático y almacenamiento de la información en servidores<sup>36</sup>. El tiempo necesario para obtener el resultado desde que se recibe una muestra es de aproximadamente 10 días (incluyendo preparación de la muestra, secuenciación y análisis bioinformático). Además del equipamiento analítico y servidores para procesado de datos, estos procesos requieren de personal técnico con formación específica y personal bioinformático para la puesta a punto de la técnica y el análisis de los datos de secuenciación.

Aunque la técnica y el proceso de datos bioinformático desarrollado identifican variantes raras y otras variantes no descritas previamente, estos resultados no se trasladan a la práctica clínica asistencial, ya que no se dispone de datos en la literatura para su interpretación. Se ha descrito que más del 80% de la variación en los genes de interés en farmacogenética se debe a la presencia de estas variantes raras y las plataformas convencionales no están orientadas a su análisis<sup>37,38</sup>. Este tipo de variantes podrán incorporarse a la práctica clínica asistencial a medida que se disponga de protocolos y consensos para su interpretación, así como de publicaciones científicas con datos de correlación genotipo-fenotipo<sup>39</sup>.

Por último, para demostrar el valor clínico de esta plataforma, es conveniente su evaluación y validación en estudios clínicos prospectivos.

## Financiación

Luis Ramudo Cela es personal cofinanciado por la Xunta de Galicia según la Resolución da Axencia Galega de Innovación del 2 de junio de 2016 (Programa Principia).

## Agradecimientos

Al personal de Health in Code S. L.

## Conflictivo de intereses

Sin conflictos de intereses.

## Aportación a la literatura científica

El presente trabajo aborda el desarrollo, implementación y validación de un panel de secuenciación masiva en paralelo para la farmacogenética clínica.

**Tabla suplementaria 1.** Listado de regiones genómicas, tamaño y cobertura promedio para cada gen

Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura	Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura
A1BG	19	1	1	392,88	ATIC	2	10	10	606,15
ABCB1	7	41	41	756,22	ATP7A	X	7	7	656,25
ABCB11	2	32	34	712,94	ATP7B	13	10	10	582,63
ABCB4	7	27	27	784,39	BCHE	3	11	11	644,37
ABCB7	X	6	6	530,52	BDKRB1	14	1	1	652,75
ABCC1	16	22	25	631,36	BDKRB2	14	2	2	394,44
ABCC2	10	45	45	924,36	C16orf54	16	1	1	512,13
ABCC3	17	6	6	613,77	C6orf10	6	1	1	460,00
ABCC4	13	17	25	646,23	C8orf34	8	1	1	709,25
ABCC5	3	9	9	700,81	CACNA1C	12	1	1	456,25
ABCC6	16	7	7	442,11	CACNA1S	1	2	2	541,63
ABCC8	11	7	8	543,95	CACNA2D1	7	1	1	395,38
ABCC9	12	6	6	570,61	CACNB2	10	2	2	488,32
ABCG1	21	7	7	690,41	CACNG2	22	1	1	987,63
ABCG2	4	11	11	695,70	CALU	7	3	3	595,33
ACE	17	9	9	551,54	CBR1	21	6	6	767,71
ADD1	4	1	1	578,25	CBR3	21	6	6	635,67
ADH1A	4	7	7	483,02	CCDC170	6	1	1	608,00
ADH1B	4	6	6	865,02	CCDC174	3	7	7	654,90
ADH1C	4	5	5	739,55	CCDC33	15	7	7	630,07
ADH4	4	6	6	667,13	CCHCR1	6	2	2	504,94
ADH5	4	5	5	460,75	CDA	1	7	7	554,68
ADH6	4	6	6	540,27	CDSN	6	1	1	0,38
ADH7	4	5	5	723,88	CES1	16	13	13	737,25
ADORA2A	22	3	3	288,09	CES2	16	7	7	478,79
ADRA2A	10	1	1	194,25	CETP	16	8	8	607,27
ADRB1	10	2	2	166,63	CFTR	7	19	24	994,56
ADRB2	5	9	9	799,66	CHRNA3	15	3	3	531,46
AGTR1	3	2	2	599,69	CHRNA4	20	2	2	476,76
AHR	7	6	6	369,71	CHST1	11	6	6	607,48
AKAP9	7	3	3	648,50	CHST10	2	14	14	890,26
ALB	4	7	7	593,29	CHST11	12	7	7	684,70
ALDH1A1	9	6	76	721,56	CHST13	3	6	6	612,98
ALDH2	12	6	6	592,00	CHST2	3	5	5	537,85
ALDH3A1	17	5	5	379,78	CHST3	10	16	16	805,07
ALDH3A2	17	7	7	757,22	CHST4	16	3	3	442,42
ALOX5	10	1	1	570,63	CHST5	16	7	7	577,66
AMPD1	1	1	1	372,13	CHST6	16	3	3	725,29
ANKK1	11	5	5	421,98	CHST7	X	5	5	409,53
AOC1	7	7	7	475,68	CHST8	19	3	3	523,96
AOX1	2	7	7	777,09	CHST9	18	5	5	686,73
APOA2	1	1	1	555,25	COL22A1	8	1	1	556,25
APOB	2	1	1	453,63	COMT	22	8	8	524,55
APOC3	11	3	3	544,67	COQ2	4	2	2	476,07
APOE	19	2	2	166,57	CPS1	2	1	1	860,25
ARG1	6	1	1	759,75	CRHR1	17	30	30	618,46
ARNT	1	7	7	618,32	CRHR2	7	6	6	581,52
ARSA	22	5	5	400,48	CRIP3	6	7	7	643,73
ASS1	9	2	2	718,32	CROT	7	4	4	597,88

**Tabla suplementaria 1 (cont.).** Listado de regiones genómicas, tamaño y cobertura promedio para cada gen

Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura	Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura
CRYAB	11	1	1	257,13	DPYD	1	18	31	834,48
CYB5R3	22	4	4	336,35	DRD1	5	1	1	232,75
CYP11B1	8	17	17	641,75	DRD2	11	25	25	584,49
CYP11B2	8	5	5	568,03	DRD3	3	2	2	523,44
CYP17A1	10	6	6	881,71	DRD4	11	1	1	100,50
CYP19A1	15	11	11	840,04	DSG2	18	1	1	421,50
CYP1A1	15	13	13	882,79	DTNBP1	6	2	2	468,82
CYP1A2	15	17	17	671,07	DYNC2H1	11	1	1	614,88
CYP1B1	2	21	106	886,47	DYRK1A	21	1	1	537,00
CYP20A1	2	5	5	661,43	EDN1	6	1	1	928,50
CYP21A2	6	3	3	280,46	EGF	4	8	8	557,36
CYP24A1	20	8	10	933,41	EPHX1	1	13	13	781,79
CYP26A1	10	5	5	560,98	EPHX2	8	5	8	523,70
CYP26C1	10	1	1	469,50	ERCC1	19	3	3	488,42
CYP27A1	2	5	5	710,73	ERCC6L2	9	1	1	649,75
CYP27B1	12	9	9	743,06	ESR2	14	1	1	287,38
CYP2A13	19	5	5	701,98	F2	11	1	1	451,25
CYP2A6	19	20	26	826,22	F5	1	3	3	590,88
CYP2A7	19	12	12	813,69	F7	13	1	1	390,38
CYP2B6	19	23	23	846,89	FAAH	1	2	2	585,19
CYP2C18	10	7	7	672,18	FCGR3A	1	1	1	408,88
CYP2C19	10	21	21	883,92	FDPS	1	1	1	527,88
CYP2C8	10	12	23	781,26	FKBP5	6	5	5	444,65
CYP2C9	10	20	32	970,51	FLOT1	6	2	2	341,75
CYP2D6	22	115	4787	930,94	FMO1	1	28	43	797,37
CYP2E1	10	10	10	593,28	FMO3	1	31	31	1009,32
CYP2F1	19	5	5	549,90	FMO4	1	6	6	790,82
CYP2J2	1	9	9	767,27	FMO5	1	7	11	756,06
CYP2S1	19	4	4	646,32	FTO	16	1	1	854,50
CYP39A1	6	4	4	631,63	G6PD	X	204	255	1164,76
CYP3A4	7	26	26	766,96	GABRA6	5	1	1	478,75
CYP3A43	7	7	7	554,39	GALC	14	1	1	555,50
CYP3A5	7	16	16	846,99	GGCX	2	5	13	561,63
CYP3A7	7	7	7	968,25	GLCCI1	7	1	1	456,00
CYP46A1	14	1	1	529,00	GNB3	12	2	2	432,13
CYP4A11	1	4	4	497,44	GOLGA7	8	1	21	35,00
CYP4B1	1	7	9	1038,11	GP1BA	17	2	2	185,32
CYP4F11	19	4	4	509,25	GRIK4	11	2	2	668,63
CYP4F12	19	6	6	709,50	GRK4	4	3	3	802,59
CYP4F2	19	11	11	579,92	GRK5	10	1	1	0,00
CYP4F3	19	14	14	584,74	GSTA1	6	12	12	772,23
CYP4Z1	1	4	4	433,35	GSTA2	6	7	7	783,13
CYP51A1	7	7	7	468,00	GSTA3	6	7	7	826,06
CYP7A1	8	7	7	455,56	GSTA4	6	7	7	676,77
CYP7B1	8	4	4	495,25	GSTA5	6	7	7	670,57
CYP8B1	3	7	7	563,56	GSTM1	1	3	3	115,50
DBH	9	1	1	412,38	GSTM2	1	5	5	837,38
DCK	4	4	4	554,01	GSTM3	1	6	6	730,84
DDTL	22	2	2	68,13	GSTM4	1	3	3	1207,38

**Tabla suplementaria 1 (cont.).** Listado de regiones genómicas, tamaño y cobertura promedio para cada gen

Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura	Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura
<b>GSTM5</b>	1	5	5	712,23	<b>NAT1</b>	8	21	29	942,85
<b>GSTO1</b>	10	5	5	748,93	<b>NAT2</b>	8	16	16	1031,74
<b>GSTP1</b>	11	5	5	433,15	<b>NEDD4L</b>	18	5	5	602,75
<b>GSTT1</b>	22	6	6	386,80	<b>NELFCD</b>	20	1	1	753,25
<b>GSTZ1</b>	14	4	4	544,97	<b>NNMT</b>	11	7	8	816,15
<b>HAS3</b>	16	1	1	271,75	<b>NOS3</b>	7	2	2	456,26
<b>HLA-A</b>	6	1	1	931,25	<b>NPPA</b>	1	1	1	760,88
<b>HLA-B</b>	6	53	3743	1061,70	<b>NQO1</b>	16	7	7	669,34
<b>HLA-C</b>	6	1	1	502,38	<b>NR1I2</b>	3	16	16	778,78
<b>HLA-DQA1</b>	6	1	1	428,75	<b>NR1I3</b>	1	6	6	803,19
<b>HLA-DRB1</b>	6	1	1	320,63	<b>NR3C1</b>	5	12	12	702,65
<b>HLA-F</b>	6	1	1	943,00	<b>NT5C2</b>	10	2	2	474,94
<b>HMGCR</b>	5	9	9	480,38	<b>NUDT15</b>	13	2	2	667,82
<b>HNMT</b>	2	6	10	453,82	<b>OPRD1</b>	1	1	1	493,38
<b>HTR1A</b>	5	5	5	505,13	<b>OPRK1</b>	8	2	2	241,88
<b>HTR2A</b>	13	10	10	591,97	<b>OPRM1</b>	6	8	8	506,55
<b>HTR2C</b>	X	7	7	371,68	<b>ORM1</b>	9	4	4	1132,94
<b>IFNL3</b>	19	7	7	409,30	<b>ORM2</b>	9	2	2	745,63
<b>IFNL4</b>	19	5	6	452,73	<b>P2RY1</b>	3	2	2	464,44
<b>IL1B</b>	2	1	1	536,75	<b>P2RY12</b>	3	1	1	399,00
<b>IMPDH1</b>	7	2	2	269,51	<b>PARP12</b>	7	25	25	878,85
<b>ITGB3</b>	17	1	1	514,13	<b>PCSK9</b>	1	1	1	298,63
<b>ITPA</b>	20	1	1	715,50	<b>PGAP3</b>	17	1	1	279,63
<b>KCNIP1</b>	5	2	2	506,07	<b>PLEC</b>	8	1	1	262,63
<b>KCNIP4</b>	4	6	6	430,36	<b>PNMT</b>	17	3	3	211,88
<b>KCNJ11</b>	11	1	1	346,13	<b>POLG</b>	15	1	1	540,63
<b>KCNMB1</b>	5	2	2	506,07	<b>PON1</b>	7	5	8	793,98
<b>KIF6</b>	6	3	3	661,13	<b>PON2</b>	7	4	4	500,51
<b>LDLR</b>	19	11	11	473,49	<b>PON3</b>	7	7	7	908,86
<b>LEP</b>	7	2	2	573,44	<b>POR</b>	7	3	3	631,96
<b>LPA</b>	6	2	2	679,82	<b>PPARD</b>	6	47	47	607,97
<b>LPIN1</b>	2	1	1	665,38	<b>PPARG</b>	3	6	6	738,11
<b>LTA</b>	6	2	2	732,13	<b>PPP1R9A</b>	7	1	1	283,88
<b>LTA4H</b>	12	1	1	560,88	<b>PRKCA</b>	17	2	2	661,88
<b>LTC4S</b>	5	1	1	289,50	<b>PROC</b>	2	3	3	650,09
<b>LUC7L2</b>	7	1	1	663,88	<b>PROS1</b>	3	1	1	481,75
<b>LYZ</b>	12	1	1	349,25	<b>PRSS53</b>	16	1	1	354,50
<b>MAOB</b>	X	13	13	584,88	<b>PSORS1C1</b>	6	1	1	0,38
<b>MAT1A</b>	10	7	7	643,02	<b>PTGFR</b>	1	2	2	660,07
<b>MC4R</b>	18	4	4	693,63	<b>PTGIS</b>	20	12	12	592,26
<b>MED12L</b>	3	1	1	399,00	<b>PTGS1</b>	9	8	8	597,71
<b>METTL1</b>	12	1	1	615,75	<b>QPRT</b>	16	1	1	519,25
<b>MICA</b>	6	1	1	408,00	<b>RALBP1</b>	18	6	6	465,84
<b>MICB</b>	6	1	1	557,38	<b>RARG</b>	12	1	1	262,88
<b>MTHFR</b>	1	6	6	668,63	<b>RGS4</b>	1	3	3	490,13
<b>MTRR</b>	5	3	3	686,63	<b>RPF2</b>	6	1	1	732,38
<b>MUC21</b>	6	1	1	616,50	<b>RPL13</b>	16	2	2	395,57
<b>MYH7</b>	14	1	1	252,25	<b>RPP21</b>	6	1	1	0,00
<b>MYLIP</b>	6	1	1	832,50	<b>RXRA</b>	9	4	4	426,91

Tabla suplementaria 1 (cont.). Listado de regiones genómicas, tamaño y cobertura promedio para cada gen

Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura	Gen	Cromosoma	Número de regiones	Tamaño total (pb)	Promedio de cobertura
RYR1	19	45	45	674,46	SLCO5A1	8	5	5	482,43
SCN1A	2	3	3	561,75	SOD2	6	2	2	548,19
SELE	1	1	1	952,38	SPECC1L	22	1	1	200,13
SELP	1	1	1	507,63	SPG7	16	8	8	620,44
SEMA3C	7	2	2	769,50	SPN	16	7	7	373,72
SERPINA7	X	7	7	763,02	SRY	Y	7	715	40,86
SERPINE1	7	1	1	367,25	SULT1A1	16	3	3	678,67
SIGLEC12	19	1	1	638,50	SULT1A2	16	5	5	587,78
SIVA1	14	1	1	269,00	SULT1B1	4	6	6	514,31
SLC10A1	14	7	7	978,81	SULT1C2	2	9	9	796,60
SLC10A2	13	8	8	711,30	SULT1C4	2	7	7	663,61
SLC13A1	7	7	7	580,45	SULT1E1	4	5	5	607,40
SLC14A2	18	2	2	735,07	SULT2A1	19	6	6	727,44
SLC15A1	13	19	19	812,04	SULT2B1	19	7	7	618,47
SLC15A2	3	10	10	906,10	SULT4A1	22	5	5	566,50
SLC16A1	1	7	7	836,32	TANC1	2	7	7	721,70
SLC19A1	21	5	5	438,13	TBL1Y	Y	1	1	125,75
SLC22A1	6	30	31	942,09	TCF7L2	10	7	7	604,73
SLC22A11	11	6	6	510,84	TGFBR3	1	1	1	97,50
SLC22A12	11	2	2	403,07	TNF	6	9	9	904,04
SLC22A13	3	7	7	507,05	TOMM40	19	1	1	516,88
SLC22A14	3	6	6	588,15	TPH2	12	2	2	627,69
SLC22A16	6	1	1	519,13	TPMT	6	9	9	554,04
SLC22A2	6	16	19	752,27	TPSG1	16	2	2	231,07
SLC22A3	6	5	5	685,65	TRIM40	6	1	1	0,13
SLC22A4	5	7	16	674,49	TSPY8	Y	1	1	122,00
SLC22A5	5	8	8	717,52	TTN	2	1	1	347,50
SLC22A6	11	7	8	447,47	TYMS	18	5	5	1027,90
SLC22A8	11	11	11	636,59	UGT1A1	2	29	62	1074,66
SLC25A27	6	3	3	410,25	UGT2A1	4	7	7	622,09
SLC28A1	15	27	27	794,69	UGT2B11	4	3	3	780,13
SLC28A2	15	7	7	950,50	UGT2B15	4	6	6	616,94
SLC28A3	9	7	7	644,59	UGT2B17	4	14	504	147,66
SLC29A2	11	6	6	484,54	UGT2B28	4	3	3	686,79
SLC35B2	6	9	9	693,46	UGT2B4	4	7	7	631,29
SLC47A2	17	3	3	338,55	UGT2B7	4	8	8	739,52
SLC5A6	2	7	7	597,11	UGT8	4	5	5	1058,18
SLC6A4	17	7	46	255,18	UMPS	3	4	4	726,35
SLC7A5	16	3	3	700,17	USP24	1	1	1	752,50
SLC7A7	14	7	7	685,41	USP40	2	8	8	1044,28
SLC7A8	14	7	7	605,40	USP9Y	Y	1	1	110,88
SLC9A7	X	2	2	383,75	VDR	12	3	3	593,46
SLCO1A2	12	16	17	540,64	VKORC1	16	22	22	613,28
SLCO1B1	12	18	18	643,56	XDH	2	8	8	666,47
SLCO1B3	12	7	11	556,13	XRCC1	19	3	3	468,79
SLCO2B1	11	7	7	648,79	YEATS4	12	1	1	535,13
SLCO3A1	15	6	7	572,52	ZBTB42	14	1	1	623,38
SLCO4A1	20	5	5	521,43	ZNRD1	6	1	1	0,00

## Bibliografía

1. Roden DM, McLeod HL, Relling MV, Williams MS, Mensah GA, Peterson JF, et al. Pharmacogenomics. *Lancet* [Internet]. 2019;373(10):1-12. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S0140673619312760>
2. Han SM, Park J, Lee JH, Lee SS, Kim H, Han H, et al. Targeted Next-Generation Sequencing for Comprehensive Genetic Profiling of Pharmacogenes. *Clin Pharmacol Ther*. 2017;101(3):396-405.
3. Ji Y, Si Y, McMillin GA, Lyon E. Clinical pharmacogenomics testing in the era of next generation sequencing: challenges and opportunities for precision medicine. *Expert Rev Mol Diagn* [Internet]. 2018;18(5):411-21. Disponible en: <https://doi.org/10.1080/14737159.2018.1461561>
4. Bank P, Caudle K, Swen J, Gammal R, Whirl-Carrillo M, Klein T, et al. Comparison of the Guidelines of the Clinical Pharmacogenetics Implementation Consortium and the Dutch Pharmacogenetics Working Group. *Clin Pharmacol Ther* [Internet]. 2017;4457. Disponible en: <http://onlinelibrary.wiley.com/doi/10.1002/cpt.762/full>
5. Bousman CA, Jaksa P, Pantelis C. Systematic evaluation of commercial pharmacogenetic testing in psychiatry. *Pharmacogenet Genomics* [Internet]. 2017;27(11):387-93. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed/28777243>
6. Kalman LV, Agúndez JAG, Appell ML, Black JL, Bell GC, Boukouvala S, et al. Pharmacogenetic Allele Nomenclature: International Workgroup Recommendations for Test Result Reporting. *Clin Pharmacol Ther* [Internet]. 2016;99(2):172-85. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed/26479518>
7. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flücke P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res* [Internet]. 2019 Oct 31. Disponible en: <https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkz950/5610347>
8. Wittig M, Anmarkrud JA, Kassens JC, Koch S, Forster M, Ellinghaus E, et al. Development of a high-resolution NGS-based HLA-typing and analysis pipeline. *Nucleic Acids Res* [Internet]. 2015;43(11):e70-e70. Disponible en: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv184>
9. Agilent Technologies. SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library [Internet]. 2019 [consultado 07/06/2020]. Disponible en: <https://www.agilent.com/cs/library/usermanuals/Public/G7530-90000.pdf>
10. Qiagen. Manual de instrucciones de uso de QIAsymphony® DSP DNA [Internet]. 2015 [consultado 07/06/2020]. Disponible en: <https://www.qiagen.com/fr/resources/download.aspx?id=8bc88dad-4140-467e-a1f0-e390fd193865&lang=es-ES>
11. Illumina. HiSeq 1500 System Guide (1503578 v03) [Internet]. 2019 [consultado 07/06/2020]. Disponible en: [https://emea.support.illumina.com/downloads/hiseq\\_1500\\_user\\_guide\\_15035788.html](https://emea.support.illumina.com/downloads/hiseq_1500_user_guide_15035788.html)
12. Ochoa JP, Sabater-Molina M, García-Pinilla JM, Mogensen J, Restrepo-Córdoba A, Palomino-Doza J, et al. Formin Homology 2 Domain Containing 3 (FHOD3) Is a Genetic Basis for Hypertrophic Cardiomyopathy. *J Am Coll Cardiol* [Internet]. 2018;72(20):2457-67. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/30442288>
13. Ortiz-Genga MF, Cuenca S, Dal Ferro M, Zorio E, Salgado-Aranda R, Climent V, et al. Truncating FLNC Mutations Are Associated With High-Risk Dilated and Arrhythmic Cardiomyopathies. *J Am Coll Cardiol* [Internet]. 2016;68(22):2440-51. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/27908349>
14. Trujillo-Quintero JP, Gutiérrez-Agulló M, Ochoa JP, Martínez-Martínez JG, de Uña D, García-Fernández A. Familial Brugada Syndrome Associated With a Complete Deletion of the SCN5A and SCN10A Genes. *Rev Esp Cardiol* (Engl Ed) [Internet]. 2019;72(2):176-8. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/29650450>
15. The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* [Internet]. 2012;491(7422):56-65. Disponible en: <http://www.nature.com/articles/nature11632>
16. Numanagić I, Malikić S, Pratt VM, Skaar TC, Flockhart DA, Sahinalp SC. Cypripi: exact genotyping of CYP2D6 using high-throughput sequencing data. *Bioinformatics* [Internet]. 2015;31(12):i27-34. Disponible en: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv232>
17. Lee S, Wheeler MM, Patterson K, McGee S, Dalton R, Woodahl EL, et al. Star-gazer: a software tool for calling star alleles from next-generation sequencing data using CYP2D6 as a model. *Genet Med* [Internet]. 2019;21(2):361-72. Disponible en: <http://dx.doi.org/10.1038/s41436-018-0054-0>
18. Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, et al. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nat Biotechnol* [Internet]. 2014;32(3):246-51. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/24531798>
19. BCFTools - Variant Calling [Internet] [consultado 08/06/2020]. Disponible en: <https://samtools.github.io/bcftools/howtos/variant-calling.html>
20. Pratt VM, Everts RE, Aggarwal P, Beyer BN, Broeckel U, Epstein-Baak R, et al. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes: A GeT-RM Collaborative Project. *J Mol Diagnostics* [Internet]. 2016;18(1):109-23. Disponible en: <http://dx.doi.org/10.1016/j.jmoldx.2015.08.005>
21. Pratt VM, Zehnbauer B, Wilson JA, Baak R, Babic N, Bettinotti M, et al. Characterization of 107 Genomic DNA Reference Materials for CYP2D6, CYP2C19, CYP2C9, VKORC1, and UGT1A1. *J Mol Diagnostics* [Internet]. 2010;12(6):835-46. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S1525157810601341>
22. Arbitrio M, Di Martino MT, Scionti F, Agapito G, Hiram Guzzi P, Cannataro M, et al. DMETTM (Drug Metabolism Enzymes and Transporters): a Pharmacogenomic platform for precision medicine. *Oncotarget* [Internet]. 2016;5(0). Disponible en: <http://www.oncotarget.com/abstract/9927>
23. Agena Bioscience: iPLEX® PGx Pro Panel Flyer [Internet] [consultado 04/06/2020]. Disponible en: <https://agenabio.com/resources/product-literature/>
24. Twist GP, Gaedigk A, Miller NA, Farrow EG, Willig LK, Dinwiddie DL, et al. Constellation: a tool for rapid, automated phenotype assignment of a highly polymorphic pharmacogene, CYP2D6, from whole-genome sequences. *npj Genomic Med* [Internet]. 2016;1(August 2015):15007. Disponible en: <http://www.nature.com/articles/npgenomed20157>
25. Gaedigk A, Turner A, Everts RE, Scott SA, Aggarwal P, Broeckel U, et al. Characterization of Reference Materials for Genetic Testing of CYP2D6 Alleles: A GeT-RM Collaborative Project. *J Mol Diagnostics* [Internet]. 2019;21(6):1034-52. Disponible en: <https://doi.org/10.1016/j.jmoldx.2019.06.007>
26. Owen RP, Sangkuhl K, Klein TE, Altman RB. Cytochrome P450 2D6. *Pharmacogenet Genomics* [Internet]. 2009;19(7):559-62. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/19512959>
27. Hicks JK, Bishop JR, Sangkuhl K, Muller DJ, Ji Y, Leckband SG, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. *Clin Pharmacol Ther*. 2015;98(2):127-34.
28. Swen JJ, Nijenhuis M, de Boer A, Grandia L, Maitland-van der Zee AH, Mulder H, et al. Pharmacogenetics: from bench to byte—an update of guidelines. *Clin Pharmacol Ther* [Internet]. 2011;89(5):662-73. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/21412232>
29. Drögemöller BI, Wright GEB, Shih J, Monzon JG, Gelmon KA, Ross CJ, et al. CYP2D6 as a treatment decision aid for ER-positive non-metastatic breast cancer patients: a systematic review with accompanying clinical practice guidelines. *Breast Cancer Res Treat* [Internet]. 2019;173(3):521-32. Disponible en: <http://dx.doi.org/10.1007/s10549-018-0527-0>
30. Bell GC, Caudle KE, Whirl-Carrillo M, Gordon RJ, Hikino K, Prows CA, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline for CYP2D6 genotype and use of ondansetron and tropisetron. *Clin Pharmacol Ther* [Internet]. 2017;102(2):213-8. Disponible en: <http://www.pharmgkb.org/page/cpic/>
31. Madadi P, Amstutz U, Rieder M, Ito S, Fung V, Hwang S, et al. Clinical practice guideline: CYP2D6 genotyping for safe and efficacious codeine therapy. *J Popul Ther Clin Pharmacol*. 2013;20(3):369-96.
32. Klein TE, Ritchie MD. PharmCAT: A Pharmacogenomics Clinical Annotation Tool. *Clin Pharmacol Ther* [Internet]. 2018;104(1):19-22. Disponible en: <http://doi.wiley.com/10.1002/cpt.928>
33. Numanagić I, Malikić S, Ford M, Qin X, Toji L, Radovich M, et al. Allelic decomposition and exact genotyping of highly polymorphic and structurally variant genes. *Nat Commun* [Internet]. 2018;9(1):828. Disponible en: <http://www.nature.com/articles/s41467-018-03273-1>
34. Barbarino JM, Krötz DL, Klein TE, Altman RB. PharmGKB summary: very important pharmacogenomic information for human leukocyte antigen B. *Pharmacogenet Genomics* [Internet]. 2015;25(4):205-21. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/25647431>
35. Hosomichi K, Shiina T, Tajima A, Inoue I. The impact of next-generation sequencing technologies on HLA research. *J Hum Genet* [Internet]. 2015;60(11):665-73. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/26311539%5Cnhttp://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC4660052>
36. Roy S, Coldren C, Karunamurthy A, Kip NS, Keele EW, Lincoln SE, et al. Standards and Guidelines for Validating Next-Generation Sequencing Bioinformatics Pipelines. *J Mol Diagnostics* [Internet]. 2018;20(1):4-27. Disponible en: <https://linkinghub.elsevier.com/retrieve/pii/S1525157817303732>
37. Hovelson DH, Xue Z, Zawistowski M, Ehm MG, Harris EC, Stocker SL, et al. Characterization of ADME gene variation in 21 populations by exome sequencing. *Pharmacogenet Genomics* [Internet]. 2017;27(3):89-100. Disponible en: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=01213011-900000000-99273>
38. Kozyra M, Ingelman-Sundberg M, Lauschke VM. Rare genetic variants in cellular transporters, metabolic enzymes, and nuclear receptors can be important determinants of interindividual differences in drug response. *Genet Med*. 2017;19(1):20-9.
39. Lauschke VM, Ingelman-Sundberg M. How to Consider Rare Genetic Variants in Personalized Drug Therapy. *Clin Pharmacol Ther* [Internet]. 2018;103(5):745-8. Disponible en: <http://dx.doi.org/10.1002/cpt.976>