

UNIVERSIDAD AUTÓNOMA DE COAHUILA
FACULTAD DE MEDICINA
DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN



**“EVALUACIÓN Y DESARROLLO DE HERRAMIENTAS BIOINFORMÁTICAS
PARA EL ANÁLISIS *IN SILICO* DE POLIMORFISMOS DEL GEN *CYP2D6*.”**

Por

Luis Emilio Vázquez Maya

Como requisito parcial para obtener el Grado de

MAESTRÍA EN INVESTIGACIÓN MULTIDISCIPLINARIA DE SALUD

TORREÓN, COAHUILA

FEBRERO 2017

RESUMEN

Introducción. La variabilidad en la respuesta a los fármacos tiene atribución en parte a las diferencias individuales en el metabolismo, que incluye factores ambientales (por ejemplo, la edad, ingesta en la dieta, estilo de vida, terapias de fármacos constantes) y por la genética del individuo en base al polimorfismo en genes que codifican las enzimas metabolizadoras de fármacos.

Una de las principales enzimas implicadas en el metabolismo y degradación de medicamentos en mamíferos es el citocromo P450 (CYP450). El complejo de enzimas del citocromo P450 está altamente distribuido en animales, plantas y protistas. Estas enzimas se encuentran presentes en tejidos tales como riñón, pulmón, piel, intestino, corteza adrenal, testículos, placenta y otros, pero principalmente activo en el hígado.

El gen *CYP2D6* es altamente polimórfico en el cual se han identificado más de cien variantes alélicas. Algunas de ellas ya han sido claramente identificadas y se encuentran publicadas en la base de datos “The Human Cytochrome P450 (CYP) Allele Nomenclature Database” (<http://www.cypalleles.ki.se/>) donde se describen las relevancias clínicas de ellas. Además, el *CYP2D6* es responsable del metabolismo y eliminación de aproximadamente el 25% de los fármacos de uso común.

Las principales modificaciones en este gen son debido a los polimorfismos de un solo nucleótido (SNP) y las variaciones en el número de copias (CNV) del gen. Aunado a ello, se encuentran alelos funcionales, alelos no funcionales y alelos con función reducida de acuerdo a su capacidad metabólica. Debido a las diferencias en la actividad enzimática que se muestra en diferentes poblaciones, con motivo de estos cambios, se suele hacer referencia a 4 grupos: el metabolizador pobre (PM), intermedio (IM), extensivo/normal (EM) y ultra-rápido (UM).

Objetivo. Diseñar y evaluar una herramienta bioinformática para el análisis *in silico* de polimorfismos del gen *CYP2D6*.

Materiales y métodos. Estudio observacional, descriptivo, prospectivo y transversal, con un diseño de investigación epidemiológico a nivel descriptivo. Se analizaron 64 muestras de secuenciación de exoma completo y 98 muestras de

microarreglos, todas las muestras pertenecientes a individuos mexicanos mestizos. La extracción de ADN genómico a partir de sangre periférica fue realizada por medio de la técnica “salting out”, para su genotipificación a través del sistema HiSeq® 4000 de Illumina® con longitudes de lectura máximas de 2 x 100pb. Adicionalmente, se genotipificaron 98 muestras por medio de la tecnología de microarreglos HumanOmniExpressExome de Illumina®, con cobertura de ~900,000 SNPs. Después, se los datos crudos obtenidos fueron transformados en formato VCF y se analizó su calidad. Posterior a ello, se realizó la identificación de los alelos y su genotipo para la inferencia del puntaje de actividad (AS). Finalmente, se creó una base de datos con los resultados y su análisis de frecuencias alélicas y genotípicas.

Resultados. El análisis de validación de calidad de las secuenciaciones mostró, en su mayoría, un contenido GC no adecuado; sin embargo, estas indicaciones no afectaron la predicción de los fenotipos. El nombramiento de las muestras de exomas mostraron una alta frecuencia para el diplotipo *CYP2D6*2/*2* (67.18%), seguida por *CYP2D6*2/*4* (6.25%), *CYP2D6*2/*82* (6.25%) y el nombramiento ambiguo *CYP2D6*2/*87* (4.68%); mientras que los microarreglos reportaron una alta frecuencia para el diplotipo *CYP2D6*1/*2* (28.57%), seguido de *CYP2D6*1/*4* (16.33%), *CYP2D6*1/*53* (15.31%), *CYP2D6*1/*41* (6.12%), *CYP2D6*2/*2* (6.12%), y *CYP2D6*2/*4* (3.06%), entre otros de menor repetición. Los datos mostraron una nula presencia de individuos con metabolismo lento y ultrar-rápido. Asimismo, en concordancia con la literatura, las mayores frecuencias se observaron en los alelos para actividad normal de la enzima (*CYP2D6*1* y **2*). Además, se observó la aparición del alelo *CYP2D6*82* que es propio de la población amerindia mexicana y del alelo *CYP2D6*4* que se considera de origen caucásico y que ha sido asociado a un pobre metabolismo.

Conclusiones. Las herramientas para la evaluación y nombramiento de alelos de enzimas metabolizadoras del citocromo CYP450 puede resultar una herramienta útil en la predicción de fenotipos, proveyendo una ayuda en la decisión del suministro de un fármaco determinado. En este trabajo de investigación se realizó un análisis *in silico* para la predicción de fenotipos en individuos mexicanos.

Palabras clave: CYP2D6, polimorfismos, fenotipo, herramientas bioinformáticas

ABSTRACT

Introduction. Variability in the response to drugs is attributable in part to individual differences in metabolism, which includes environmental factors (age, dietary intake, lifestyle, constant drug therapies) and polymorphisms in the genes encoding the enzymes metabolizers of drugs.

One of the main enzymes involved in the metabolism and degradation of drugs in mammals is cytochrome P450 (CYP). The cytochrome P450 enzyme complex is highly distributed in animals, plants and protists. It is present in tissues such as: kidney, lung, skin, intestine, adrenal cortex, testes, placenta and others, but mainly active in the liver.

CYP2D6 is the name by which a specific enzyme in the broad family of CYP450 is known. It takes its name directly from cytochrome p450, family 2, subfamily D, member 6.

The *CYP2D6* gene is highly polymorphic and up to now more than one hundred allelic variants and mutations have been identified. Some of them have already been clearly identified and are published in The Human Cytochrome P450 (CYP) Allele Nomenclature Database (<http://www.cypalleles.ki.se/>) which describes the clinical findings of them.

The main modifications are due to polymorphisms of a single nucleotide (SNP) and variations in the number of genes copied (CNV). In addition, functional alleles, non-functional alleles and alleles with reduced function are found. Due to the differences in enzymatic activity shown in different populations due to these changes, 4 groups are usually referred to: poor metabolizer (PM), intermediate (IM), extensive (EM) and ultrarapid (UM).

Objective. Design and evaluate a bioinformatics tool for the in silico analysis of CYP2D6 gene polymorphisms.

Materials and methods. Observational, descriptive, prospective and cross-sectional study, with an epidemiological research design at a descriptive level. We analyzed 64 samples of complete sequencing of the exome and 98 samples of

microarrays, all Mexican mestizos. The genomic DNA was isolated from peripheral blood by means of a standardized process of “salting out”, for its subsequent genotyping by the HiSeq® 4000 system of Illumina® with maximum reading lengths of 2 x 100 bp. The SNPs were determined by means of the HumanumExpressExome microarray technology from Illumina®, with coverage of ~ 900,000 SNPs. The files obtained were modified until they were in VCF format and their quality and validation were analyzed. After that, the alleles and their diplotype were named, for the subsequent inference of the activity score (AS) corresponding to the phenotype. Finally, a database was created with the results and their analysis of frequencies, allelic and phenotypic.