



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

“DESARROLLO DE PRUEBAS GENÓMICAS PRIORITARIAS PARA LA SALUD PÚBLICA DE MÉXICO”

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS

PRESENTA:

OMAR FERNANDO CRUZ CORREA

TUTOR PRINCIPAL:

Dr. Francisco Xavier Soberón Mainero Instituto de Biotecnología, UNAM

MIEMBROS DEL COMITÉ TUTOR:

Dra. María Dolores Correa Beltrán Instituto Nacional de Pediatría
Dr. Gabriel del Río Guerra Instituto de Fisiología Celular, UNAM

MÉXICO, D.F. JUNIO, 2013



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

AGRADECIMIENTOS:

En primer lugar al doctor Xavier Soberón por todo el apoyo que me brindó y toda la confianza que depositó en mí y en mi trabajo. También agradezco a los miembros de mi comité tutor por sus útiles comentarios y observaciones que me fueron haciendo a lo largo de todo el proyecto, particularmente al doctor Ruy López Ridaura del Instituto Nacional de Salud Pública quien por desgracia no pudo ser reconocido oficialmente ante la UNAM como parte del comité.

Al doctor Francisco Quiñones Falconi del Instituto Nacional de Enfermedades Respiratorias y a todo su equipo por compartir con nosotros los aspectos prácticos actuales del diagnóstico y tratamiento de las enfermedades infecciosas respiratorias en esta institución, así como también por el aislamiento, cultivo y caracterización de las muestras de *Mycobacterium tuberculosis* analizadas en este trabajo.

También a todos mis amigos y compañeros en el Instituto Nacional de Medicina Genómica (INMEGEN), por compartir su tiempo conmigo y ofrecerme su amistad, en especial a los miembros de la Unidad de Secuenciación e Identificación de Polimorfismos: la maestra Karol Carrillo, Julio Canseco, Salvador Hernández y Haydee Miranda.

A mis padres les agradezco no sólo el haberme inculcado hábitos de estudio, sino por mantenerse a mi lado, y junto a mi hermano apoyarme todo este tiempo, alentándome siempre a alcanzar mis metas. Así mismo quiero agradecer a Lourdes Ramírez Hobak el que me brindara su compañía en tiempos difíciles y una hermosa sonrisa tranquilizadora siempre que lo necesitara.

A todos aquellos amigos, profesores e investigadores que de una manera u otra me ofrecieron su consejo, apoyo y amistad e hicieron posible la realización de este trabajo: Gracias.

CONTENIDO:

	<u>Página</u>
<u>1. RESUMEN</u>	5
<u>2. ABSTRACT</u>	6
<u>3. INTRODUCCIÓN</u>	7
3.1 Características fundamentales de las pruebas genómicas	8
3.1.1 <i>Objetivos de las pruebas genómicas</i>	9
3.1.2 <i>Naturaleza de la muestra</i>	11
3.1.3 <i>Moléculas informacionales “blanco”</i>	11
3.1.4 <i>Enfoque de la prueba genómica</i>	12
3.1.5 <i>Tipo o nivel de la información buscada</i>	13
3.1.6 <i>Reportes de resultados posibles</i>	14
3.1.7 <i>Método de análisis genómico</i>	15
<u>4. ANTECEDENTES</u>	19
<u>5. HIPÓTESIS</u>	21
<u>6. OBJETIVOS</u>	21
6.1 Objetivo principal	21
6.2 Objetivos secundarios	21
<u>7. METODOLOGÍA</u>	22
7.1 Estrategia	22
7.2 Priorización de enfermedades	25
7.2.1 <i>Definición de parámetros para la priorización de enfermedades</i>	25
7.2.2 <i>Guías para la determinación del valor de cada parámetro</i>	26
7.2.3 <i>Grupo de padecimientos a evaluar</i>	31
7.3 Propuesta de una prueba genómica para la enfermedad seleccionada	32
7.4 Validación analítica de la prueba propuesta	34
7.4.1 <i>Amplificación por PCR</i>	34
7.4.2 <i>Purificación de los productos de PCR</i>	35
7.4.3 <i>Cuantificación de los productos de PCR y preparación de una mezcla equimolar</i>	35
7.4.4 <i>Preparación de bibliotecas y secuenciación con Ion Torrent PGM</i>	35
7.4.5 <i>Validación de la detección de mutaciones</i>	35

7.4.6 Validación de la detección de resistencia a antibióticos mediante secuenciación	36
7.4.7 Concordancia con datos fenotípicos	36
<u>8. RESULTADOS</u>	37
8.1 Priorización de enfermedades	37
8.1.4 Evaluación de las enfermedades	37
8.1.5 Selección de una enfermedad prioritaria	38
8.2 Propuesta de una prueba genómica para la enfermedad seleccionada	41
8.2.1 Necesidades específicas del ámbito clínico	41
8.2.2 Definición de un esquema de proceso particular	43
8.3 Validación analítica de la prueba propuesta	46
8.3.1 Validación de la detección de mutaciones	46
8.3.2 Validación de la detección de resistencia a antibióticos mediante secuenciación	49
8.3.3 Concordancia entre la resistencia inferida por secuenciación y los datos fenotípicos	59
<u>9. DISCUSIÓN Y CONCLUSIONES</u>	62
9.1 Priorización de enfermedades	62
9.2 Propuesta de una prueba genómica	64
<u>10. REFERENCIAS</u>	68
<u>11. ANEXOS</u>	A1
ANEXO 1. Resultados de la evaluación por enfermedad	A1
ANEXO 2. Secuencias de los cebadores utilizados para la amplificación de las regiones analizadas	A9
ANEXO 3. Antibióticos para el tratamiento de tuberculosis y mecanismos de resistencia considerados en este trabajo	A10
ANEXO 4. Protocolo detallado de la prueba propuesta	A12
ANEXO 5. Profundidad de cobertura para las regiones del genoma de <i>Mycobacterium tuberculosis</i> analizadas	A23
ANEXO 6. Mutaciones detectadas en las 16 muestras de <i>Mycobacterium tuberculosis</i> analizadas con la prueba propuesta	A32
<u>12. REFERENCIAS DE LOS ANEXOS</u>	A36

1. RESUMEN:

Consideramos que el uso de las tecnologías de análisis genómico en el cuidado de la salud debe darse a través del desarrollo de nuevas pruebas que sean consideradas útiles para la práctica clínica y, a la vez, prioritarias para la salud pública de México. Para ello, en el presente trabajo proponemos y aplicamos una estrategia basada en la priorización de diferentes padecimientos para la selección de un área de enfermedad, el análisis de las necesidades actuales de la práctica clínica en el área de enfermedad seleccionada y la definición de un esquema de proceso particular que cumpla con dichas necesidades.

En primera instancia, nos enfocamos a elaborar un catálogo de las enfermedades con mayor costo social en cuanto a la viabilidad y pertinencia del desarrollo de pruebas con tecnologías de análisis genómico enfocadas a su manejo clínico (ya sea diagnóstico, pronóstico, predisposición o selección de tratamiento). Identificando mediante este análisis algunas enfermedades donde existe mayor probabilidad de que la traducción de los descubrimientos y la aplicación de los recursos tecnológicos del área de la genómica generen beneficios importantes, los cuáles se puedan ver reflejados rápidamente en la salud de los mexicanos.

Posteriormente, seleccionamos las infecciones de las vías respiratorias inferiores (uno de los grupos de padecimientos encontrados como prioritarios según la clasificación respecto a la perspectiva genómica) como modelo para proponer una nueva prueba que mediante el uso de tecnología de análisis genómico asista al manejo clínico de los pacientes en esta área.

Como siguiente punto presentamos las características y requerimientos específicos del ámbito clínico en una institución dedicada en gran parte al diagnóstico y tratamiento de las infecciones respiratorias inferiores. Estas características resaltaron entre otras cosas la importancia de la farmacorresistencia en infecciones por *Mycobacterium tuberculosis* y funcionaron como punto de anclaje para desarrollar una propuesta de prueba que mediante el uso de secuenciación de nueva generación (en la plataforma PGM de Ion Torrent) puede asistir en el manejo clínico de este problema en particular.

Después, mediante el análisis de un grupo de muestras control de *Mycobacterium tuberculosis* caracterizadas previamente para resistencia a fármacos en el INER, evaluamos la capacidad analítica del esquema de proceso de secuenciación masiva definido para esta nueva prueba. En primer lugar validamos la detección de mutaciones, mediante la comparación entre dos plataformas de secuenciación diferentes, determinando con este análisis que parámetros tales como el contexto genómico y la frecuencia de variante permiten estimar cuáles de las variantes detectadas representan mutaciones verdaderamente presentes en las muestras y cuáles se deben únicamente a errores característicos de la plataforma de secuenciación utilizada.

También validamos la detección de resistencia a los principales antibióticos antituberculosos al comparar con los datos fenotípicos previamente obtenidos en el INER, obteniendo valores de sensibilidad de 0.8, 0.93, 0.44, 0.87, 0.8, 0.5 y 0.5 y especificidad de 1, cero, 0.6, cero, 0.86, 0.8 y 1

para la detección de resistencia a fluoroquinolonas, rifampicina, estreptomina, isoniazida, pirazinamida, etambutol y amikacina respectivamente.

Como se muestra, con la estrategia utilizada en este trabajo es posible proponer, desarrollar y validar prototipos de nuevas pruebas con tecnología de análisis genómico que se encuentren enfocadas al manejo clínico de una de las enfermedades prioritarias para la salud pública de México.

2. ABSTRACT:

In our view, the use of genomic analysis technology in health care settings must materialize through the development of new tests that are considered useful by the clinical practice and also of high priority according to Mexico's public health. To that effect, we propose and apply a strategy based on the prioritization of different health problems for the selection of a disease, the analysis of current needs of the clinical practice regarding the selected disease and the definition of a specific process pipeline capable of meeting those needs.

First of all, we set to catalogue the diseases with higher social cost according to the viability and appropriateness of developing genomic technology based tests focusing on their clinical management (diagnosis, prognosis, predisposition or treatment selection). Through this analysis we identified several diseases where the translation and application of genomic discoveries and technologic resources presents a higher probability of generating substantial benefits, which could be readily reflected in Mexican people's health.

Next, we selected the lower tract respiratory infections (one of the group of healthcare issues of first importance found through our genomics' perspective classification) as a working model for the proposal of a new test which employs genomic analysis technology to assist in the clinical management of patients within this area of disease.

Then, we present the characteristics and specific needs of clinical practice in an institution dedicated to a considerable extent to the diagnosis and treatment of lower respiratory infections. These characteristics point out, among other things, to the relevance of drug resistance development in *Mycobacterium tuberculosis* infections and provide a basis to propose a next-generation sequencing test (on the PGM sequencer from Ion Torrent) that may assist in the clinical management of this particular issue.

Afterwards, we analyzed a set of *Mycobacterium tuberculosis* control samples previously evaluated for drug resistance to assess the analytic capabilities of the next-generation sequencing pipeline defined for this new test. First, we validated the detection of mutations through the comparison between two different sequencing platforms, from this analysis we determined that parameters such as genomic context and variant frequency allow to estimate which of the variants

detected represent real mutations in the samples and which of them are due only to systematic errors characteristic of the sequencing platform used.

Also, we validated the detection of drug resistance to the principal antibiotics used in the treatment of tuberculosis by comparison with the fenotipic data previously reported, obtaining sensibility with values of 0.8, 0.93, 0.44, 0.87, 0.5 and 0.5 and specificity values of 1, zero, 0.6, zero, 0.86, 0.8 and 1 for the detection of resistance to fluoroquinolones, rifampin, streptomycin, isoniazid, pyrazinamide, ethambutol and amikacin, respectively.

As we show, with the strategy used here is possible to propose, develop and validate new prototypic tests based on genomic analysis technology focused on the clinical management of one of Mexico's public health top priorities.

3. INTRODUCCIÓN:

El desarrollo de las tecnologías utilizadas para el análisis genómico ha sido realmente vertiginoso¹ y tanto en el ámbito académico como en el comercial aparecen constantemente nuevas tecnologías, enfoques y variantes metodológicas que permiten abordar el análisis de las moléculas informacionales (ADN, ARN, proteínas) de manera más precisa y masiva.

En el mercado mexicano es posible encontrar actualmente productos y servicios que utilizan herramientas de análisis genómico para el diagnóstico de una gran gama de padecimientos y que, aunque pueden resultar muy útiles en ciertos aspectos, debemos considerar que tales productos y servicios han sido diseñados principalmente en base a las necesidades y características poblacionales de Europa, Japón y Estados Unidos. Como consecuencia de su diseño estos productos no necesariamente están optimizados para su utilización en México² y por lo tanto, es deseable no depender exclusivamente de la oferta de las empresas que los proveen y, en su lugar más bien, adaptar las nuevas tecnologías que ellas utilizan y diseñar nuevas pruebas genómicas (ya sea de diagnóstico, pronóstico, predisposición o selección de tratamiento) más acordes a las condiciones particulares de México y su población.

Sin embargo México, por su calidad de país en desarrollo, cuenta con servicios de salud donde los recursos son limitados y por lo tanto el desarrollo de pruebas con tecnología de análisis genómico debe estar enfocado de manera eficiente para maximizar los beneficios obtenidos de su inversión en él. Por esta razón resulta necesario contar con un sistema que, manteniendo presentes las características y necesidades específicas de nuestro país, nos permita identificar las posibilidades para nuevas pruebas genómicas que tengan la mayor relevancia posible para la práctica clínica de México, asegurando así un impacto positivo en la salud pública de nuestro país.

En este contexto consideramos que los beneficios que pueden ser obtenidos de manera más rápida y oportuna con la inversión hecha en tecnología genómica lo constituyen aquellos que involucren el desarrollo de pruebas para detección temprana y manejo clínico de enfermedades

que presenten problemas serios de salud pública para nuestro país para las cuales se hayan descubierto y corroborado asociaciones a blancos genéticos.

En cuanto a la implementación de las tecnologías de análisis genómico en la práctica clínica a través del desarrollo de nuevas pruebas para el diagnóstico o manejo clínico de las enfermedades, actualmente podemos distinguir dos enfoques extremos:

El primero de ellos se centra en el desarrollo de pruebas específicas que analizan por separado cada una de las mutaciones o genes relevantes para el manejo clínico de los pacientes, identificando de manera muy rápida, específica y confiable un número pequeño de blancos asociados a un subgrupo particular de enfermedades, por ejemplo, la detección por PCR de la infección por *Staphylococcus aureus* resistente a meticilina³ o de los virus de influenza A y B⁴.

El segundo enfoque, en contraste, es la tendencia al desarrollo de pruebas mucho más generales, cuyas plataformas analíticas son capaces de detectar simultáneamente todos los blancos genéticos conocidos que puedan ser relevantes para una patología o un grupo de padecimientos con características similares, como en el caso de una prueba descrita recientemente que identifica a los portadores de 448 padecimientos recesivos severos mediante secuenciación⁵ o de algunos microarreglos para detección de un gran número de patógenos respiratorios⁶ e incluso microarreglos que están diseñados para la identificación de prácticamente todos los virus conocidos^{7,8}.

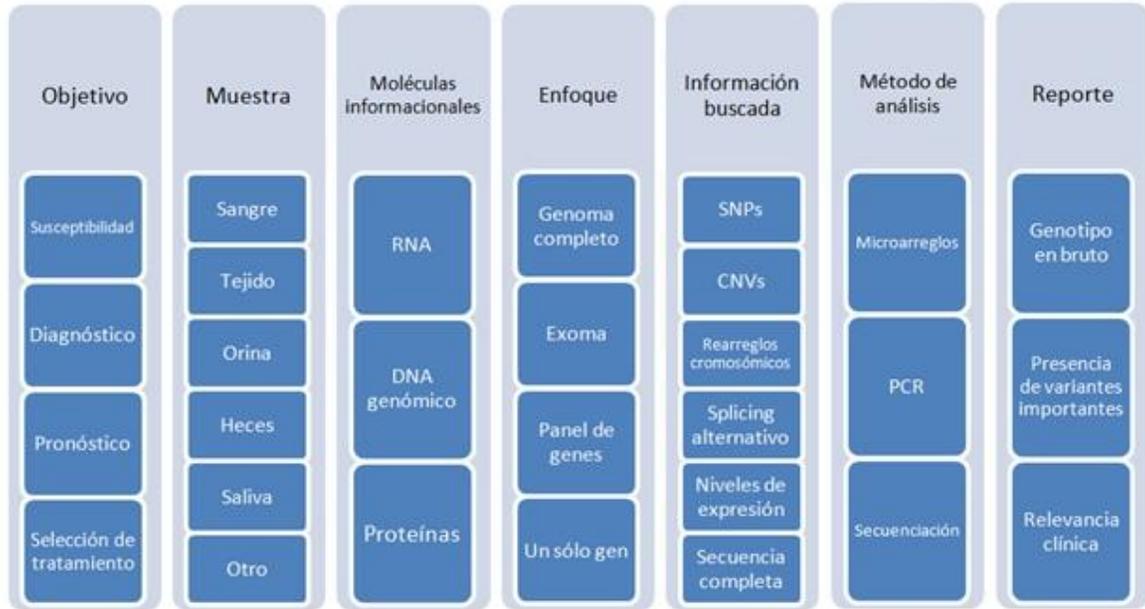
El desarrollo de este tipo de plataformas generales es un área prometedora para el futuro cercano debido al rápido abaratamiento del análisis genómico y al continuo surgimiento de nuevas tecnologías y variantes metodológicas para realizar dicho análisis con mayor precisión.

Todos estos recursos tecnológicos para el análisis genómico auguran, en principio, la posibilidad de diseñar sistemas que impacten de manera contundente la práctica clínica⁹, pero plantean el complejo reto de cómo seleccionar el mejor sistema analítico aplicable a cada problema en particular.

3.1 Características fundamentales de las pruebas genómicas

A continuación desglosaremos algunas características que son determinantes en el diseño de un proceso particular para una prueba con tecnología genómica y mencionaremos las diferentes alternativas que existen para cada una de ellas (Figura 1), tomando como ejemplo las enfermedades infecciosas de las vías respiratorias inferiores para ayudarnos a identificar las ventajas y limitaciones de las tecnologías disponibles para llevar a cabo los diferentes pasos del proceso de análisis, así como también la forma en la cual podemos adecuarlas al volumen de pruebas necesario y lograr obtener resultados suficientemente claros en un tiempo clínicamente relevante para la práctica clínica en México.

Figura 1. Características fundamentales para el diseño de una prueba con tecnología genómica.



3.1.1 Objetivos de las pruebas genómicas.

En la práctica clínica las pruebas de laboratorio, donde se incluyen aquellas con tecnología genómica, son utilizadas para obtener información acerca de una gran variedad de patologías, ya que se utilizan como herramientas para la toma de decisiones en cuanto al diagnóstico, pronóstico, tamizaje, monitoreo y evaluación del riesgo de los individuos.

Sin embargo el objetivo de una prueba con tecnología genómica particular es un factor determinante de la cantidad de individuos que son candidatos a la aplicación de dicha prueba (Figura 2). Esto, a su vez, decidirá el volumen de pruebas que tendrán que aplicarse en un tiempo determinado y por lo tanto influirá de manera importante en la elección de un método de análisis genómico.

Las pruebas de susceptibilidad buscan estratificar a los individuos según la magnitud del riesgo que tienen de presentar un padecimiento en el futuro y pueden aplicarse en la población general o sólo a aquellos individuos que presenten factores de riesgo adicionales para la enfermedad, por ejemplo determinado género, grupo de edad, niveles de exposición a contaminantes ambientales, etc. En este tipo de pruebas es donde se realizarían cotidianamente el mayor número de pruebas, debido a las grandes poblaciones que son candidato a su aplicación.

En cuanto a las pruebas de diagnóstico, éstas podrían aplicarse directamente a la población general en forma de tamizaje, pero en la mayoría de los casos se obtendrán mayores beneficios al realizarla en individuos en los cuales se sospeche la presencia de la enfermedad de acuerdo a los signos y síntomas que presente o en aquellos que presentan factores de riesgo importantes asociados al padecimiento.

Tanto las pruebas con objetivo pronóstico, que predicen la probabilidad de que los pacientes presenten mayores complicaciones y/o secuelas de los padecimientos como aquellas de selección de tratamiento que anticipan las reacciones adversas o la ineffectividad de éste deben, evidentemente, realizarse en individuos que han sido previamente diagnosticados para la enfermedad, por lo que se aplicarían a un grupo de individuos mucho más específico que para pruebas con otros objetivos.

Figura 2. Proporción de la población general que son candidatos a la aplicación de una prueba según su objetivo.



En el caso concreto de las enfermedades infecciosas de las vías respiratorias inferiores podría establecerse como objetivo de una prueba genómica la realización del diagnóstico diferencial en un solo ensayo, identificando el agente causante de entre la gran cantidad y diversidad de patógenos que producen síntomas similares.

Por otro lado, también se consideraría relevante el desarrollo de pruebas genómicas para la selección del tratamiento más efectivo para cada paciente a través de la identificación de genes de los organismos patógenos asociados a la resistencia ante los principales fármacos utilizados para tratar estas enfermedades. Este esquema resulta particularmente interesante debido a la creciente preocupación global sobre las consecuencias de la emergencia de resistencia a antibióticos entre las que se cuentan enfermedades más prolongadas, mayor riesgo de muerte y contagio aunados a gastos mayores en tratamiento y hospitalización¹⁰.

3.1.2 Naturaleza de la muestra.

Para asegurar la eficiencia de una prueba es requisito indispensable que el tipo de muestra a analizar sea adecuado para el objetivo buscado, factor especialmente evidente en enfermedades órgano específicas como las infecciosas y las neoplásicas, o cuando los blancos genéticos a interrogar revelan estados funcionales de tejidos particulares, ya que las muestras obtenidas fuera del área afectada resultan completamente inadecuadas para encontrar los blancos genéticos buscados (de hecho el espectro de mutaciones entre los genomas de células tumorales y células normales del mismo individuo es tan diferente que su comparación es un procedimiento acostumbrado en el estudio de los mecanismos oncogénicos¹¹).

En cuanto a los tipos de muestra existen muchas opciones, cada una con ventajas en ciertos contextos, por ejemplo, la obtención de sangre es un método llevado a cabo rutinariamente por muchos laboratorios clínicos, y por lo tanto, las metodologías para la extracción de ácidos nucleicos y proteínas a partir de este fluido se encuentran bien establecidas y es esta experiencia en su manejo lo que favorece su adopción en muchos casos como primera opción.

Mientras tanto, otros fluidos y secreciones corporales, por ejemplo saliva u orina, presentan la gran ventaja de que los métodos para su obtención son menos invasivos que para otros tipos de muestra, y pueden ser alternativas útiles en muchos casos. Nuevamente, existen protocolos muy bien establecidos para su procesamiento debido a su amplia adopción.

Por otro lado, los cortes de tejido pueden revelar información única, y se obtienen normalmente en aquellos casos donde el paciente necesita ser sometido a algún proceso invasivo, por ejemplo cuando es necesario realizar una biopsia para su análisis histológico. La reciente posibilidad de utilizar no sólo tejidos frescos para su análisis sino también aquellos previamente fijados en parafina, ha posibilitado un mayor uso de este tipo de muestras.

Los tipos de muestras utilizados normalmente para la detección de infecciones de las vías respiratorias inferiores son principalmente esputo espontáneo o inducido, lavado broncoalveolar y otras técnicas endoscópicas, mientras que los especímenes obtenidos mediante métodos menos complicados e invasivos, como podrían ser exudados bucales y nasofaríngeos, no son adecuadas para este tipo de padecimientos sino, más bien, para infecciones de las vías superiores. Estos tipos de muestra se encuentran en una situación análoga en cierta medida a la extracción de sangre, ya que también son obtenidos comúnmente para otros análisis en la práctica clínica actual y por lo tanto se cuenta con experiencia en su manejo (en hospitales de cierto nivel de especialidad).

3.1.3 Moléculas informacionales “blanco”.

Una variable fundamental para el diseño de una prueba genómica es el tipo de moléculas que se analizarán, ya que es necesario que éstas contengan la información que nos interesa acerca de las características de las enfermedades.

El ADN genómico nos puede proporcionar información útil sobre mutaciones somáticas y germinales, en todo tipo de regiones del genoma (regiones codificantes, intrones y regiones regulatorias), pero no nos proporciona datos acerca de los niveles de expresión génica.

Por otro lado, el ARN resulta indispensable cuando se desean analizar los niveles de expresión de genes específicos, así como eventos postranscripcionales (por ejemplo “splicing” alternativo), y también es posible analizarlo en busca de mutaciones, sin embargo, no proporciona información acerca de las regiones no transcritas del genoma y además es necesario un paso de transcripción inversa a ADN complementario previo a su análisis debido a que se degrada con mayor facilidad.

Además existe la posibilidad de utilizar también como moléculas informacionales a las proteínas, pero cabe mencionar que, por el momento, nos enfocamos en lo referente al análisis de ácidos nucleicos y dejamos deliberadamente al área de la proteómica fuera del alcance de este trabajo. Es evidente que el repertorio de proteínas expresadas puede reflejar de mejor manera el estado real de los organismos, pero todavía existen algunos obstáculos que superar antes de que sus aplicaciones clínicas avancen de manera importante y se inserten en la práctica clínica cotidiana¹², por lo que no abordaremos a profundidad en las características del análisis proteómico.

En el caso específico de las infecciones respiratorias de las vías inferiores se necesita considerar que muchos de los virus más importantes en estos padecimientos llevan su información genética en forma de ARN. Esto implica que la selección de moléculas informacionales depende de los patógenos específicos que sean incluidos en el análisis: para bacterias, hongos y parásitos puede utilizarse ADN o ARN mientras que para virus que utilizan ARN como material genético será indispensable utilizar ARN como molécula informacional.

3.1.4 Enfoque de la prueba genómica.

En cuanto al enfoque de las pruebas genómicas un abordaje de genoma completo nos permite el análisis simultáneo de toda la información genética asociada a la salud de un individuo, sin embargo, su costo es todavía demasiado elevado para un uso rutinario en la clínica. Además, para analizar la gran cantidad de datos generados es necesario contar con equipos de cómputo bastante robustos y costosos, así como también con personal especializado en el área de bioinformática para su manejo.

Como respuesta a estas dificultades, el genoma también podría dividirse en subconjuntos de tal manera que abarquen sólo ciertas regiones de las cuales se pueda obtener información mejor validada. Una de las ventajas de hacer esto es el mejor aprovechamiento de la altísima capacidad de procesamiento de las nuevas tecnologías para analizar un mayor número de muestras, al mismo tiempo que se facilita la interpretación de los resultados, debido a la disminución de la información y complejidad de los mismos¹³.

Dentro de estos enfoques, uno de los mejor desarrollados es el análisis a nivel de todos los exones (exoma), que se concentra sólo en las regiones codificantes para proteínas, donde el efecto

funcional de las variantes genéticas encontradas será, generalmente, más fácil de predecir y con mayor significancia biológica que en sus contrapartes no codificantes¹⁴.

Sin embargo, es posible que la mayor utilidad de estos enfoques se encuentren al centrar una prueba sólo en el subconjunto del genoma que abarque todas aquellas regiones relevantes para el diagnóstico y manejo clínico de una enfermedad en particular (que análogamente podría llamarse el “relevantoma” de dicho padecimiento), ya sea que se encuentren en regiones codificantes, intrónicas o regulatorias de gran cantidad de genes y en diversas regiones del genoma.

En cuanto a las enfermedades respiratorias inferiores podrían utilizarse enfoques de microbioma completo, analizando todos los ácidos nucleicos presentes en la muestra, o mediante un panel de genes escogido de tal manera que posibilite alcanzar el objetivo de la prueba (ya sea la identificación inequívoca de los patógenos presentes o de la resistencia a antibióticos).

3.1.5 Tipo o nivel de la información buscada.

La información que se obtiene mediante una prueba con tecnología de análisis genómico podría quizás definirse como la presencia, ya sea relativa o absoluta, de los blancos genéticos asociados a las características particulares de un padecimiento. Sin embargo, es importante que definamos el tipo al cual pertenecen los blancos de nuestro interés debido a que existen plataformas analíticas diseñadas específicamente para identificar los diferentes tipos.

A continuación se mencionan los principales tipos de blancos genéticos.

- Los “Single Nucleotide Polymorphisms” (SNPs) son variaciones en un solo nucleótido encontradas comúnmente dentro de una población y que pueden utilizarse como marcadores del genotipo de la región del genoma adyacente a donde se encuentran. En este hecho es donde se encuentra su mayor ventaja, ya que con relativamente pocos marcadores se puede cubrir una región extensa del genoma, sin embargo la principal desventaja de los SNPs es que en la mayoría de los casos solamente revelan efectos pequeños sobre los fenotipos asociados a una enfermedad.
- “Copy Number Variations” (CNVs) pueden ser tanto el aumento o la disminución en el número de copias que existen de una región dentro del genoma, estas variaciones son el resultado de duplicaciones o deleciones. La importancia de los CNVs se encuentra relacionada a la zona en la cual se producen, ya que cuando ocurre en una región codificante para proteína la variación en el número de copias normalmente resulta en niveles de expresión anormales y productos génicos alterados.
- Los rearrreglos cromosómicos son variaciones en la estructura de los cromosomas a partir de inversiones y translocaciones que no alteran el contenido génico. Sin embargo, cuando esto sucede normalmente se ve afectada la regulación de la expresión génica de la región donde ocurren y también es común la interrupción de secuencias codificantes.

- Eventos de “splicing” alternativo, es decir variaciones durante el proceso de edición de intrones que modifican la secuencia final de los ARNm maduros y, por lo tanto, dan lugar a diferentes isoformas de las proteínas codificadas en ellos. Los efectos en la función celular se deben a que estas isoformas varían en estabilidad, afinidad por sustrato y/o actividad.
- Los cambios en los niveles de expresión génica que se producen ante algún estímulo de interés. Estos son medidos indirectamente a través de la abundancia relativa de los transcritos bajo las condiciones de interés y se comparan contra aquellos obtenidos en condiciones de referencia. Analizando los cambios en los niveles de expresión es posible identificar las vías metabólicas y procesos celulares que se activan o reprimen durante la patogenia de las enfermedades.
- La secuencia completa, es decir, el orden de los ácidos nucleicos que componen las moléculas de ADN o ARN. Al obtener la secuencia completa de los ácidos nucleicos es posible identificar simultáneamente (*in silico*) todas las variantes que se encuentran en la región del genoma analizada, incluyendo aquellas que no han sido reportadas con anterioridad.

3.1.6 Reportes de resultados posibles.

La forma en la que se presentarán los resultados de una prueba genómica es un aspecto de gran importancia, ya que la facilidad de uso e interpretación resultan ventajas fundamentales para lograr que una prueba sea utilizada extensamente. Mientras más sencillos se presenten los resultados se tendrá una mayor influencia en la toma de decisiones, sin embargo, también se necesitará invertir más recursos en el procesamiento previo de los datos.

Una opción para la entrega de resultados es el reporte del genotipo en bruto, es decir, las características genéticas en las regiones interrogadas por la prueba, carentes de toda interpretación. Esta opción deja al receptor de los resultados la tarea de compararlos contra los genotipos de referencia para discernir la presencia de las variantes o blancos buscados. Esto es particularmente útil si se desea comparar los resultados contra diferentes referencias personalizadas (que podrían ser bases de datos específicas para varias subpoblaciones). Sin embargo, para que este esquema resulte práctico es necesario que los médicos cuenten con amplios conocimientos y experiencia en el análisis e interpretación de este tipo de resultados, así como que dispongan de los recursos necesarios para llevarlos a cabo.

Por otro lado, también podrían reportarse solamente la presencia o ausencia de los blancos buscados por la prueba, en lugar de presentarse todo el genotipo de las regiones estudiadas. Esto podría entregarse en forma de un sí/no o un número sencillo de entender, reduciendo la complejidad de la interpretación de los resultados.

Incluso podría llegarse más lejos y, como otra opción, ofrecer los resultados después de haber sido interpretados por un profesional de la salud. En este reporte de resultados se presentarían

recomendaciones con relevancia clínica emitidas en base a la prueba de forma estandarizada. Esta opción ayudaría a reducir ambigüedades en la interpretación, sin embargo requeriría de mayores recursos por parte del proveedor de la prueba, lo que implicaría a su vez un aumento en el costo de la misma.

Es evidente que el ámbito de la interpretación y entrega pertinente de los resultados será un área sumamente relevante en un ambiente clínico como el mexicano, donde podemos encontrar fuertes contrastes en cuanto a la experiencia en el manejo de la información genética con que cuentan los médicos que atienden directamente a los pacientes, así como también en cuanto a los recursos de los cuales disponen para ello.

En el caso específico de las pruebas para enfermedades infecciosas de las vías respiratorias inferiores podría fácilmente enlistarse los blancos genéticos encontrados, así como también añadir su interpretación para indicar la presencia de patógenos específicos o resistencia a ciertos fármacos. En el caso de identificar genes de resistencia también podría proporcionarse una recomendación de aquellos fármacos que no deben ser utilizados para presentar de forma más clara los resultados y facilitar la selección el tratamiento más adecuado.

3.1.7 Método de análisis genómico.

Un punto fundamental en el diseño de una prueba genómica es la selección del método más adecuado para identificar la presencia de los blancos genéticos buscados, ya que actualmente existe un gran número de plataformas tecnológicas, tanto comerciales como a nivel de investigación, que se encuentran disponibles para llevar a cabo el análisis genómico basado en ácidos nucleicos¹⁵.

Con base en el tipo de técnica utilizada para la identificación de las regiones genómicas objetivo podemos agrupar de manera general a las metodologías de análisis genómico en tres tecnologías principales:

- Las que utilizan reacción en cadena de la polimerasa (“PCR”) para la detección de secuencias específicas asociada a la amplificación de los blancos. En esta categoría agruparemos tanto esquemas de PCR “de punto final”, donde los productos de la reacción son analizados en base a su presencia y tamaño para obtener información puramente cualitativa, como la utilizada en los análisis forenses para la identificación de individuos¹⁶ como aquellos esquemas de PCR “en tiempo real”, donde se proporciona información cuantitativa al monitorear la cinética de reacción a través de diversos tipos de sondas y agentes intercalantes de ADN los cuales se unen a las secuencias objetivo y emiten fluorescencia conforme éstas regiones son amplificadas¹⁷.
- Aquellas que utilizan la hibridación de secuencias complementarias para la detección de los blancos, las cuales incluyen aquellas donde la hibridación se realiza en solución, pero

destacan por su frecuente uso las plataformas de alta densidad en matriz sólida (“microarreglos”)¹⁸.

- Las que utilizan la secuenciación directa de los nucleótidos de la muestra a analizar, para después compararlos *in silico* con secuencias de referencia, detectando de esta manera los blancos buscados cuando se encuentren presentes. Dentro de la amplia gama de plataformas disponibles para realizar la secuenciación nos enfocaremos más en aquellas de “nueva generación”, que se caracterizan por la producción masiva de información gracias a la secuenciación en paralelo de millones de moléculas de ADN.

Las tecnologías basadas en amplificación de ácidos nucleicos por PCR son muy robustas y se encuentran bien establecidas, pero normalmente permiten analizar simultáneamente un número limitado de variantes o blancos genéticos, debido principalmente a la interferencia entre los diferentes pares de cebadores (“primers”) necesarios para la amplificación de varios blancos a un mismo tiempo. Para aumentar el número de variantes que se pueden interrogar simultáneamente se ha recurrido a esquemas de compartimentación, por ejemplo, compañías como RainDance Technologies y Fluidigm Corporation han propuesto plataformas comerciales donde se divide cada muestra a analizar en una gran cantidad de “micro volúmenes” los cuales se ponen en contacto con diferentes grupos o “sets” de cebadores que interrogan una o más variantes por “micro volumen”, realizando cientos de PCR tradicionales en paralelo y evitando así mucha de la interferencia entre cebadores. Algo que se debe tomar en cuenta en este tipo de tecnologías es que al aumentar el número de muestras que se pueden procesar simultáneamente (“multiplex”) disminuye el número de blancos que se pueden interrogar por ensayo o experimento, por ejemplo, si se utilizara la plataforma Access Array IFC 48.48 de Fluidigm Corporation para 48 muestras podrían analizarse simultáneamente 480 variantes, mientras que en una sola muestra podría utilizarse la misma plataforma para analizar 23,040 variantes¹⁹.

Una de las mayores ventajas de las tecnologías de análisis basadas en PCR sobre otros métodos es que requieren un menor tiempo de procesamiento para la obtención de los resultados (aproximadamente un par de horas para PCR, algunos días para microarreglos y semanas enteras para la secuenciación de genoma completo). Sin embargo, actualmente empiezan a comercializarse tecnologías de secuenciación que utilizan tiempos mucho más comparables a los utilizados por una PCR convencional, por ejemplo la plataforma Personal Genome Machine de la compañía Life Technologies es capaz de secuenciar entre 10 y 100 Mb en dos horas (sin contar el tiempo que se necesita invertir previamente en la preparación de la muestra)²⁰.

Mientras tanto, los métodos basados en la detección por hibridación complementaria (“microarreglos”) nos adentran en mayor medida en el ámbito de la genómica, ya que como ventaja principal nos permiten analizar un mucho mayor número de blancos genéticos que aquellos que se encuentran basados en amplificación por PCR (llegando incluso a encontrarse diseños comerciales de genotipificación y de análisis de expresión génica que interrogan varios millones de blancos simultáneamente). Por otro lado, la principal limitación que presentan es el

hecho de que son capaces de procesar un número bastante limitado de muestras a un mismo tiempo. Por ejemplo, dentro de los productos de la compañía Illumina se encuentra la plataforma HumanOmni2.5-8 que es capaz de detectar 2.5 millones de SNPs en ocho muestras y la plataforma HumanOmni5-Quad que puede identificar cerca de 5 millones de estas variantes pero solamente en cuatro muestras de manera simultánea²¹.

Como particularidad especial de los microarreglos se ha observado que en el análisis de muestras clínicas los ácidos nucleicos contaminantes que presenten cierta complementariedad de secuencia hibridarán inespecíficamente con las sondas del microarreglo, aún cuando las sondas hayan sido diseñadas *in silico* de tal forma que debieran maximizar la hibridación específica (y minimizar la hibridación cruzada⁷. Sin embargo, esta característica de los microarreglos en ocasiones puede convertirse en una ventaja al permitir la detección de los blancos deseados aún en contextos de gran variabilidad genética. Por lo tanto, la hibridación inespecífica se ha aprovechado incluso para diseñar, en base a organismos cuyo genoma es conocido, microarreglos que detectan patógenos desconocidos y emergentes^{7,8}.

Hasta hace algunos años el obstáculo principal para utilizar enfoques de secuenciación era el costo alto (ahora ya relativamente estable), de la clásica y robusta técnica de Sanger, que sigue siendo el estándar de oro para un buen número de ensayos²². Sin embargo, el surgimiento de las tecnologías llamadas de “siguiente generación” (“Next Generation Sequencing”) ha resultado en un dramático e inesperado abaratamiento del costo por base secuenciada²³ que permite un análisis con gran “profundidad” o “cobertura” (parámetro que se ha constituido como fundamental para determinar la validez de un análisis particular utilizando estas nuevas tecnologías²⁴), mediante la secuenciación de cientos o miles de veces las regiones genómicas de interés.

Estas nuevas tecnologías tienen peculiaridades que exigen un examen cuidadoso para lograr su adaptación a requerimientos específicos de la práctica clínica. Estas peculiaridades incluyen el tamaño corto de las lecturas obtenidas, la variación en el tipo y tasa de error como consecuencia del uso de las diferentes plataformas comerciales existentes, y los largos tiempos por experimento que hasta ahora necesitan estas tecnologías^{1,25}.

Para aprovechar al máximo la capacidad masiva de secuenciación que ofrecen las nuevas plataformas es posible realizar el análisis simultáneo de múltiples muestras en la misma corrida del equipo (“multiplexing”). Esto se puede lograr mediante la introducción de secuencias etiqueta (“bar codes”)²⁶, oligonucleótidos diseñados de tal manera que identifican específicamente a una muestra y que se ligan a los ácidos nucleicos antes de analizarlos para que al obtenerse las secuencias pueda determinarse también a qué muestra pertenecen.

Además, existen técnicas que reducen la complejidad de las muestras a solamente un subconjunto de los ácidos nucleicos presentes y que pueden utilizarse para adaptar las nuevas tecnologías de secuenciación a las necesidades particulares de cada proyecto. Por ejemplo, en lugar de secuenciar el genoma completo de un individuo, se pueden “capturar” solamente los fragmentos

que contienen exones (regiones codificantes para proteína) en lo que se ha llamado secuenciación del “exoma”²⁷ que constituye menos del 2% del total de material genético humano. Análogamente, también es posible diseñar sistemas de captura o amplificación diferencial para otros subconjuntos de fragmentos del genoma que pueden resultar interesantes en un esquema analítico particular¹³.

La mayor ventaja de las técnicas de secuenciación sobre otros métodos es el hecho de que constituyen el máximo nivel de caracterización posible para una muestra a nivel de ADN o de ARN y por lo tanto, posibilitan la detección de todos los blancos genéticos con relevancia clínica, sin importar de qué tipo de blancos genéticos se trate y aún sin conocimiento previo de su existencia. Gracias a este hecho la secuenciación masiva ha sido utilizada exitosamente para determinar la causa de algunas enfermedades de origen genético desconocido que no habían podido ser diagnosticadas previamente^{28,29}.

Además, debido a que la identificación de los blancos buscados entre las secuencias obtenidas se realiza informáticamente en un paso posterior, es posible agregar la interpretación de nuevos genes y variantes importantes conforme sus efectos patológicos vayan siendo mejor descritos sin tener que modificar el procedimiento ya establecido de la prueba, haciendo de las pruebas basadas en secuenciación las más adaptables y evolucionables³⁰.

Por todas estas consideraciones puede inferirse que el empleo de técnicas de secuenciación de última generación va resultando cada vez más versátil y atractiva, pero uno de los principales factores que desaniman su uso es, como ya se ha mencionado anteriormente, su costo relativamente elevado por experimento y la complejidad del análisis de la inmensa cantidad de datos generada por estas plataformas.

Aunque además existen muchas otras tecnologías (normalmente especializadas en aplicaciones concretas y que utilizan alguna variante de los métodos de extensión o ligación de ácidos nucleicos para analizar una muestra³¹), nos hemos basado en las tecnologías de análisis mencionadas en los párrafos anteriores para proponer algunos elementos básicos que se pueden tomar como referencia para seleccionar un método de análisis genómico (Figura 3).

Figura 3. Características generales de los métodos de análisis genómico.

	Número de muestras (multiplex)	Variantes interrogadas	Tiempo de proceso	Análisis computacional	Límite de detección	Obtención de información nueva	Interferencia de variantes no buscadas
Secuenciación	Secuenciación	PCR	PCR	PCR	Secuenciación	Secuenciación	
PCR	Microarreglos	Microarreglos	Microarreglos	Microarreglos	Microarreglos	PCR	
Microarreglos	PCR	Secuenciación	Secuenciación	Secuenciación	PCR	Microarreglos	

Como se puede observar en la figura anterior, donde se resumen las características de los métodos de análisis genómico, por el momento no hay una tecnología que se prefiera en todos los casos, ya que cada una presenta ventajas y desventajas dependiendo del contexto particular y en cierto modo son complementarias entre sí.

4. ANTECEDENTES:

El desarrollo de nuevas pruebas diagnósticas o pronósticas mediante el uso de tecnologías de análisis genético y genómico surge de la necesidad de solventar las limitaciones de los métodos tradicionales, y mejorar entre otras cosas los tiempos diagnósticos, la sensibilidad y la especificidad de éstos.

Por ejemplo, los métodos fenotípicos para el diagnóstico de farmacorresistencia en infecciones por *Mycobacterium tuberculosis*, requieren de 6 a 9 semanas de cultivo previo, aún cuando se utilizan sofisticadas plataformas de automatización como BACTEC³². En cambio, los métodos genotípicos tales como secuenciación capilar, PCR en tiempo real, ensayos de sonda en línea (Line Probe Assay), etc. ofrecen la ventaja de no requerir el cultivo del bacilo y por lo tanto reducen el tiempo necesario para la detección de farmacorresistencia a sólo unos cuantos días u horas³³.

Los métodos de análisis genómico ofrecen también la oportunidad de mejorar los límites de detección, por ejemplo para el diagnóstico de tuberculosis por baciloscopía se requieren entre 5,000-10,000 bacilos por mililitro mientras que los métodos basados en PCR requieren menos de 100 bacilos por muestra³⁴.

Aunque estos avances son de gran utilidad, han surgido de las necesidades muy particulares de diferentes grupos que se desarrollan en el área clínica y de investigación, que normalmente adoptan soluciones más específicas y limitadas a la entidad que les ocupa. Por ejemplo, muchos de los métodos para la detección de farmacorresistencia en tuberculosis se enfocan a uno o dos antibióticos (rifampicina e isoniazida) dejando de lado muchas otras alternativas de primera y segunda línea para el tratamiento de esta enfermedad.

Tenemos la convicción de que los beneficios significativos para los sistemas de salud del país se obtendrán del desarrollo de plataformas analíticas que sean generales y flexibles en cuanto a su evolución en el tiempo y que aborden padecimientos que representen los más grandes problemas de salud pública para México.

Debido a esto, deseamos establecer una base de comparación que permita priorizar diferentes padecimientos según la importancia y pertinencia de desarrollar e implementar una prueba con tecnología genómica para ayudar a su manejo clínico, así como también según la magnitud de los beneficios potenciales de este desarrollo.

No tenemos conocimiento de trabajos previos que a nuestro entender busquen realizar una priorización de enfermedades respecto a la pertinencia del uso de tecnologías de análisis genómico para ayudar a su manejo clínico. Sin embargo, la evaluación de enfermedades respecto a su contribución a la pérdida de vida saludable ha sido utilizada desde un largo tiempo para fijar prioridades en programas de Salud Pública y proponer nuevas pruebas e intervenciones.

Tradicionalmente, esta contribución se ha calculado con base en información de mortalidad, y de la prevalencia e incidencia de las enfermedades, o utilizando diferentes indicadores que son específicos para cada padecimiento, lo que dificulta el cálculo estandarizado y comparable del impacto en salud pública de enfermedades diferentes.

Por esta razón en 1994 la Organización Mundial de la Salud describió detalladamente la metodología para el cálculo de un parámetro que estima el amplio impacto de las enfermedades en la salud en una población de forma estandarizada y comparable entre diferentes regiones: los Disability Adjusted Life Years (DALYs)³⁵.

Estos “Disability Adjusted Life Years”, o DALYs, buscan incorporar en la medida de lo posible todas las situaciones en salud que afectan el bienestar social en un solo indicador general, combinando en un mismo concepto los años de vida potencialmente perdidos por muerte prematura y el tiempo en años de vida “saludables” perdidos a causa de estados de salud pobre e incapacidad.

Técnicamente, los DALYs combinan el tiempo de vida en discapacidad y el tiempo perdido debido a mortalidad prematura. La duración del tiempo perdido a causa de mortalidad se calcula utilizando años de vida estándar perdidos utilizando una tabla de vida modelo West, con una expectativa de vida al nacimiento de 82.5 años para mujeres y 80 para hombres.

En el cálculo de los DALYs se otorga mayor importancia al tiempo vivido en edad productiva mediante la utilización de una función exponencial de la forma $Ce^{-\beta x}$ (con un valor de $\beta=0.04$ y $C=0.16243$) que concuerda con el cambio en el nivel de dependencia con la edad, mientras que los resultados futuros se descuentan con una tasa del 3%, a través de una función continua de la forma $e^{-r(x-a)}$, donde r es la tasa de descuento y a es la edad de aparición.

Por su parte, el tiempo vivido con discapacidad se multiplica por un factor que refleja la severidad de la misma en una escala de 0 (saludable) a 1 (muerte), dichos factores de ponderación son arbitrarios, aunque fueron seleccionados por consenso de dos grupos de especialistas en cada una de las áreas del cuidado de la salud.

La fórmula general para el cálculo de los DALYs perdidos por individuo es:

$$\int_{x=a}^{x=a+L} DCxe^{-\beta x} e^{-r(x-a)} dx = - \left[\frac{DCe^{-\beta a}}{(\beta+r)^2} \left[e^{-(\beta+r)(L)} (1 + (\beta+r)(L+a)) - (1 + \beta+r)a \right] \right]$$

Donde D es el factor de severidad de discapacidad (1 en el caso de muerte prematura), r es la tasa de descuento (0.03), C es la constante de corrección por edad (0.16243), β es el parámetro de la

función de corrección por edad (0.04), a es la edad de aparición en el individuo y L es la duración de la discapacidad o el tiempo perdido debido a muerte prematura.

5. HIPÓTESIS:

Utilizar una estrategia de priorización de enfermedades es de utilidad para proponer pruebas con tecnología de análisis genómico tanto importantes para la salud pública como consideradas de utilidad por la práctica clínica en México.

6. OBJETIVOS:

6.1 Objetivo principal:

Desarrollar y validar un prototipo de prueba con tecnología de análisis genómico enfocada al manejo clínico de una de las enfermedades prioritarias para la salud pública de México.

6.2 Objetivos secundarios:

Definir un esquema que permita priorizar las enfermedades de la población mexicana y evaluar la pertinencia del desarrollo de pruebas genómicas para su manejo clínico.

Proponer una estrategia para desarrollar pruebas genómicas útiles para la práctica clínica en México y prioritarias para la salud pública del país.

7. METODOLOGÍA:

7.1 Estrategia:

En este proyecto proponemos y aplicamos una estrategia para el desarrollo de nuevas pruebas con tecnología de análisis genómico que sean consideradas prioritarias para la salud pública de México y simultáneamente sean apreciadas como útiles por las instancias actuales del cuidado de la salud. Dicha estrategia se basa en la priorización de diferentes enfermedades o grupos de padecimientos respecto a su importancia en salud pública y la pertinencia del ámbito genómico en su manejo clínico, el análisis de las necesidades actuales más apremiantes de los servicios de salud dedicados a su cuidado y la posterior definición de un esquema de proceso específico que cubra estas necesidades de la mejor manera posible.

En primer lugar definimos un grupo de parámetros y proponemos guías para su evaluación de tal forma que nos permitan priorizar diferentes padecimientos según las necesidades en la Salud Pública de México, el potencial para desarrollar e implementar una prueba con tecnología genómica para su diagnóstico o manejo clínico y la magnitud de los posibles beneficios de salud que podrían obtenerse de dicho desarrollo.

Utilizando los parámetros propuestos y con base en investigación bibliográfica y la opinión de expertos en las áreas clínica y de investigación, evaluamos la viabilidad y pertinencia del desarrollo de sistemas analíticos en el área de la genómica que asistan en el manejo clínico de los padecimientos con mayor impacto en la Salud Pública de México.

A partir de esta evaluación seleccionamos una enfermedad relevante para la salud pública de México, que presenta un adecuado potencial para desarrollar e implementar una prueba con tecnología genómica que apoye en su manejo clínico (ya sea diagnóstico, pronóstico, predisposición o selección de tratamiento).

Como siguiente paso analizamos los requerimientos específicos del ámbito clínico en una institución de los servicios de salud mexicanos respecto a la enfermedad prioritaria seleccionada y proponemos una nueva prueba con tecnología genómica, definiendo un esquema de proceso particular que cubra las necesidades del ámbito clínico.

Habiendo utilizado esta estrategia para desarrollar una nueva prueba genómica específicamente para la detección de resistencia a los principales fármacos utilizados en el tratamiento de infecciones por *Mycobacterium tuberculosis* mediante secuenciación de nueva generación, nos dedicamos posteriormente a evaluar la capacidad analítica que ésta presenta mediante la secuenciación de acuerdo a nuestro protocolo (con la plataforma PGM de Ion Torrent) de un grupo de 16 muestras de *Mycobacterium tuberculosis* previamente caracterizadas como susceptibles y resistentes a diversos antibióticos importantes en el Instituto Nacional de Enfermedades Respiratorias (INER).

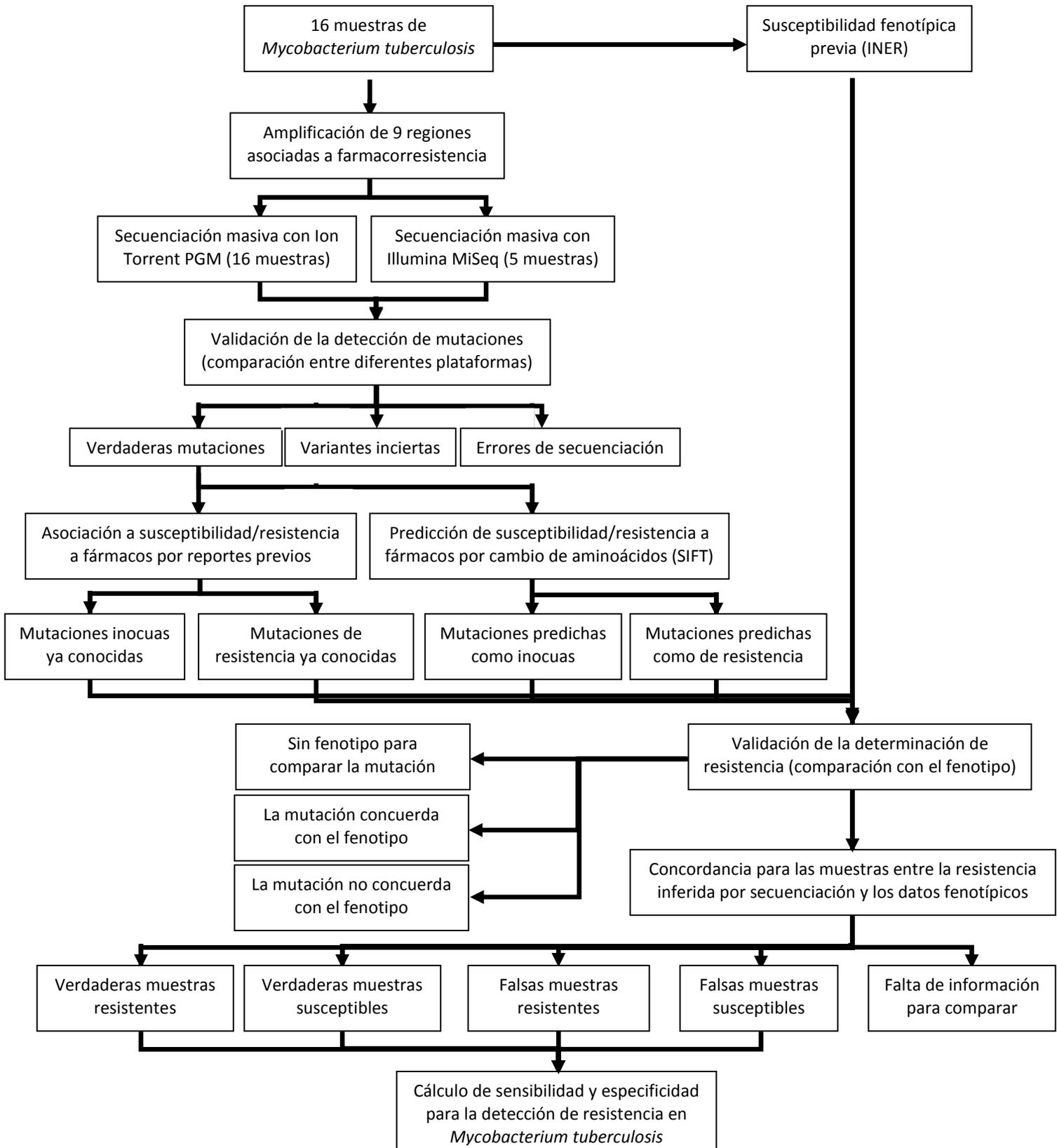
Como primer punto validamos la detección de mutaciones mediante nuestro protocolo de secuenciación con el PGM de Ion Torrent mediante una comparación con los resultados obtenidos a través de otra plataforma de secuenciación diferente (MiSeq de Illumina). Después realizamos también una evaluación de la determinación de susceptibilidad o resistencia a diferentes antibióticos con base en las mutaciones detectadas mediante secuenciación masiva y por último evaluamos la concordancia entre la resistencia inferida mediante secuenciación masiva para cada una de las muestras con los resultados de la caracterización fenotípica realizada previamente en el INER.

Los pasos seguidos durante la realización del proyecto, así como también las actividades necesarias para cada uno de ellos que se mencionaron anteriormente se muestran de forma gráfica en en las figuras 4 y 5.

Figura 4. Diagrama de flujo del proyecto.



Figura 5. Diagrama de flujo de la validación analítica de la prueba propuesta



7.2 Priorización de enfermedades.

7.2.1 Definición de parámetros para la priorización de enfermedades.

En este trabajo se desea establecer una base de comparación que permita priorizar diferentes padecimientos en cuanto a la importancia y pertinencia, en un plazo corto, de desarrollar e implementar una prueba con tecnología genómica para ayudar a su manejo clínico. Así, la base de comparación propuesta se encuentra conformada por varios parámetros que, en conjunto, toman en cuenta lo siguiente: las mayores necesidades en salud pública de México, el mayor potencial para desarrollar y establecer una prueba de tecnología genómica y los mayores beneficios de salud que se podrían obtener de este desarrollo.

Los parámetros de priorización propuestos son:

1. "Disability Adjusted Life Years" (DALYs) por cien mil habitantes: Un indicador ampliamente aceptado sobre el costo individual y social del problema de salud causado por el padecimiento que se está estudiando. Es una medida del impacto de las enfermedades sobre la población mexicana que resulta ventajosa gracias a que representa los años de vida saludables que se pierden debido a la mortalidad, pero también debido a la discapacidad provocada por una enfermedad en particular³⁵. Su valor numérico es igual a los años de vida saludables perdidos en un año por cada cien mil habitantes que componen la población de dicho país en ese año.

Utilizando este parámetro se puede obtener una idea de que tan grave es el problema de salud causado por el padecimiento que se está estudiando y así darles mayor peso a aquellos que representen un problema más serio.

2. Factibilidad en el corto plazo: Este parámetro se refiere a la disponibilidad de los recursos científicos necesarios para emplear un enfoque genómico en la determinación del riesgo de presentar la enfermedad, su diagnóstico o la selección de un tratamiento para ella dentro de un periodo corto de tiempo.

El valor de este parámetro es fundamental debido a que permite identificar áreas donde se hayan descubierto y corroborado asociaciones a blancos genéticos, indicándonos aquellas en las cuales exista más claridad acerca de la interpretación de su presencia, es decir, aquellas donde se cuente actualmente con una base sólida de "casos de éxito" o las muy necesarias "pruebas de campo" a partir de las cuales se pueda emprender el desarrollo de plataformas con tecnologías de análisis genómico, evitando así crear falsas expectativas sobre el valor concreto de la medicina genómica en la práctica clínica de aquellos padecimientos donde todavía es necesario invertir varios años más de investigación.

3. Efecto esperado: Se refiere a la magnitud de los beneficios potenciales que se obtendrían al desarrollar pruebas con tecnologías de análisis genómico para un padecimiento en particular.

Los beneficios que se pueden obtener a partir del desarrollo de pruebas genómicas están determinados por la medida en la cual éstas sean actuables clínicamente, es decir, que existan acciones preventivas o correctivas basadas en ellas que puedan tomarse para evitar o reducir las secuelas del padecimiento (lo cual no siempre es el caso).

7.2.2 Guías para la determinación del valor de cada parámetro.

Una vez que los parámetros han sido definidos es necesario contar con ciertas guías básicas que nos ayuden a asignarle un determinado valor a cada parámetro para algún padecimiento en particular, aspecto que se cubre en esta sección.

Determinación del valor de DALYs por cien mil habitantes.

La Organización Mundial de la Salud ha realizado esfuerzos en coordinación con todos los países que pertenecen a ella (como es el caso de México) para calcular los DALYs de gran variedad de padecimientos de forma estandarizada en el marco del proyecto “Global Burden of Disease” (GBD). La determinación del número de DALYs por cien mil habitantes que corresponde a una enfermedad en particular es bastante directa debido a que los resultados más actuales de estos esfuerzos se encuentran públicamente disponibles en el sitio web del proyecto GBD³⁶.

Determinación del valor de la “Factibilidad en el corto plazo”.

No existe una medida cuantitativa para un parámetro como éste, a diferencia del caso de los DALYs por cien mil habitantes. Por ello decidimos utilizar una escala de cuatro valores posibles para la factibilidad en el corto plazo que, ordenados de mayor a menor, definimos como: “Segura”, “Probable”, “Difícil” y “No identificada”.

Utilizando estos cuatro valores podemos abarcar desde aquellos casos en que la enfermedad ya haya sido estudiada extensivamente y se haya realizado mucha investigación en el desarrollo de pruebas de este tipo, donde probablemente sólo sea necesario adaptar pruebas existentes a las características de la población mexicana, hasta aquellos casos donde no existan factores genéticos claramente identificados que pudieran estar involucrados en las características clínicas de la enfermedad.

Para determinar cuál de los cuatro valores posibles para este parámetro le corresponde a un padecimiento particular utilizamos un diagrama guía basado en unas pocas preguntas (Figura 6).

La primera pregunta mostrada en el diagrama se centra en cuestionar la existencia de información acerca de asociaciones entre blancos genéticos y la enfermedad que se está evaluando, ya sea que

los blancos modifiquen la probabilidad de presentar la enfermedad, indiquen el mejor o peor pronóstico de unos pacientes sobre otros o una mayor o menor efectividad del tratamiento. En caso de no haberse encontrado blancos genéticos asociados al problema de salud, resulta poco factible desarrollar una prueba con tecnología genómica para su manejo clínico en un plazo breve, y por lo tanto, la factibilidad en el corto plazo correspondiente sería “No identificada”.

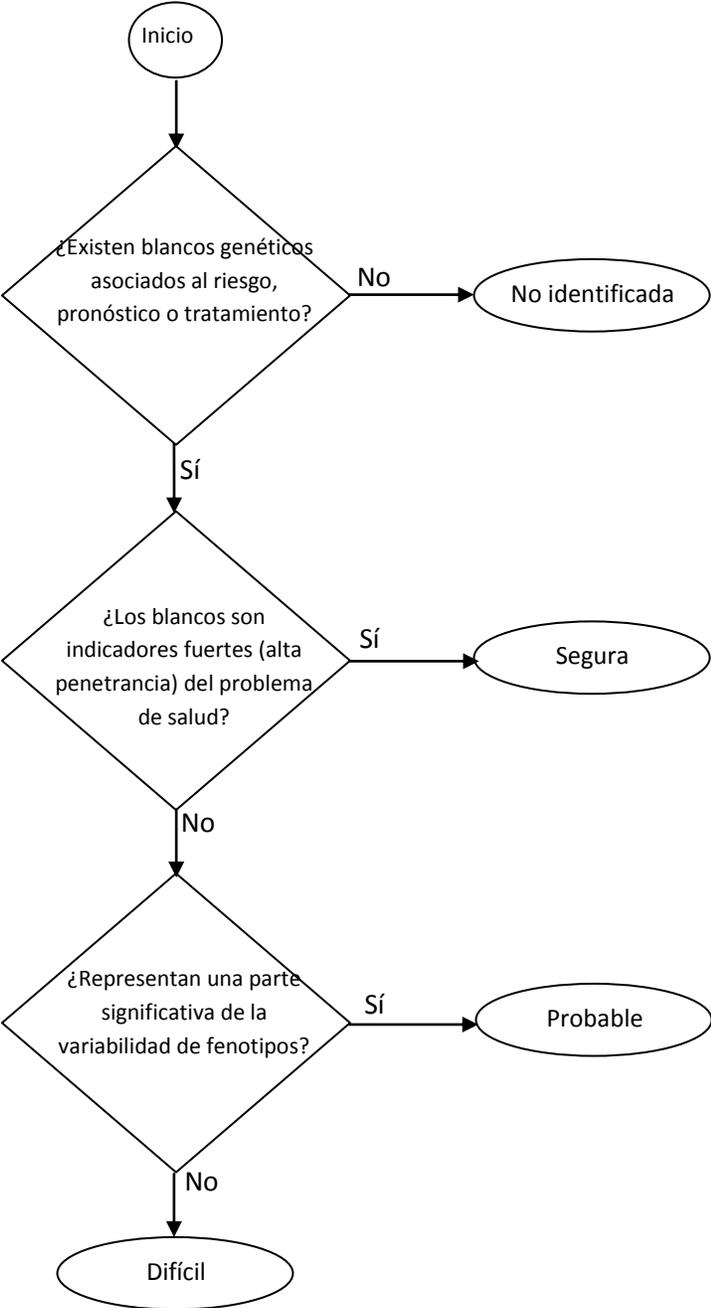
Las respuestas a las siguientes preguntas que formulamos en el diagrama están determinadas por la magnitud de las asociaciones que se presentan entre el padecimiento y los blancos genéticos relacionados, si es que las hay.

Si la presencia de los blancos genéticos se puede utilizar como indicador sólido o seguro de la presencia de la enfermedad ya que (debido a su alta penetrancia) cuentan con una buena correlación con los fenotipos importantes del padecimiento, es posible asignar un valor de “Segura” a la factibilidad en el corto plazo del problema de salud que se está evaluando. Algunos blancos que presentan esta correlación podrían ser, por ejemplo, niveles de expresión de genes muy particulares, un alto título de patógeno o llegando incluso a variantes con un patrón de herencia mendeliana. En caso de que los blancos genéticos no presenten tan buena correlación pero que sí expliquen una parte significativa de la variabilidad de los fenotipos relacionados a un problema de salud (en cuanto a susceptibilidad, pronóstico o efectividad de un tratamiento) entre los diferentes individuos, definimos una factibilidad “Probable”, en el corto plazo, de dicho padecimiento.

Cuando dentro de las áreas antes mencionadas sólo existen blancos genéticos con pequeños efectos, por ejemplo con razón de posibilidades (“odds ratio”; OR) muy cercanas a la unidad, será más complicado encontrar una combinación adecuada de ellos que logre aportar una utilidad clínica importante, en especial en un corto tiempo y, por lo tanto correspondería a una factibilidad en el corto plazo de “Difícil”.

Es muy importante tener en cuenta que los problemas de salud que actualmente tienen una baja factibilidad en el corto plazo podrían en el futuro (conforme vaya avanzando la investigación científica de dichos padecimientos) ser vulnerables al uso de tecnologías genómicas para ayudar a su manejo clínico.

Figura 6. Diagrama guía para determinar la factibilidad en el corto plazo.



Determinación del valor del “Efecto esperado”.

Para éste parámetro se proponen tres valores posibles, que en orden descendente son: “Impacto drástico”, “Impacto medio” e “Impacto marginal”. Decidimos también utilizar un diagrama guía que consta de unas pocas preguntas generales para determinar el valor correspondiente al efecto esperado de un determinado problema de salud (Figura 7).

Como referente hemos considerado que las aplicaciones de pruebas con tecnología genómica con beneficios más importantes a corto plazo se sitúan en la transición hacia una medicina más preventiva, como son el área de diagnóstico temprano, donde se permitiría la detección de una enfermedad en una etapa presintomática o poco avanzada (para iniciar el manejo clínico del paciente y evitar así complicaciones que dificulten y hagan más costoso su tratamiento) así como también el área de estratificación de pacientes, donde se podría identificar a los individuos con una alta susceptibilidad a los padecimientos (para minimizar su aparición a través de medidas preventivas adecuadas) e identificar *a priori* la efectividad y probabilidad de reacciones adversas que tendrá algún tratamiento específico (para seleccionar el tratamiento más adecuado para cada paciente en particular).

Debido a lo anterior, en el diagrama se pregunta acerca de la existencia de beneficios potenciales ante un diagnóstico más temprano, que son de gran importancia en enfermedades crónicas cuya etapa temprana puede cursar en forma asintomática y que podrían ser mucho más limitados en aquellas agudas y cuyo periodo libre de síntomas sea muy corto. Un ejemplo de estas enfermedades cuya etapa temprana cursa de manera asintomática y que por lo tanto se beneficiarían de un diagnóstico en una etapa más temprana es el caso de diabetes mellitus tipo 2 donde alrededor de 30 a 50% de las personas afectadas desconocen su problema por meses o años, lo que condiciona a que la mayoría de los pacientes ya hayan iniciado daño macrovascular y entre 25 y 40% neuropatía diabética al momento del diagnóstico³⁷.

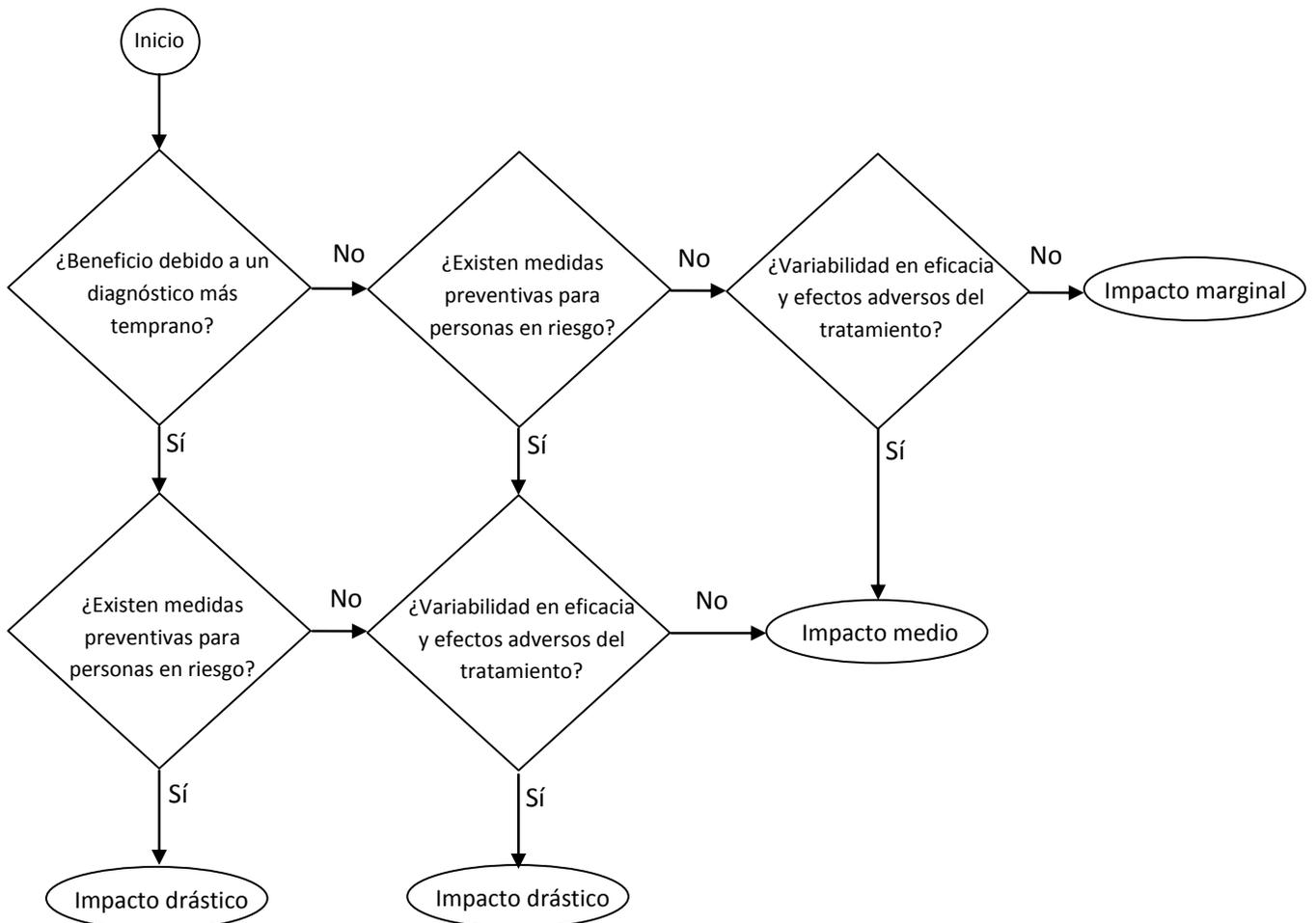
Para determinar si, a partir de identificar a los individuos más susceptibles a un padecimiento, podrían obtenerse beneficios para el sector salud, se pregunta en nuestro diagrama si existen medidas preventivas recomendadas para personas en riesgo de sufrirlo. En caso de que sí existan este tipo de medidas (que podrían ayudar a disminuir la severidad e incluso evitar por completo el desarrollo de la enfermedad) los beneficios potenciales son mayores que en el caso contrario.

Se pregunta después si existe una variabilidad evidente en la eficacia de tratamiento o en el riesgo de reacciones adversas entre algunos grupos de pacientes respecto a otros. De ser así podría haber beneficios importantes en el desarrollo de una prueba genómica, mientras que en otros casos donde se cuente con tratamientos más estandarizados y con menos efectos secundarios los beneficios serán más modestos. Además, hemos considerado que cuando se puedan obtener beneficios en al menos dos de estas áreas el efecto esperado de desarrollar una prueba genómica para un padecimiento particular corresponde a un “Impacto drástico”, mientras que si sólo existen beneficios potenciales en una de ellas se esperaría un “Impacto medio” como efecto esperado, y

en caso de que no existan beneficios claros a partir de alguna de estas aplicaciones se debería asignar un valor de “Impacto marginal” a este parámetro.

Este parámetro es de gran importancia ya que representa, en cierta manera, los máximos beneficios para los servicios de salud del país que podrían obtenerse mediante el desarrollo de pruebas genómicas para afrontar la problemática de una enfermedad determinada.

Figura 7. Diagrama guía para determinar el efecto esperado



7.2.3 Grupo de padecimientos a evaluar.

Se consideraron como padecimientos convenientes a evaluar los grupos de causas definidos para el cálculo de los DALYs dentro de la iniciativa “Global Burden of Disease” de la Organización Mundial de la Salud (OMS).

Se ordenaron las causas a partir de los datos de DALYs por cada cien mil habitantes para México en el año 2004 (los más recientes con los que se cuenta a la fecha de realización de este trabajo) y se limitó el grupo de enfermedades a evaluar a aquellas quince que presentaran un mayor valor (Tabla 1). De esta manera se asegura el considerar en esta fase del proyecto los problemas de salud que afectan en mayor medida la vida de la población. Cabe mencionar, además, que se decidió desagregar los padecimientos de diabetes mellitus en dos categorías (subtipos monogénicos y poligénicos) debido a que por sus diferencias intrínsecas hacían imposible asignar un mismo y único valor a cada parámetro si se les consideraba parte de un mismo conjunto.

Tabla 1. Quince causas con mayor número de DALYs por cien mil habitantes para México.

Causas	DALYs/100,000 hab*
Heridas no intencionales	1356
Diabetes mellitus	801
Neoplasias malignas	789
Trastornos depresivos unipolares	772
Anomalías congénitas	713
Enfermedad isquémica del corazón	538
Heridas intencionales	519
Cirrosis hepática	443
Prematurez y bajo peso al nacer	376
Infecciones de las vías respiratorias inferiores	354
Asma	328
Asfixia y trauma del nacimiento	327
Infecciones neonatales	322
Desórdenes endócrinos	293
Esquizofrenia	277

*Valores obtenidos de la tabla: Estimated DALYs per 100,000 population by cause, and Member State, 2004 dentro de las estadísticas de estimados de lesiones y enfermedad por país del “Burden of Disease Project”. http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/index.html

7.3 Propuesta de una prueba genómica para la enfermedad seleccionada

Para proponer una nueva prueba enfocada en el manejo clínico de una enfermedad prioritaria que la práctica clínica considere útil es necesario conocer las necesidades específicas de las instituciones encargadas del cuidado de los individuos afligidos específicamente por esos padecimientos.

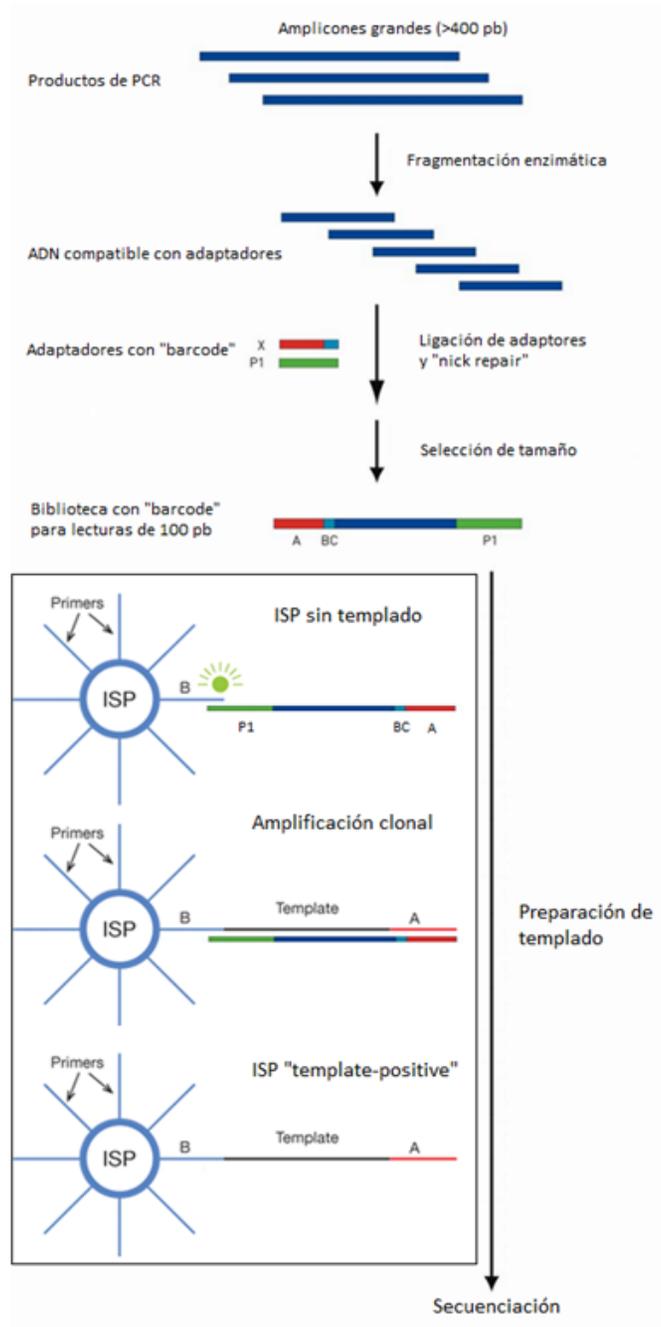
Por esta razón analizamos los requerimientos específicos del ámbito clínico respecto a las infecciones de las vías respiratorias inferiores (el padecimiento seleccionado en base a la priorización de enfermedades) a través de la interacción con el laboratorio de microbiología clínica del Instituto Nacional de Enfermedades Respiratorias “Ismael Cosío Villegas” y seleccionamos en base a dichos requerimientos las alternativas más adecuadas con que se cuenta para cada una de las características fundamentales en el diseño de una prueba con tecnología genómica (presentadas en la sección de introducción de este trabajo).

Con base en este análisis seleccionamos las tecnologías de secuenciación de nueva generación como método de análisis genómico, para permitir la identificación de mutaciones asociadas a la resistencia ante algunos de los principales antibióticos utilizados en el tratamiento de tuberculosis (isoniazida, rifampicina, etambutol, pirazinamida, amikacina, estreptomycin y fluoroquinolonas como cirpofloxacin y ofloxacin) de forma simultánea.

El esquema de secuenciación se planteó en la plataforma PGM (Ion Torrent) y requiere de la amplificación por PCR de nueve regiones del genoma de *Mycobacterium tuberculosis* que se encuentran asociadas a farmacorresistencia, la preparación de bibliotecas de secuenciación con longitud de 100 pb³⁸ y su posterior secuenciación en el equipo.

A su vez, el proceso de preparación de bibliotecas consiste en mezclar de manera equimolar los productos de PCR correspondientes a una misma muestra, su fragmentación enzimática a un tamaño cercano a 100 pb, la adición de adaptadores mediante ligación (donde se incluye la región “barcode” que identificará a dicha muestra). Posteriormente, se hace una selección de tamaño para conservar sólo los fragmentos de aproximadamente 100-200 pb los cuales se someten a una PCR en emulsión para su unión a las perlas de soporte sobre las que se lleva a cabo la reacción de secuenciación (Figura 8).

Figura 8. Esquema de proceso de preparación de bibliotecas de secuenciación con IonTorrent a partir de amplicones.



7.4 Validación analítica de la prueba propuesta

Para evaluar el desempeño de la prueba se utilizó el esquema de secuenciación masiva propuesto (descrito a detalle en el Anexo 4) para analizar la resistencia a diversos antibióticos en muestras de *Mycobacterium tuberculosis*. Como una primera aproximación, se decidió que las muestras utilizadas en esta sección fueran cultivos de *Mycobacterium tuberculosis* provenientes de una colección aislada previamente en el Instituto Nacional de Enfermedades Respiratorias. Ya que esto nos permitió realizar una comparación de los resultados de la secuenciación con la susceptibilidad y resistencia fenotípica a los antibióticos de interés, que ya ha sido caracterizada para estas muestras de forma previa en el INER.

Otro aspecto particular de esta validación es el hecho de que debido a la disponibilidad de reactivos, las muestras fueron procesadas en dos grupos, primero tres muestras en un mismo chip de secuenciación y después 13 en otro, para alcanzar un total de 16 muestras, incluyendo cepas resistentes a cada uno de los antibióticos (isoniazida, rifampicina, etambutol, pirazinamida, amikacina, estreptomina y alguna fluoroquinolona (ofloxacina o moxifloxacina).

De cada una de estas muestras se realizó la amplificación de los genes de resistencia y la secuenciación masiva de dichos amplicones, para después comparar las secuencias obtenidas con el genoma de referencia de *Mycobacterium tuberculosis H37Rv* y determinar de esta manera las mutaciones presentes en cada muestra, tal como se propone en el Anexo 4 y se resume a continuación.

7.4.1 Amplificación por PCR.

Para llevar a cabo la amplificación de las nueve regiones del genoma de *Mycobacterium tuberculosis* asociadas a farmacorresistencia se prepararon de manera independiente mezclas de reacción de 25 μ L con las siguientes concentraciones de reactivos: 20 mM Tris-HCL [pH 8.4], 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM dNTPs y "primers", 0.025 u/ μ L de polimerasa Taq (Invitrogen), 0.5 μ M de cebadores específicos correspondientes a cada región (Anexo 2), y >6 ng de DNA molde. Las mezclas de reacción se colocaron en un termociclador (GeneAmp PCR System 9700 Applied Biosystems) y se sometieron a un protocolo de termociclado consistente de una desnaturalización inicial a 95°C por 15 minutos, 35 ciclos de desnaturalización a 95 °C por 30s, alineamiento a 60 °C por 30 s y elongación 72 °C por 30 s, seguidos de una elongación final 72 °C durante 7 minutos. Para verificar la correcta amplificación de las reacciones se realizó una electroforesis en gel de agarosa al 2% utilizando el sistema E-gel EX (Invitrogen) según las instrucciones del fabricante.

7.4.2 Purificación de los productos de PCR

Para la purificación de los productos de PCR por inmovilización reversible en fase sólida sobre perlas magnéticas: se agregaron 25 μL de perlas magnéticas AMPure XP (Agencourt) a 20 μL de cada producto de PCR, se incubó la mezcla por 15 min a temperatura ambiente, después se colocó en un soporte magnético para remover el sobrenadante y se efectuaron dos lavados con etanol al 80%, se secaron las perlas a temperatura ambiente por 10 min y finalmente, se eluyó el DNA en 25 μL de agua libre de nucleasas.

7.4.3 Cuantificación de los productos de PCR y preparación de una mezcla equimolar

Los amplicones purificados fueron cuantificados utilizando el fluorómetro Qubit con un kit dsDNA High Sensitivity (Invitrogen), siguiendo las instrucciones del fabricante. Utilizando los valores de concentración obtenidos se calcularon los volúmenes necesarios para obtener una mezcla equimolar de todos los amplicones (conteniendo 0.5 pmol de cada amplicón) y se prepararon dichas mezclas.

7.4.4 Preparación de bibliotecas y secuenciación con Ion Torrent PGM

Se prepararon bibliotecas de secuenciación para lecturas de 100 pb con barcode mediante fragmentación enzimática a partir de 100 ng totales de cada mezcla equimolar de amplicones, utilizando los reactivos Ion Shear Plus y Ion Xpress Plus (Life Technologies), siguiendo las instrucciones del fabricante. La preparación y enriquecimiento del templado se llevó a cabo utilizando el sistema automatizado Ion OneTouch con Enricher System y fue secuenciado en el secuenciador Ion Torrent PGM, según las instrucciones del fabricante.

7.4.5 Validación de la detección de mutaciones

Para comprobar que las mutaciones detectadas a través de la secuenciación se encuentran verdaderamente presentes en las muestras y no se deben solamente a errores característicos de la plataforma seleccionada (Ion Torrent PGM) se decidió, además, secuenciar cinco de las muestras a gran profundidad de cobertura utilizando una plataforma con tecnología diferente: el secuenciador MiSeq de la compañía Illumina.

Se prepararon bibliotecas de secuenciación para lecturas “paired end” de 2x150 pb con barcode a partir de 50 ng totales de cada mezcla equimolar de amplicones, utilizando los reactivos de Nextera, y fueron secuenciadas en el secuenciador MiSeq (Illumina) siguiendo las instrucciones del fabricante.

Al realizar la comparación entre las variantes detectadas con ambas tecnologías de secuenciación, consideramos que aquellas que fueron detectadas por ambas plataformas representan mutaciones reales que se encuentran presentes en las muestras, mientras que aquellas que fueron detectadas sólo por una de las plataformas podrían ser errores específicos provenientes del procesamiento y tecnología que utiliza cada una de ellas, ya sea el MiSeq o el Ion Torrent PGM.

7.4.6 Validación de la detección de resistencia a antibióticos mediante secuenciación.

Se realizó el control de calidad de la corrida de secuenciación utilizando la aplicación FASTQC (Babraham Institute)³⁹ y se visualizaron los archivos BAM en el Integrative Genomics Viewer (Broad Institute)⁴⁰ para analizar la cobertura de las regiones y curar manualmente el llamado de variantes presentado por Variant Caller plug in de Ion Torrent en el archivo VCF.

Después se infirió la resistencia a fármacos de cada muestra a partir de la presencia de mutaciones previamente reportadas en fuentes de información selectas^{41,42} o se realizó una predicción con base en el análisis de la tolerancia a los cambios de aminoácidos mediante SIFT⁴³ si éstas no habían sido reportadas previamente.

7.4.7 Concordancia con datos fenotípicos

Como paso final comparamos la susceptibilidad o resistencia a los fármacos utilizados en el tratamiento de tuberculosis que inferimos a través del análisis de secuenciación masiva con los resultados de las pruebas de sensibilidad fenotípica a los antibióticos realizadas previamente en el Instituto Nacional de Enfermedades Respiratorias.

8. Resultados:

8.1 Priorización de enfermedades

8.1.4 Evaluación de las enfermedades

Utilizando en conjunto los parámetros antes descritos pueden compararse diferentes padecimientos y así obtener una idea de para cuáles de ellos sería más pertinente desarrollar pruebas genómicas con fines de diagnóstico o manejo clínico.

Específicamente, el padecimiento que cuente con un valor elevado en todos los parámetros establecidos será el objetivo más conveniente para el desarrollo de pruebas genómicas. Un gran número de años de vida saludables perdidos nos indica que el padecimiento es un problema importante de salud pública, un valor elevado de factibilidad en el corto plazo nos dice que es posible diseñar pruebas genómicas y comenzar a recibir beneficios en poco tiempo y, además, un valor alto de efecto esperado nos habla de que los beneficios que se pueden obtener de dichas pruebas son considerables.

Para llevar a cabo la evaluación de las enfermedades anotadas en la tabla 1 se realizó, como primer paso, una investigación bibliográfica con el objetivo de obtener un panorama general de los blancos genéticos asociados a las enfermedades. Dicha investigación bibliográfica se centró en artículos recientes (prefiriendo aquellos publicados en los últimos dos años) que tratan sobre los componentes genéticos de cada una de las enfermedades a evaluar, sin embargo no pretendió ser una revisión exhaustiva de la literatura acerca de este tema.

Además de la investigación bibliográfica se tomó en consideración también la opinión de expertos tanto en el área clínica como de investigación acerca del uso de la genómica para el diagnóstico, evaluación del riesgo o selección de tratamiento de algunas enfermedades. Dicha opinión fue recabada a partir de la participación de estos expertos en varios Núcleos de Discusión Académica realizados a lo largo del 2011 en el Instituto Nacional de Medicina Genómica donde se discutieron las posibilidades actuales y futuras de las técnicas de análisis genómico en el manejo clínico de las enfermedades.

Toda esta información fue utilizada para asignar valores específicos a los parámetros de priorización para los distintos grupos de padecimientos seleccionados, utilizando como referencia las guías presentadas en la sección anterior de este trabajo (Anexo 1). Los resultados de esta asignación de valores se muestran a continuación en la tabla 2.

Tabla 2. Valores asignados a los parámetros de priorización para cada padecimiento.

Padecimientos evaluados	DALYs/100,000 hab	Factibilidad en el corto plazo	Efecto esperado
Heridas no intencionales	1356	No identificada	Medio
Diabetes mellitus (subtipos poligénicos)	789	Probable	Drástico
Neoplasias malignas	789	Probable	Drástico
Trastornos depresivos unipolares	772	Difícil	Drástico
Anomalías congénitas	713	Segura	Medio
Enfermedad isquémica del corazón	538	Probable	Drástico
Heridas intencionales	519	Probable	Medio
Cirrosis hepática	443	Probable	Drástico
Prematurez y bajo peso al nacer	376	No identificada	Medio
Infecciones de las vías respiratorias inferiores	354	Segura	Drástico
Asma	328	Probable	Drástico
Asfixia y trauma del nacimiento	327	No identificada	Medio
Infecciones neonatales	322	Segura	Drástico
Desórdenes endócrinos	293	Probable	Drástico
Esquizofrenia	277	Probable	Drástico
Diabetes mellitus (subtipos monogénicos)	12	Segura	Drástico

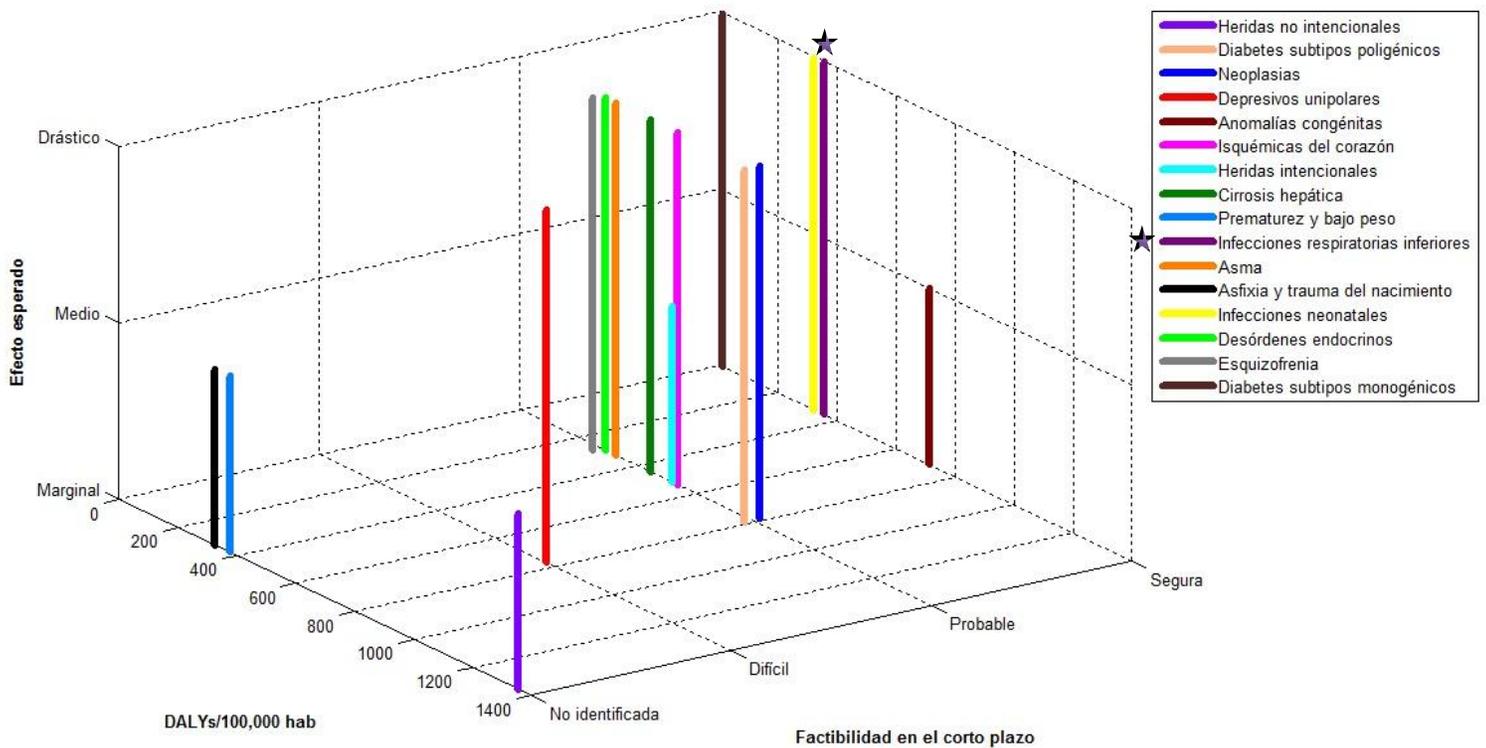
8.1.5 Selección de una enfermedad prioritaria

Si comparamos los valores obtenidos para los tres parámetros de priorización (DALYs/100,000 hab., factibilidad en el corto plazo y efecto esperado) otorgándoles igual importancia a todos ellos podemos obtener un gráfico que nos ayude a identificar los grupos de padecimientos prioritarios (Figura 9).

En dicha gráfica podemos observar algunos aspectos importantes que ilustran la naturaleza del análisis objeto de este trabajo:

- El primero de ellos es que para los padecimientos con un mayor impacto en la vida de la población mexicana, es decir, las heridas no intencionales (accidentales) no hay factibilidad previsible de que las pruebas genómicas tengan algún impacto en su prevención o atención en el corto plazo.
- Es ilustrativo también el caso de las enfermedades de genética Mendeliana, entre los que se encuentran los subtipos monogénicos de diabetes, y donde muchas veces se conocen los genes y variantes genéticas que determinan claramente la presencia de la enfermedad⁴⁴. Por lo tanto, es posible desarrollar fácilmente pruebas que identifiquen estos padecimientos pero, al afectar a una baja proporción de la población (en su conjunto estos subtipos sólo representan entre el 1 y 2% de todos los casos de diabetes⁴⁵), su impacto en salud pública medido a través de los años de vida saludables perdidos es marginal.

Figura 9. Gráfica comparativa de las enfermedades según los valores asignados a los parámetros de priorización.



Resulta evidente entonces que los padecimientos más interesantes para el enfoque genómico son aquellos que presentan a la fecha valores más balanceados para los parámetros: así, en la zona central de la gráfica, encontramos algunos padecimientos que son reconocidos actualmente por la comunidad científica, los medios de comunicación y el público en general como los problemas de salud más grandes de México, pero también otros que podrían ser menos evidentes.

Dentro de estos padecimientos que resultan particularmente atractivos para el enfoque genómico se encuentran las enfermedades neoplásicas, donde un gran interés de la comunidad científica internacional ha llevado al descubrimiento de blancos genéticos asociados a una mayor predisposición para presentar ciertos tipos de cáncer, de su potencial metastásico y de algunos otros asociados a la respuesta ante algunos fármacos antineoplásicos⁴⁶. Los resultados del continuo y creciente interés de la comunidad científica y médica están permitiendo establecer números crecientes de pruebas comerciales útiles para anticipar la efectividad de fármacos o ayudar al pronóstico en algunas poblaciones, y será relativamente accesible tomar estas experiencias como base para desarrollar nuevas pruebas que respondan mejor a las necesidades de nuestro país. Hasta ahora, las pruebas genéticas y genómicas para estos padecimientos se ofrecen de manera individual, debido a la gran heterogeneidad que existe en cuanto a la presencia

de mutaciones, tanto entre diferentes tipos de cáncer como entre diferentes tumores de los mismos linajes celulares^{47,48}. Sin embargo, recientemente se han propuesto programas para el desarrollo de esquemas más generales para la detección de mutaciones en un amplio panel de genes asociados de manera importante a la respuesta ante diversos fármacos antineoplásicos (donde podría utilizarse la misma prueba para analizar varios tipos de tumores diferentes)⁴⁹.

Otro de los grupos más pertinentes de enfermedades lo constituyen los subtipos poligénicos de diabetes ya que un buen número de blancos genéticos han sido asociados a la susceptibilidad de presentar estos padecimientos. Sin embargo, su utilidad clínica todavía es incierta cuando se compara con el poder discriminatorio de los factores de riesgo utilizados actualmente en la práctica clínica (tales como los antecedentes familiares de la enfermedad o el índice de masa corporal) que, en general, son más fáciles de evaluar^{50,51}. Por otro lado, también han sido identificados blancos genéticos asociados a una respuesta mayor o menor al tratamiento farmacológico, pero estos se limitan a dos de las nueve grandes clases de fármacos antidiabéticos, dejando todavía incompleto un esquema de selección del tratamiento para este tipo de padecimientos^{52,44}.

Las enfermedades infecciosas, a su vez, son particularmente interesantes debido a que los blancos genéticos buscados (ácidos nucleicos de bacterias, virus y otros organismos patógenos) se encuentran en un contexto de genomas de menor complejidad, lo que repercute en mayor facilidad para su análisis e identificación. Dentro de este grupo destacan las infecciones de las vías respiratorias inferiores, por contar con un valor de DALYs por cien mil habitantes algo más elevado que otras enfermedades infecciosas y también debido a que en la práctica clínica actual se realiza ya la identificación de los agentes causantes e incluso de resistencia a diversos fármacos a través de la detección de blancos específicos para diferentes patógenos (que podrían mejorarse drásticamente en términos de velocidad, especificidad y versatilidad mediante la utilización de tecnología genómica en dichas pruebas).

En este trabajo seleccionamos a las infecciones de las vías respiratorias inferiores como enfermedad objetivo para el desarrollo de nuevas pruebas genómicas, ya que presenta valores de “Segura” para la factibilidad en el corto plazo y “Drástico” para el efecto esperado debidos principalmente a la facilidad de análisis de los blancos genéticos y a la claridad de sus repercusiones en el manejo clínico de los pacientes.

8.2 Propuesta de una prueba para la enfermedad seleccionada

Una vez que hemos identificado un padecimiento que sobresale a los demás con base en los criterios antes establecidos, el siguiente paso es diseñar una prueba con tecnología de análisis genómico, traduciendo las características y naturaleza del padecimiento a un esquema de proceso adecuado.

8.2.1 Necesidades específicas del ámbito clínico

Para lograr que una prueba con tecnología de análisis genómico impacte en los sistemas nacionales de salud es necesario obtener la intersección entre lo que es posible realizar a partir de la genómica y lo que el área clínica considera realmente útil.

En base a este planteamiento nos acercamos al laboratorio de microbiología clínica del Instituto Nacional de Enfermedades Respiratorias (INER), que además de contar con amplia experiencia en el manejo de técnicas moleculares es un ejemplo de la práctica clínica actual más avanzada y cuyo ámbito de competencia es justamente los padecimientos del aparato respiratorio.

Esta interacción con una entidad tan importante del cuidado de la salud en México, ayudó a destacar los aspectos más importantes de las infecciones de las vías respiratorias inferiores, así como para puntualizar las necesidades (particularmente en cuanto al desarrollo de nuevas pruebas moleculares), en las áreas de diagnóstico simultáneo de patógenos y detección de resistencia a antibióticos, siendo los más importantes los que explicaremos a continuación:

En pacientes pediátricos o en estado crítico sería beneficioso permitir el diagnóstico diferencial de forma rápida, identificando a los patógenos responsables del padecimiento mediante la búsqueda de blancos genéticos para todos los agentes que pueden causar este tipo de enfermedades. Sin embargo, se considera de potencialmente mayor utilidad el desarrollo de pruebas para ayudar a la selección de los tratamientos más adecuados.

Esto se debe a que en la práctica clínica, generalmente, no se realizan pruebas para todos los patógenos, sino que más bien se buscan sólo aquellos que se sospecha pudieran encontrarse en el paciente. Entre los sistemas que utilizan para comprobar estas sospechas se encuentran: metodologías de secuenciación para la detección y diagnóstico de infecciones por hongos, microarreglos de ADN para aquellas causadas por virus respiratorios y pruebas de caracterización bioquímica y fenotípica para la identificación de bacterias. Además, se considera que estos sistemas cubren con sus necesidades actuales, incluso en el caso de las pruebas basadas en cultivo ya que cuentan con sólidas plataformas de automatización para llevarlas a cabo.

En cuanto a la propuesta de nuevas pruebas para la detección de resistencia a antibióticos (mediante el análisis de los genes donde se presenten mutaciones que confieren resistencia o que codifican proteínas de resistencia) se piensa que los mayores beneficios se obtendrían en el caso de patógenos que presentan con frecuencia resistencia a múltiples agentes terapéuticos y que además son considerados como de lento crecimiento, debido a que esto dificulta la realización de

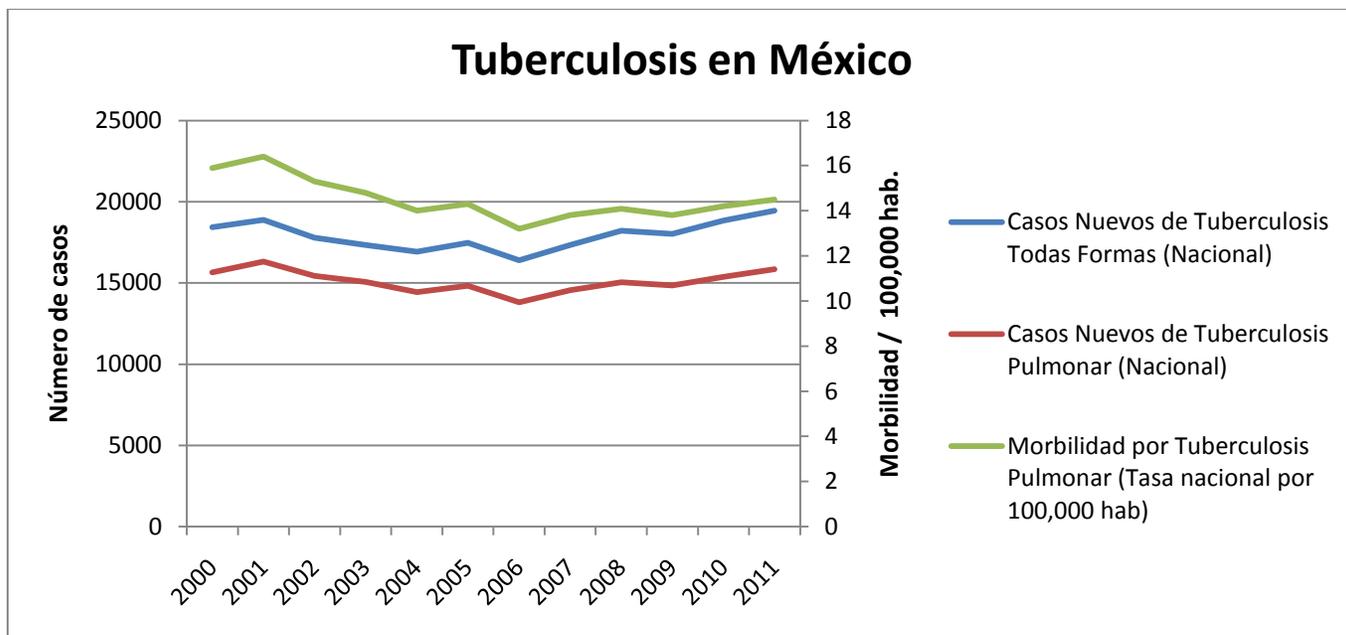
las pruebas bioquímicas dependientes de cultivo con las cuales se determina normalmente la susceptibilidad o resistencia.

Dentro de este tipo de microorganismos un ejemplo particularmente importante son aquellos pertenecientes al complejo de *Mycobacterium tuberculosis*, ya que han causado preocupación internacional debido a las bajas tasas de detección y curación en países en vías de desarrollo⁵³ y a la emergencia y aumento de resistencia a múltiples fármacos que dificulta y vuelve más costoso el control de esta enfermedad^{54,55}.

En el caso particular de México, cada año se presentan alrededor de 15,000 nuevos casos de tuberculosis respiratoria, los cuales no han disminuido de manera notable en los últimos años (Figura 10) a pesar de la instauración de políticas para el control de esta enfermedad.

En cuanto a la prevalencia de resistencia a antibióticos, en nuestro país sólo el 2.8% de los casos de tuberculosis se encuentran dentro de la definición de MDR-TB (cepas resistentes al menos a isoniacida y rifampicina, dos de los cuatro antibióticos de primera línea utilizados en su tratamiento). Sin embargo, el 18% de todos los casos es resistente al menos a un fármaco antituberculoso y este porcentaje se eleva hasta el 27% cuando el paciente cuenta con un historial de tratamiento previo⁵⁶.

Figura 10. Casos nuevos y morbilidad de la tuberculosis en México (2000-2011).



Fuente: Cifras oficiales del Centro Nacional de Vigilancia Epidemiológica y Control de Enfermedades (2000-2011).

8.2.2 Definición de un esquema de proceso particular

Como se mencionó anteriormente, dentro del grupo de padecimientos encontrado como prioritario en la primera sección de este trabajo (las enfermedades infecciosas de las vías respiratorias inferiores), resultaría particularmente útil desarrollar una prueba para la detección de resistencia en patógenos pertenecientes al complejo de *Mycobacterium tuberculosis* debido a su lento crecimiento y frecuente resistencia a múltiples antibióticos.

Para que esta prueba pueda, en un futuro, insertarse en la práctica clínica es necesario que utilice como materia prima alguno de los tipos de muestra más comúnmente utilizados en el INER, por ejemplo, esputo inducido o espontáneo y lavado broncoalveolar (prefiriendo el esputo por ser menos invasivo) para llevar a cabo la extracción de los ácidos nucleicos (ADN total) mediante sus protocolos estandarizados.

Por otro lado, un aspecto fundamental que se debe tomar en cuenta para el desarrollo de pruebas genómicas para enfermedades infecciosas es el hecho de que los blancos genéticos que son de interés (provenientes del genoma de los patógenos) conforman una proporción muy baja de los ácidos nucleicos totales en una muestra clínica, ya que se tendrá gran cantidad de ácidos nucleicos provenientes del genoma del paciente, lo que complicará la detección de los blancos de interés. Para superar este obstáculo se han propuesto diferentes alternativas metodológicas, entre las que destacan la amplificación selectiva de ciertos genes de interés⁵⁷, la captura por hibridación del material genético del patógeno⁵⁸ e incluso la secuenciación a gran cobertura de todo el material genético presente en la muestra⁵⁹.

La mayor parte de la resistencia ante los antibióticos de primera línea (isoniazida, rifampicina, etambutol y pirazinamida), así como otras alternativas de tratamiento utilizadas cotidianamente en el tratamiento de tuberculosis (estreptomina, amikacina y algunas fluoroquinolonas como ciprofloxacina y ofloxacina) se ha encontrado consistentemente asociada a mutaciones en unos cuantos genes bien identificados (Tabla 3).^{55,60,61} Por esta razón la metodología más adecuada sería la amplificación selectiva de estas regiones del genoma de *M tuberculosis* que se asocian a la presencia de resistencia ante los principales fármacos antituberculosos por reacción en cadena de la polimerasa de forma previa al análisis genómico.

Tabla 3. Principales locus asociados a la resistencia a los antibióticos de interés en *Mycobacterium tuberculosis*.

Antibiótico	Locus	Producto	Relación con el fármaco	Resistencia explicada por mutaciones en el gen
Isoniazida	<i>katG</i>	Catalasa peroxidasa	Activa pre-fármaco	90.5% * ⁶¹
	<i>mabA/inhA (promotor)</i>	Enoil- ACP reductasa	Blanco del fármaco	
Rifampicina	<i>rpoB</i>	Subunidad beta de la RNA polimerasa	Blanco del fármaco	97.10% ⁶¹
Etambutol	<i>embB</i>	Arabinosil transferasa	Blanco del fármaco	80.50% ⁶¹
Pyrazinamida	<i>pncA</i>	Pirazinamidasa	Activa pre-fármaco	84.60% ⁶¹
Amikacina	<i>rrs</i>	16S rRNA	Blanco del fármaco	90% ⁶¹
Quinolonas	<i>gyrA</i>	Subunidad A de DNA girasa	Blanco del fármaco	89.5% ⁶²
	<i>gyrB</i>	Subunidad B de DNA girasa	Blanco del fármaco	<11.6% ⁶²
Estreptomycin	<i>rpsL</i>	Proteína ribosomal S12	Blanco del fármaco	52-59% ⁶³
	<i>rrs</i>	16S rRNA	Blanco del fármaco	8-21% ⁶³

* 69.2% con mutaciones exclusivamente en *katG* y 12.2% con mutaciones sólo en *inhA*

Por otro lado, aunque la mayor parte de la resistencia se encuentra determinada por estas regiones relativamente restringidas del genoma de *M. tuberculosis* (Anexo 3), existe una amplia diversidad de mutaciones que pueden ocasionar resistencia a los fármacos y se desconoce cuáles de ellas son más importantes en México, debido a que la frecuencia de mutaciones específicas varía según la región geográfica^{55,60,61} y además a que existe la posibilidad de encontrar mutaciones no reportadas en la literatura.

Considerando lo anterior sería útil seleccionar las tecnologías de secuenciación como método de análisis genómico, para permitir la identificación de todas las mutaciones (incluso si no han sido reportadas previamente) que pudieran encontrarse en estas regiones de forma simultánea.

Debido a que, en conjunto, el tamaño de las regiones asociadas a resistencia que deben analizarse no sobrepasa las 6 Kb (utilizando los cebadores presentados en el Anexo 2 para la amplificación por PCR) lo más conveniente sería utilizar un secuenciador como el Personal Genome Machine (PGM) de Ion Torrent con un “Chip 314”, ya que sus características principales son una menor capacidad de producción de información (10 Mb de secuencia), un relativamente bajo costo por experimento y un tiempo de procesamiento reducido al compararse con otras plataformas de secuenciación masiva²⁵.

Además, dicha capacidad de producción de información permitiría reunir hasta 32 muestras en la misma corrida de secuenciación mediante la utilización de un kit comercial de “multiplex vía

barcodes”, reduciendo el costo por muestra analizada pero, al mismo tiempo, manteniendo una profundidad de cobertura promedio suficiente para la determinación de variantes de todo tipo en las regiones de interés (profundidad de cobertura mayor a 50x).

Una vez se realiza la secuenciación en el Ion Torrent PGM se lleva a cabo el análisis informático mediante el uso de la paquetería incluida en esta plataforma, donde las lecturas obtenidas son agrupadas según correspondan a cada una de las muestras a través de los “barcodes” y son procesadas para eliminar las secuencias de los adaptadores. Después, las lecturas son utilizadas para ensamblar las secuencias correspondientes a las regiones estudiadas y alineadas con el genoma de *Mycobacterium tuberculosis H37Rv* (NCBI Reference Sequence: NC_000962.2) como referencia para determinar las mutaciones presentes en cada una de las muestras.

Como paso final se realiza una predicción computacional o se utiliza la información reportada previamente para analizar las mutaciones encontradas en cada una de las muestras y determinar si éstas confieren resistencia o no a los antibióticos estudiados.

En el Anexo 4 se presenta un protocolo detallado del esquema de proceso definido para llevar a cabo esta prueba y a continuación se muestran sus características principales (Figura 11).

Figura 11. Características de la prueba propuesta.



8.3 Validación analítica de la prueba propuesta

Para evaluar el desempeño de la prueba se utilizó el esquema de secuenciación masiva descrito en el Anexo 4 y la sección de metodología para analizar la resistencia a los principales antibióticos en muestras de *Mycobacterium tuberculosis* previamente caracterizadas fenotípicamente en el Instituto Nacional de Enfermedades Respiratorias (INER).

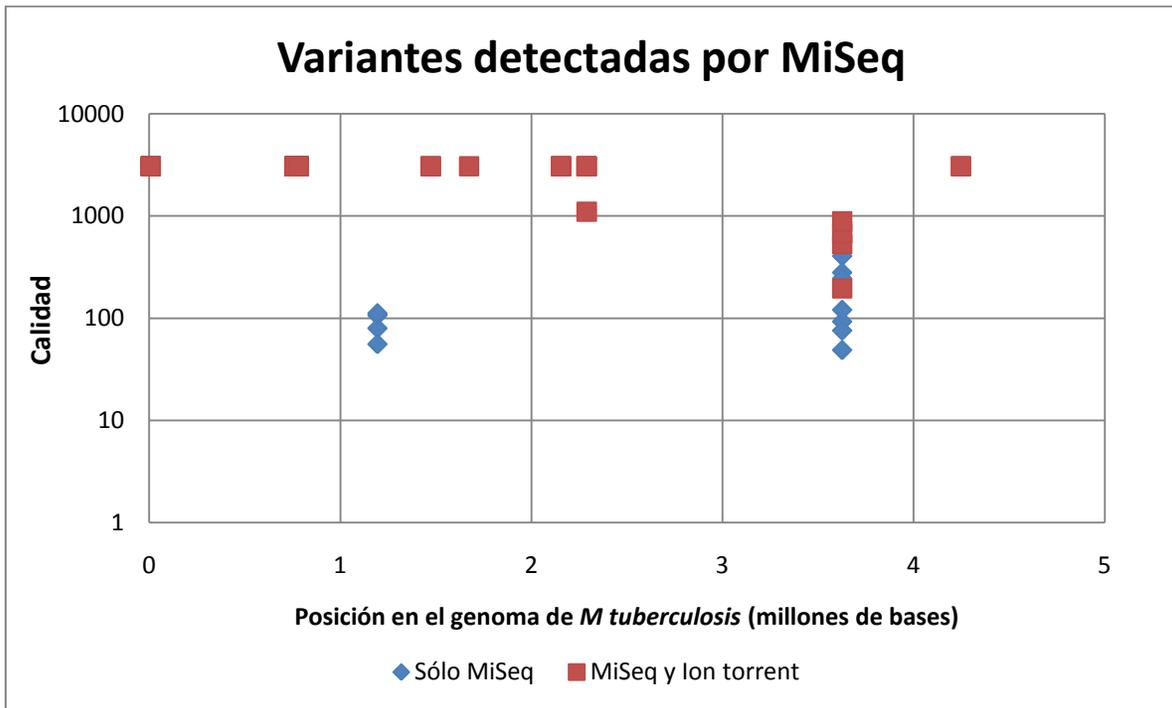
8.3.1 Validación de la detección de mutaciones

Para comprobar que las mutaciones detectadas a través de la secuenciación con la plataforma seleccionada (Ion Torrent PGM) representan verdaderas mutaciones presentes en las muestras se realizó una comparación con los resultados de secuenciar cinco de las muestras con una plataforma de secuenciación que utiliza tecnología diferente (Illumina MiSeq), considerando que aquellas variantes que son detectadas por ambas plataformas son mutaciones verdaderamente presentes en las muestras.

En las 5 muestras de *Mycobacterium tuberculosis* analizadas con ambos secuenciadores se encontraron 76 variantes respecto a la cepa de referencia, que incluyen tanto cambios puntuales de nucleótidos como inserciones y deleciones. De éstas, 36 fueron detectadas en ambas plataformas de secuenciación, mientras que 29 fueron detectadas únicamente en el Ion torrent PGM y 11 sólo con el MiSeq de Illumina.

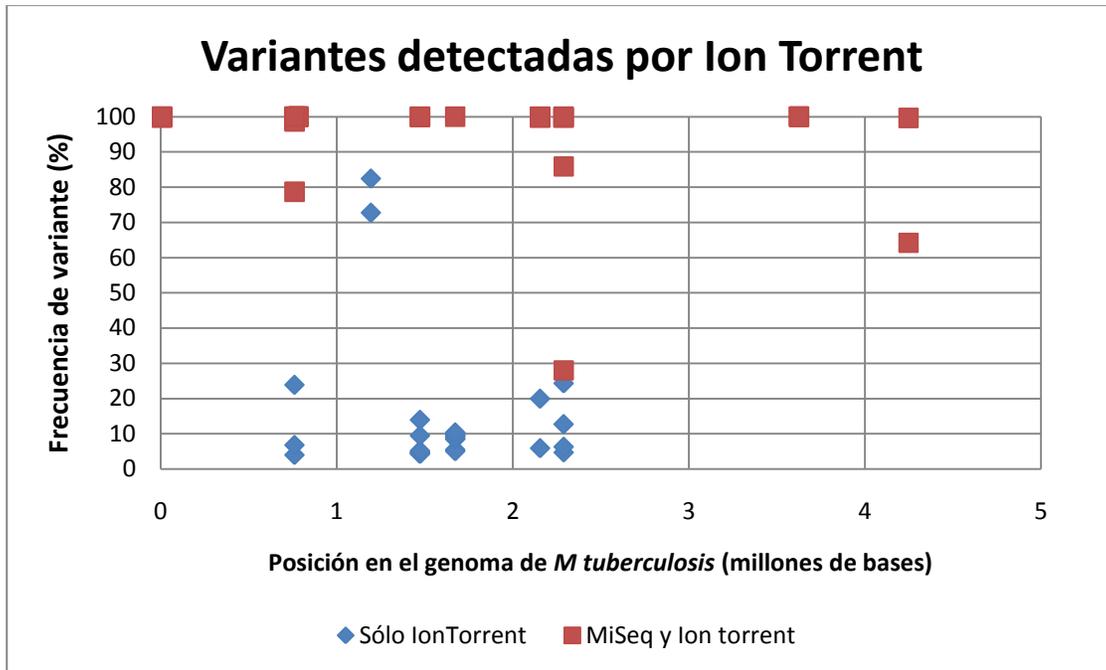
En el caso del secuenciador MiSeq, las variantes que sólo fueron detectadas en esta plataforma, tienden a tener una menor calidad que aquellas que fueron encontradas además con Ion Torrent, lo que indica que probablemente se traten de errores específicos de la química utilizada por Illumina para la secuenciación (figura 12).

Figura 12. Calidad de las variantes detectadas por el secuenciador MiSeq (Illumina).



Por su parte, las variantes detectadas por ambas plataformas de secuenciación tienden a encontrarse en una alta proporción de las lecturas obtenidas por el Ion Torrent PGM para una misma posición (presentando una alta frecuencia de variante), sin embargo, como se muestra en la figura 13, éste no fue el caso para todas ellas.

Figura 13. Frecuencia de variante de las variantes detectadas por el secuenciador PGM (Ion Torrent).

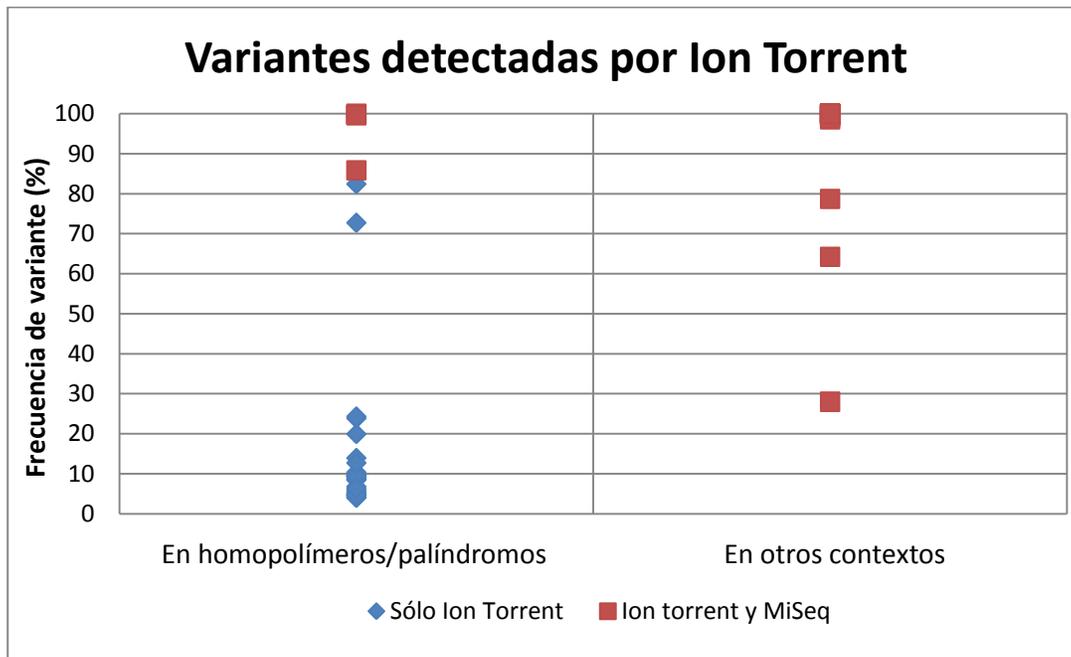


Todas las variantes que fueron detectadas únicamente por el Ion Torrent PGM presentaron una frecuencia de variante menor al 85% y se encontraron en determinados contextos genómicos (figura 14), específicamente homopolímeros (secuencias desde cuatro nucleótidos iguales adyacentes a la posición donde se encontró la variante) y palíndromos (dos secuencias de cuatro o más nucleótidos que son complementarias entre sí y que flanquean menos de siete nucleótidos que incluyen la posición donde se detectó la variante).

Mientras tanto, las variantes que fueron detectadas tanto con MiSeq como con Ion Torrent se encontraron con una frecuencia de variante mayor al 85% en regiones homopoliméricas y palindrómicas, mientras que en otros contextos genómicos presentaron frecuencias de variante apenas por encima del 20%.

Este es un hecho particularmente notable, ya que indica que en cierto número de las lecturas que abarcan este tipo de regiones, la tecnología de secuenciación utilizada por Ion Torrent interpretará una base incorrecta (según nuestras observaciones la correspondiente al nucleótido homopolimérico o una base complementaria que aumente la longitud del palíndromo).

Figura 14. Frecuencia de variante, según contexto genómico de las variantes detectadas por el PGM (Ion Torrent).



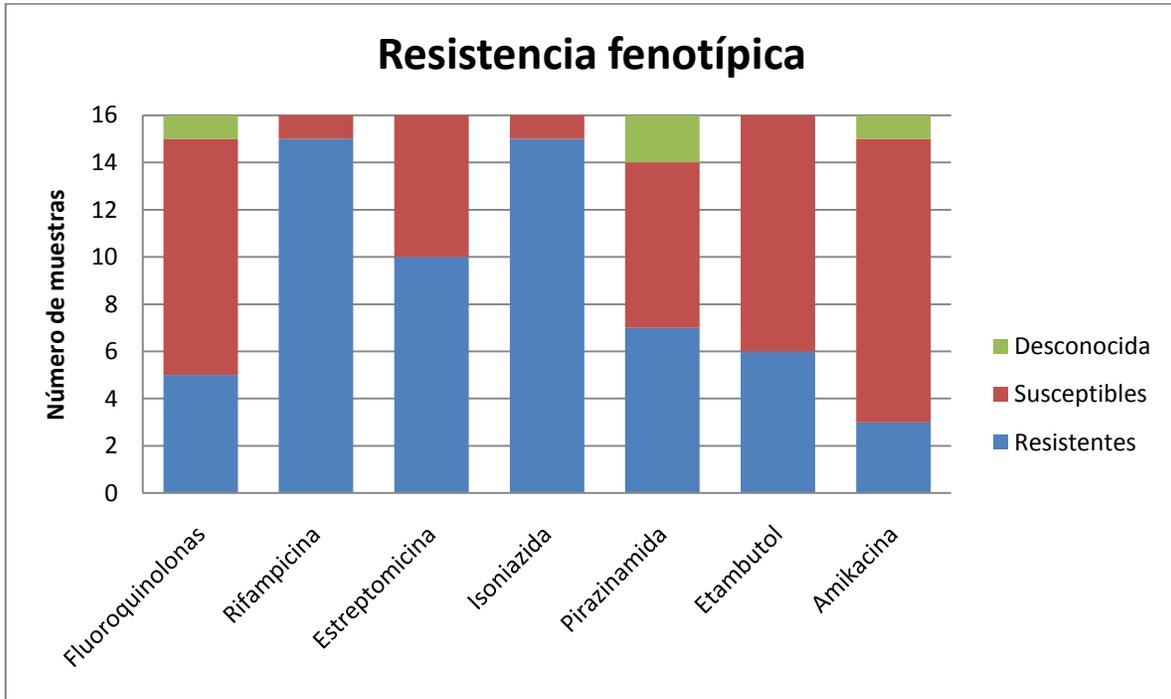
A partir de esta información es posible realizar un análisis del contexto genómico y frecuencia de variante, que nos permitirá estimar de mejor forma cuáles de entre todas las variantes detectadas representan verdaderas mutaciones presentes en las muestras y reducir los errores específicos de la secuenciación mediante el Ion Torrent PGM.

8.3.2 Validación de la detección de resistencia a antibióticos mediante secuenciación.

Respecto a la fenotipificación realizada previamente en el INER de las cepas de *Mycobacterium tuberculosis* (n=16, figura 15), se consideró que cinco (31.3%) son fenotípicamente resistentes a fluoroquinolonas al presentar resistencia a ofloxacina y/o moxifloxacina, diez (62.5%) se consideraron susceptible a toda la clase de antibióticos al ser fenotípicamente susceptible a ambos fármacos (siete cepas), o a uno de ellos cuando no se contaba con resultados de las pruebas fenotípicas para ambos antibióticos (tres muestras), y únicamente se consideró desconocido el fenotipo respecto a la resistencia ante fluoroquinolonas para una muestra (6.3%) al carecer de pruebas de susceptibilidad para los dos fármacos de esta familia ya mencionados. En el caso de rifampicina, se determinó que 15 (93.8%) eran resistentes y una (6.3%) susceptible a dicho antibiótico. Además, diez (62.5%) son fenotípicamente resistentes a estreptomycin, mientras que seis (37.5%) son susceptibles al mismo fármaco. Además, 15 (93.8%) presentaron resistencia fenotípica a isoniazida, mientras que una (6.3%) fue susceptible a ella. Siete (43.8%) son resistentes a pirazinamida, otras siete (43.8%) susceptibles a este antibiótico y no se cuenta con resultados de pruebas fenotípicas para dos (12.5%) de las muestras. Para etambutol, se determinó que seis (37.5%) eran resistentes y diez (62.5%) eran susceptibles a este fármaco. Por último, tres

(18.8%) son resistentes a amikacina, 12 (75%) susceptibles a ella, y se carece de resultados fenotípicos para una (6.3%).

Figura 15. Resultados de la determinación de resistencia mediante secuenciación.



Aún con cierta heterogeneidad entre la profundidad de cobertura para cada una de los amplicones secuenciados (figura 16 y anexo 5), se logró obtener una cobertura completa y con profundidad suficiente para la detección de variantes en las nueve regiones asociadas a resistencia en todas las muestras, con la excepción de sólo dos: la cepa 9071/10 donde no se obtuvo completa la región correspondiente a *pncA* y la muestra 6717/09 donde no se obtuvo la cobertura completa del amplicón de *katG*, esto se debe a que no fue posible obtener una correcta amplificación específicamente de esas regiones para dichas muestras, probablemente debido a la existencia de mutaciones que evitaron que los cebadores utilizados hibridaran adecuadamente y propiciaran la subsecuente amplificación inespecífica de otras regiones genómicas. Cabe resaltar que las proteínas codificadas en *pncA* y *katG* son necesarias para la activación de los precursores de pirazinamida e isoniazida respectivamente, y que al carecer de información sobre la presencia de mutaciones en estas regiones nuestra prueba no fue concluyente para dichos antibióticos en esas muestras específicas.

Figura 16. Profundidad de cobertura de las regiones analizadas para cada muestra.

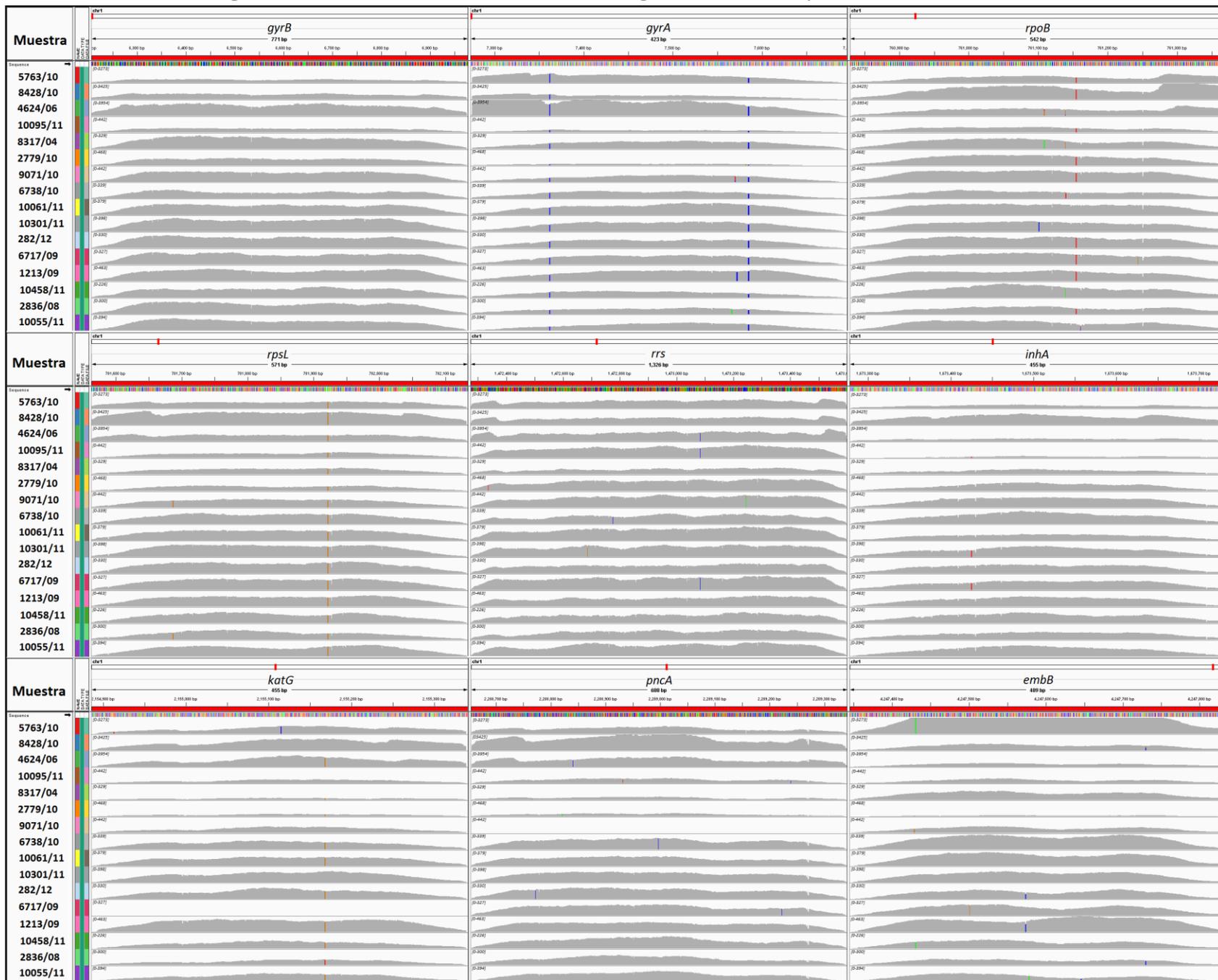
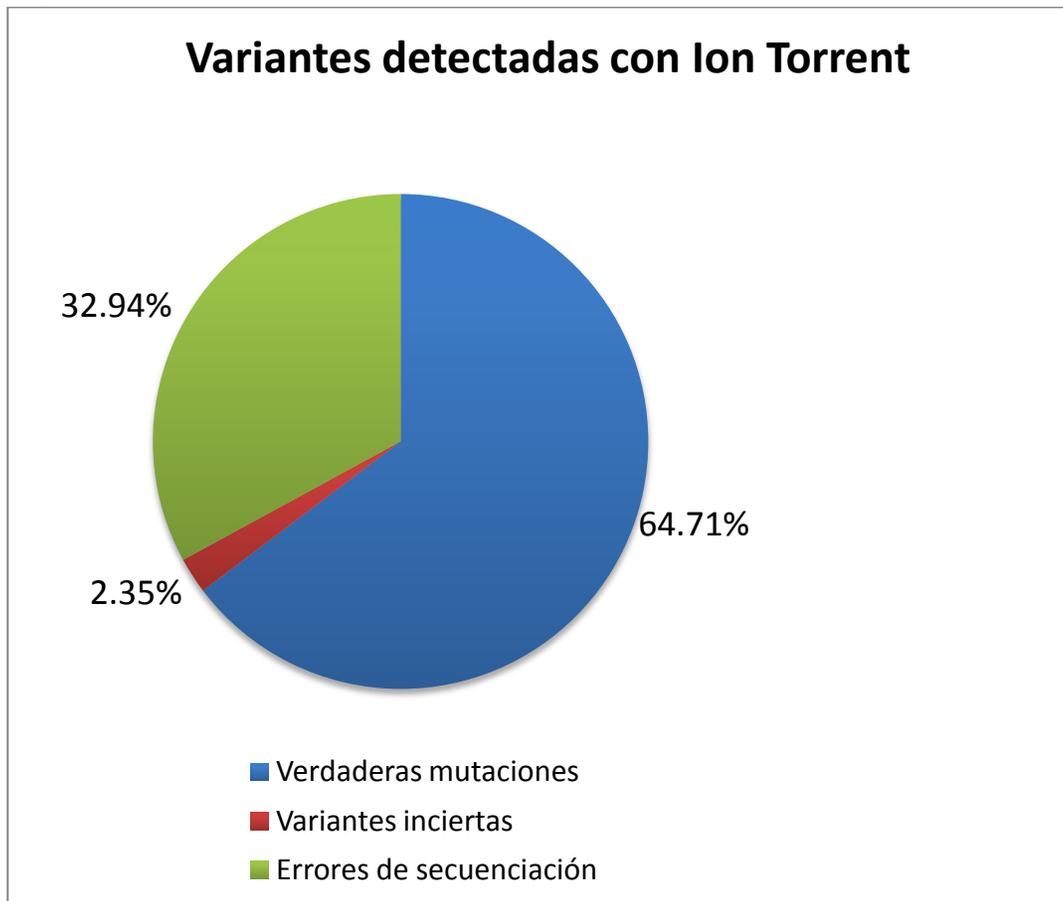


Gráfico de profundidad de cobertura para las nueve regiones del genoma de *Mycobacterium tuberculosis* analizadas. La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además, se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

En el total de nueve regiones y 16 muestras de *Mycobacterium tuberculosis* analizadas, se encontraron 170 variantes respecto a la referencia (figura 17). Con base en el análisis de validación de la detección de mutaciones descrito en la sección anterior, removimos 56 (32.94%) errores de secuenciación y determinamos que 110 (64.71%) de ellas representan mutaciones verdaderas presentes en las muestras. Sin embargo, este análisis no fue concluyente para 4 (2.35%) de ellas.

Para dicho análisis consideramos específicamente que las verdaderas mutaciones fueron detectadas en el Ion torrent PGM con una frecuencia de variante mayor al 85%, los errores de secuenciación con frecuencia de variante menor al 20% y el contexto genómico (la presencia de homopolímeros y palíndromos) fue el factor determinante para clasificar las variantes con frecuencia de variante entre estos valores.

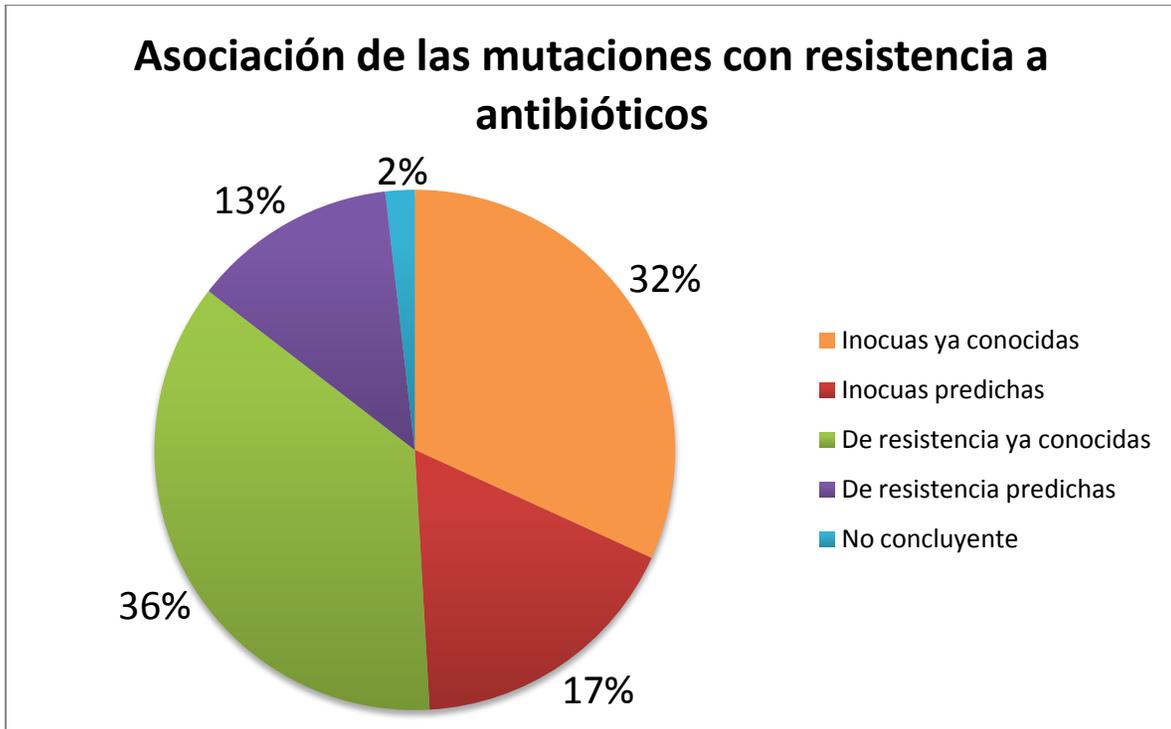
Figura 17. Clasificación de las variantes detectadas por Ion Torrent para las 16 muestras.



Después se asociaron las 110 mutaciones verdaderas a la susceptibilidad o resistencia ante los antibióticos según la información reportada y depositada en bases de datos especializadas tales como tbdb.org⁴¹ y tbdreamdb⁴², artículos recientes de secuenciación de estos genes⁶¹ y a través de la predicción del efecto funcional de los cambios de aminoácido mediante SIFT⁴³ (figura 18). En

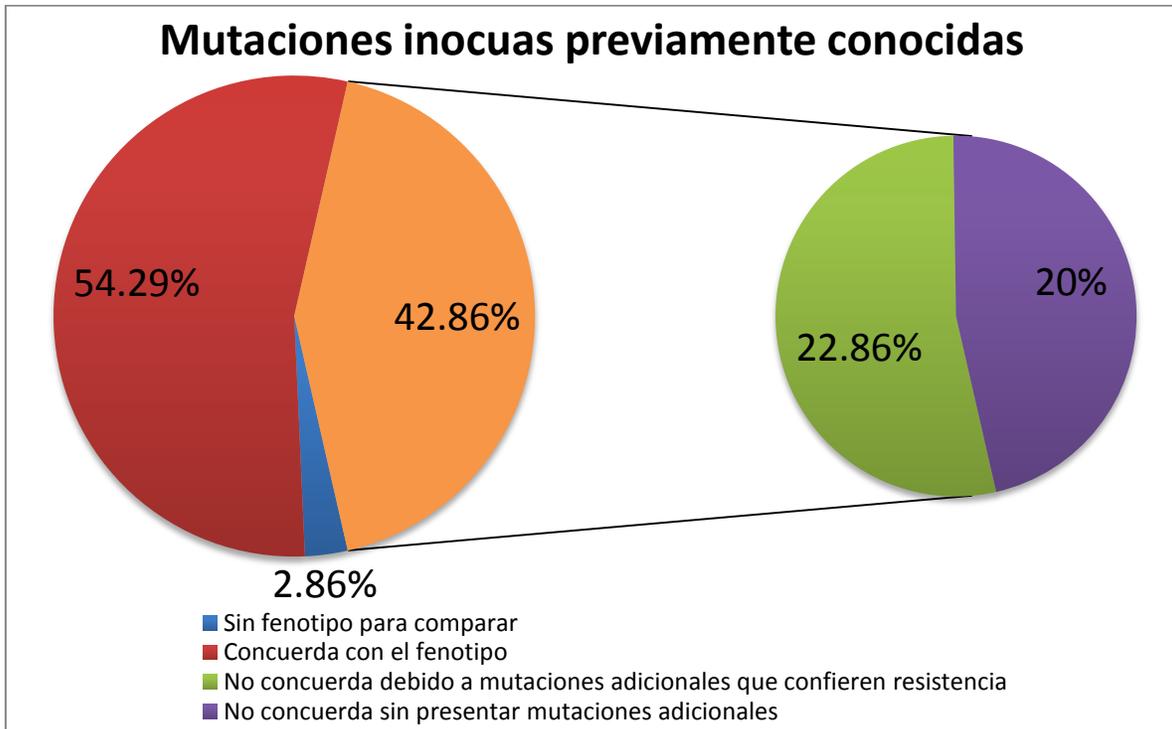
todos los casos se pudo concluir acerca de la asociación a resistencia con la excepción de dos mutaciones en el gen *rrs*, donde no fue posible predecir si se encuentran relacionadas o no con resistencia a estreptomycin y/o amikacina. La primera es una transición de adenina a guanina en el nucleótido 842 y la segunda una transversión de adenina a citosina en el nucleótido 931, estas mutaciones no fueron encontradas en reportes previos y no fue posible realizar el análisis de conservación de aminoácidos en secuencias homólogas a través de SIFT ya que este gen no codifica una proteína sino el 16S rRNA.

Figura 18. Asociación de las mutaciones con susceptibilidad o resistencia a los antibióticos.



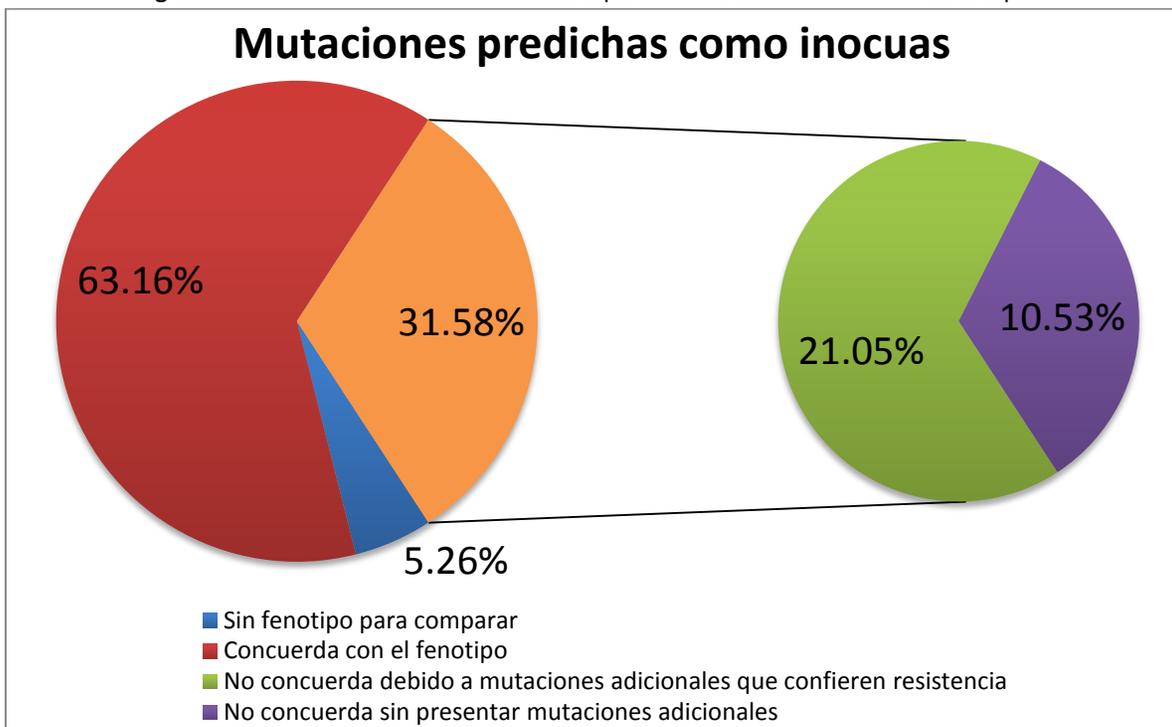
En el análisis de las 110 verdaderas mutaciones se encontraron 35 variantes que representan mutaciones conocidas y que no han sido asociadas al desarrollo de resistencia ante los antibióticos analizados (figura 19). De éstas, 19 (54.29%) concuerdan con la susceptibilidad fenotípica determinada en el INER, ocho (22.86%) no concuerdan con el fenotipo debido a la presencia (en la misma muestra) de otras mutaciones que confieren resistencia al mismo antibiótico, y no se contaba con el fenotipo de una (2.86%) donde no se pudo realizar la comparación. Sin embargo, para siete de estas mutaciones inocuas (20%) no se encontraron otras mutaciones que pudieran explicar la resistencia reportada, por lo que es probable que haya alguna mutación fuera de las regiones secuenciadas que determine la resistencia a esos antibióticos. Cabe notar además, que en seis de estos siete casos se trataba de resistencia a estreptomycin, y que en las regiones asociadas a este fármaco sólo se encontraron mutaciones sinónimas que difícilmente podrían conferir resistencia.

Figura 19. Concordancia de las mutaciones previamente reportadas como inocuas con el fenotipo.



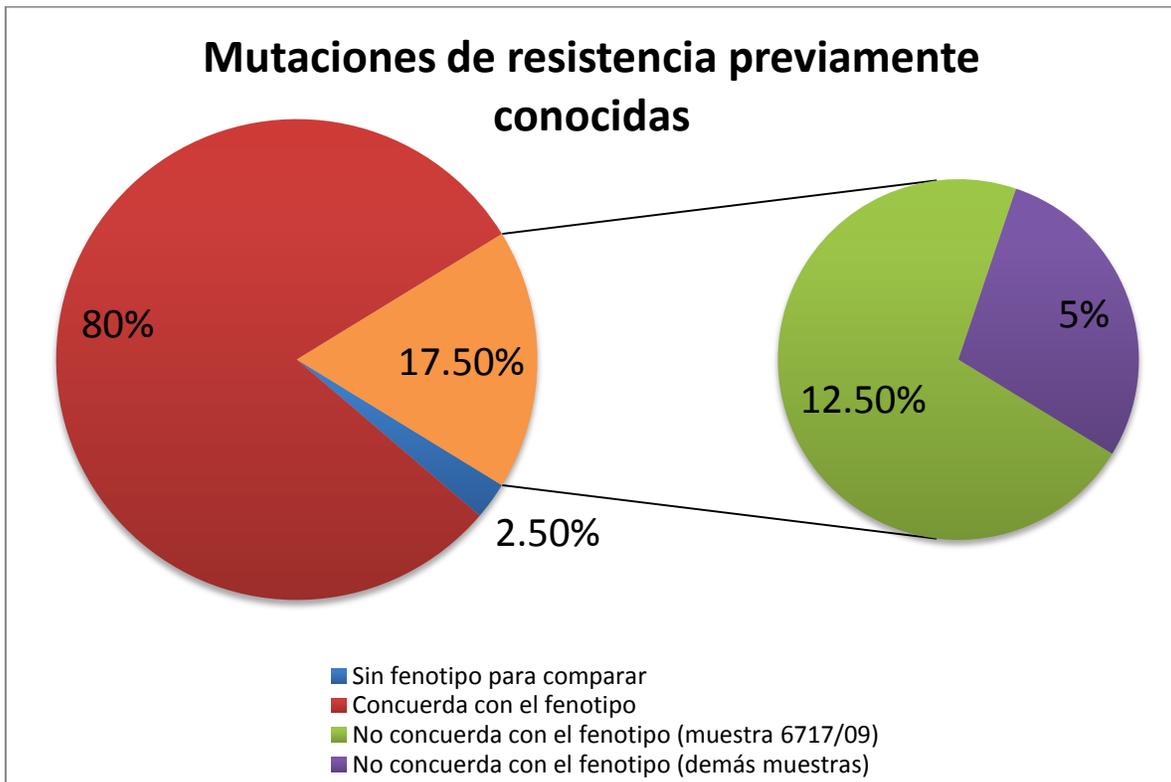
También se predijo que otras 19 mutaciones serían inocuas y no conferirían resistencia (figura 20), de las cuales 12 (63.16%) concuerdan con un fenotipo susceptible, 4 (21.05%) no concuerda ya que en la misma muestra se presentaron otras mutaciones que confieren resistencia al mismo fármaco y sólo en otras 2 (10.53%) la no concordancia con los datos fenotípicos podría deberse a errores en la predicción o a que pudieran existir mutaciones fuera de las regiones analizadas que intervengan en el desarrollo de resistencia al fármaco.

Figura 20. Concordancia de las mutaciones predichas como inocuas con el fenotipo.



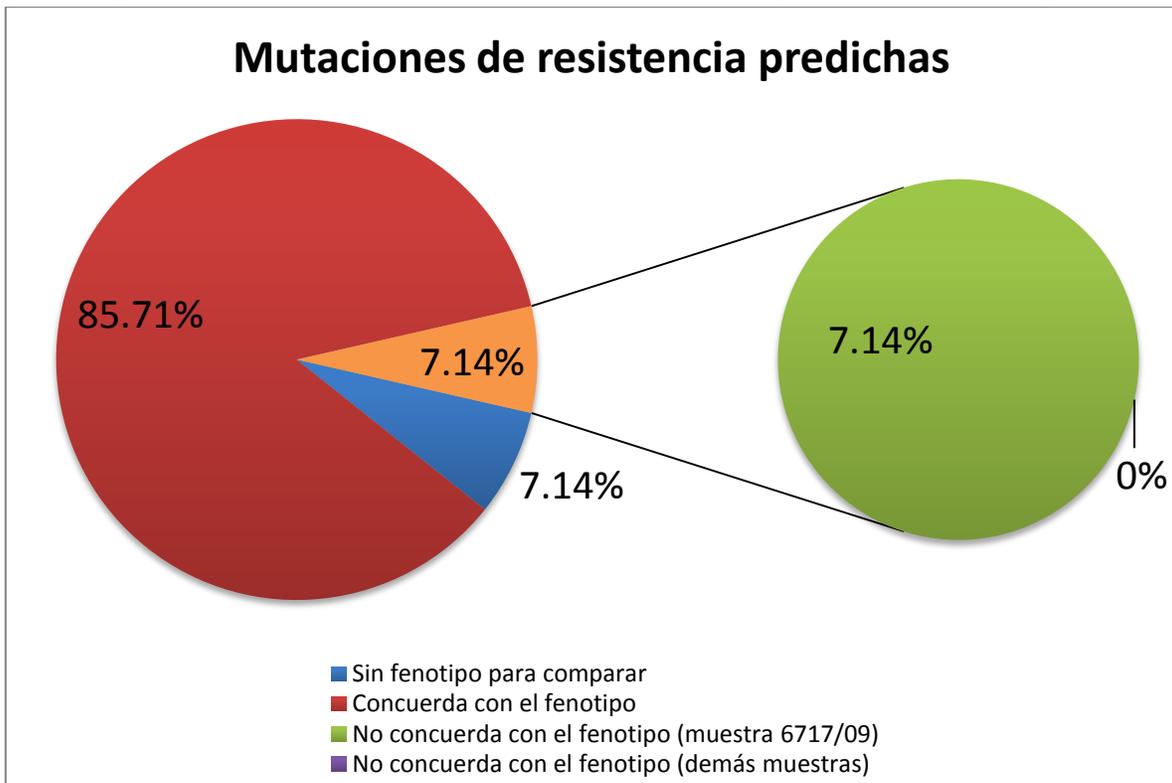
Por otro lado, se detectaron 40 mutaciones previamente reportadas como asociadas a la resistencia a los antibióticos estudiados (figura 21), de las cuales 32 (80%) concuerdan con el fenotipo, el cual es desconocido para una (2.5%) y no concordante para 7 (17.5%). Notablemente, cinco de los casos donde no hubo concordancia fueron detectados en la misma muestra (6717/09), lo que podría indicar algún problema con la caracterización fenotípica de dicha muestra, mientras que sólo uno de ellos concierne a una mutación que ha sido previamente asociada a resistencia pero que recientemente se ha determinado que entre 10 y 20% de las cepas de *Mycobacterium tuberculosis* con esta mutación pueden ser todavía susceptibles al fármaco⁶⁴.

Figura 21. Concordancia de las mutaciones de resistencia previamente reportadas con el fenotipo.



Además se predijo que 14 mutaciones conferirían resistencia (figura 22), de las cuáles 12 (85.71%) concuerdan con el fenotipo reportado por el INER, se tiene discordancia con los datos fenotípicos para una (7.14%) y se carece de este tipo de datos y no se pudo realizar la comparación en solamente una (7.14%) de ellas. Cabe resaltar también que la mutación que se predijo otorgaría resistencia y que no corresponde con el fenotipo reportado se encuentra en la muestra 6717/09 que, como se mencionó anteriormente presenta también discordancia entre la presencia de varias mutaciones ya conocidas y los datos fenotípicos.

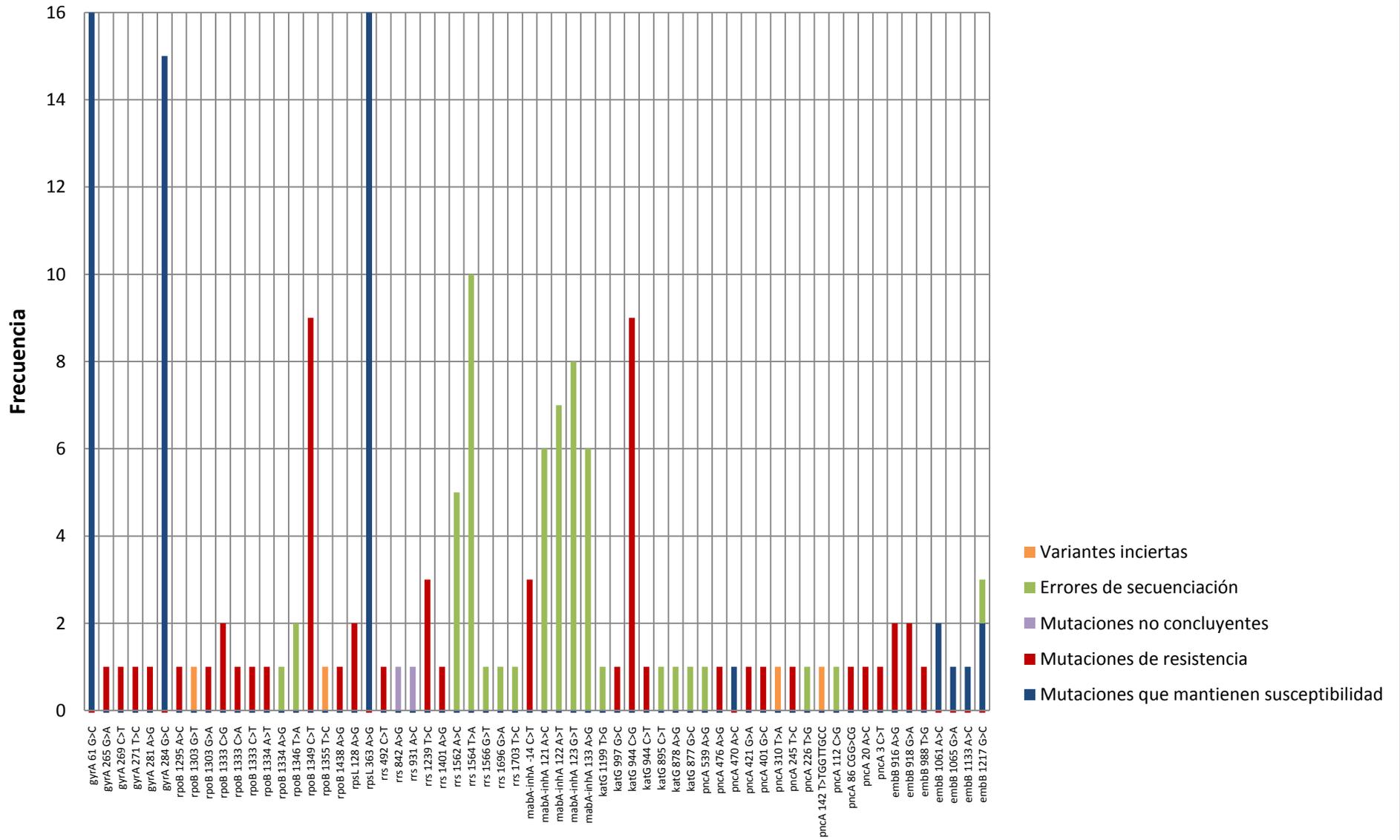
Figura 22. Concordancia de las mutaciones de resistencia predichas con el fenotipo.



Cabe resaltar que las predicciones que realizamos acerca de la asociación a susceptibilidad o resistencia, correspondieron a los datos fenotípicos en una proporción muy similar a aquella de las mutaciones previamente reportadas, lo que resalta la prometedora utilidad de llevar a cabo este tipo de inferencias cuando no se cuenta con información previa del efecto funcional de las mutaciones.

Figura 23. Resumen de variantes detectadas en las 16 muestras de *Mycobacterium tuberculosis*.

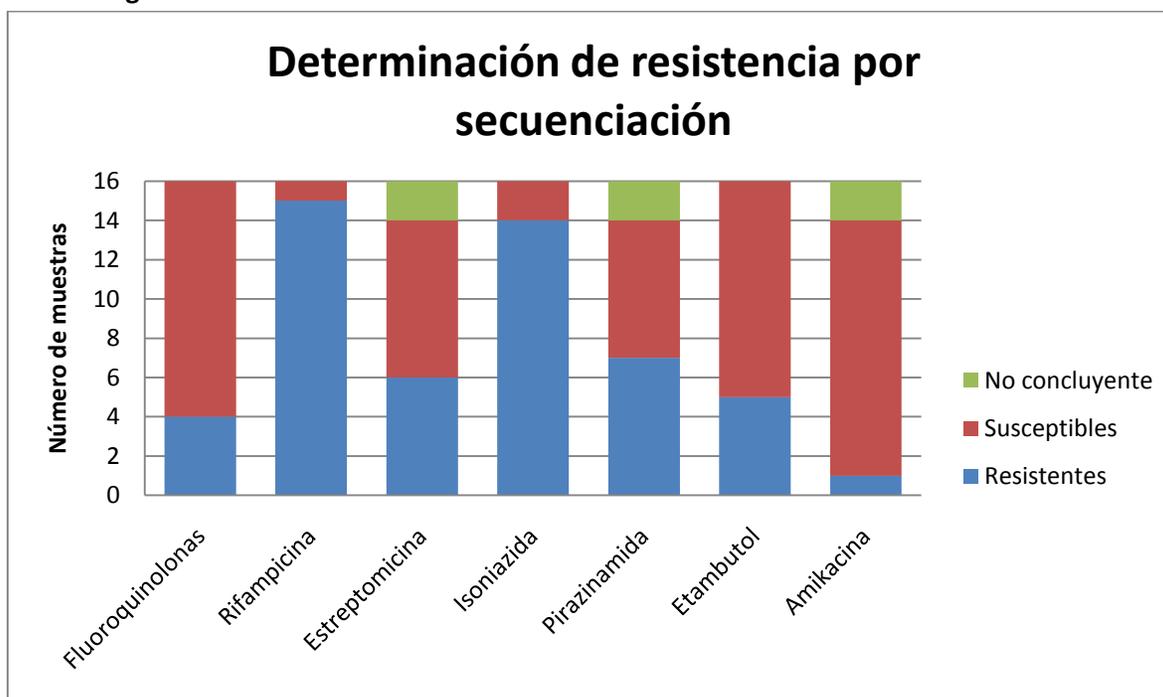
Variantes detectadas con Ion Torrent



8.3.3 Concordancia entre la resistencia inferida por secuenciación y los datos fenotípicos.

De las muestras de *Mycobacterium tuberculosis* analizadas mediante nuestro protocolo (n=16), cuatro (25%) se predijeron resistentes a fluoroquinolonas, mientras que 12 (75%) no presentaron mutaciones asociadas a la resistencia frente a esta clase de antibióticos. En 15 (93.8%) se detectaron mutaciones asociadas a la resistencia ante rifampicina, mientras que una (6.3%) se predijo susceptible a dicho fármaco. En base a las mutaciones presentes, se infirió también que seis (37.5%) serían resistente a estreptomina y que ocho (50%) serían susceptibles, sin embargo para dos (12.5%) no fue posible concluir acerca de susceptibilidad o resistencia ante este antibiótico. Además, 14 (87.5%) presentaron mutaciones asociadas a resistencia ante isoniazida, mientras que dos (12.5%) no lo hicieron. En cuanto a la resistencia a pirazinamida, para dos (12.5%) de las muestras no se obtuvo un resultado concluyente, siete (43.8%) fueron clasificadas como resistentes y las siete (43.8%) restantes como susceptible a este fármaco. Cinco (31.3%) presentaron mutaciones asociadas a la resistencia ante etambutol y once (68.8%) carecieron de ellas. Mientras tanto, sólo una (6.3%) presentó mutaciones asociadas a la resistencia ante amikacina, 13 (81.3%) no presentaron dichas mutaciones y para dos (12.5%) el resultado no fue concluyente.

Figura 24. Resultados de la determinación de resistencia mediante secuenciación.



Analizaremos ahora la correlación, a nivel de muestras, entre la resistencia a fármacos inferida a través de la secuenciación de nueva generación con IonTorrent y la fenotipificación realizada previamente en el Instituto Nacional de Enfermedades Respiratorias (Tabla 4): De las 15 cepas que contaban con resultados concluyentes tanto de secuenciación como de fenotipificación para resistencia a fluoroquinolonas, 14 fueron clasificadas correctamente (4 como resistentes y 10 como susceptibles). En el caso de rifampicina, 16 cepas contaron con resultados concluyentes y 14 de ellas fueron clasificadas correctamente, todas como resistentes al fármaco. Para las 14 cepas que contaban con resultados concluyentes respecto a estreptomicina, sólo 7 fueron clasificadas correctamente, 4 como resistentes y 3 como susceptibles. De las 16 cepas que contaban con resultados concluyentes para resistencia a isoniazida, 13 fueron correctamente clasificadas (todas como resistentes al fármaco). Además, se clasificaron correctamente 10 de las 12 cepas que contaban con resultados concluyentes para pirazinamida, 4 como resistentes y 6 como susceptibles a este antibiótico. De las 16 cepas con resultados concluyentes para resistencia a etambutol, 11 fueron correctamente clasificadas (3 como resistentes y 8 como susceptibles). Mientras que en el caso de amikacina, se clasificaron correctamente 12 de las 13 cepas con resultados concluyentes, sólo 1 como resistente y otras 11 cepas como susceptibles al fármaco.

Tabla 4. Resistencia a fármacos inferida mediante secuenciación y evaluada fenotípicamente.

Muestra	Datos	Fluoroquinolonas	Rifampicina	Estreptomicina	Isoniazida	Pirazinamida	Etambutol	Amikacina	Concordancia
8428/10	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (S)	PZA (S)	EMB (S)	AMK (S)	6/7
	Fenotipo	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	
5763/10	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (R)	EMB (R)	AMK (S)	5/7
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (S)	
4624/06	Secuenciación	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (S)	4/6
	Fenotipo	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (Sin información)	EMB (R)	AMK (S)	
10095/11	Secuenciación	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (S)	6/7
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (R)	AMK (S)	
8317/04	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (R)	EMB (S)	AMK (S)	5/6
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (Sin información)	EMB (S)	AMK (S)	
2779/10	Secuenciación	FLQ (R)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (R)	AMK (S)	7/7
	Fenotipo	FLQ (R)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (R)	AMK (S)	
9071/10	Secuenciación	FLQ (R)	RIF (R)	SM (R)	INH (S)	PZA (No concluyente)	EMB (R)	AMK (R)	5/6
	Fenotipo	FLQ (R)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (R)	AMK (R)	
6738/10	Secuenciación	FLQ (S)	RIF (R)	SM (No concluyente)	INH (R)	PZA (R)	EMB (S)	AMK (No concluyente)	5/5
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (R)	
10061/11	Secuenciación	FLQ (S)	RIF (S)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	4/7
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (S)	EMB (S)	AMK (R)	
10301/11	Secuenciación	FLQ (S)	RIF (R)	SM (No concluyente)	INH (R)	PZA (S)	EMB (S)	AMK (No concluyente)	3/5
	Fenotipo	FLQ (R)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (R)	AMK (S)	
282/12	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	5/7
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (S)	
6717/09	Secuenciación	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (R)	AMK (S)	0/5
	Fenotipo	FLQ (Sin información)	RIF (S)	SM (S)	INH (S)	PZA (S)	EMB (S)	AMK (Sin información)	
1213/09	Secuenciación	FLQ (R)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	7/7
	Fenotipo	FLQ (R)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	
10458/11	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (R)	AMK (S)	6/7
	Fenotipo	FLQ (S)	RIF (R)	SM (R)	INH (R)	PZA (S)	EMB (R)	AMK (S)	
2836/08	Secuenciación	FLQ (R)	RIF (R)	SM (R)	INH (R)	PZA (No concluyente)	EMB (S)	AMK (S)	6/6
	Fenotipo	FLQ (R)	RIF (R)	SM (R)	INH (R)	PZA (R)	EMB (S)	AMK (S)	
10055/11	Secuenciación	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	7/7
	Fenotipo	FLQ (S)	RIF (R)	SM (S)	INH (R)	PZA (S)	EMB (S)	AMK (S)	

Se resalta en verde la susceptibilidad (S), en rojo la resistencia (R) y en amarillo la carencia de información fenotípica (Sin información) o la inconclusividad de la secuenciación (No concluyente) para los antibióticos analizados: FLQ: fluoroquinolonas, RIF: rifampicina, SM: estreptomina, INH: isoniazida, PZA: pirazinamida, EMB: etambutol y AMK: amikacina.

Con la información mencionada en los párrafos anteriores se realizó un cálculo de la sensibilidad y especificidad que alcanzó la prueba para la detección de resistencia ante los principales antibióticos utilizados en el tratamiento de la tuberculosis, obteniendo los valores que se presentan a continuación en la tabla 5.

Tabla 5. Sensibilidad y especificidad obtenidas en la validación analítica de la prueba propuesta.

	Fluoroquinolonas	Rifampicina	Estreptomicina	Isoniazida	Pirazinamida	Etambutol	Amikacina
Sensibilidad*	0.80	0.93	0.44	0.87	0.80	0.50	0.50
Especificidad[†]	1.00	0.00	0.60	0.00	0.86	0.80	1.00

*La sensibilidad representa la proporción en la cual la prueba arroja un resultado de “resistente” cuando se aplica a muestras fenotípicamente resistentes.

[†]La especificidad representa la proporción en la cual la prueba arroja un resultado de “susceptible” cuando se aplica a muestras fenotípicamente susceptibles.

9. Discusión y conclusiones

9.1 Priorización de enfermedades

Utilizando en conjunto los parámetros descritos en la primera sección de este trabajo es posible comparar diferentes grupos de enfermedades con el objeto de identificar aquellos para los cuáles sería más pertinente desarrollar pruebas con tecnologías de análisis genómico para fines de diagnóstico o manejo clínico. Además, estos parámetros también pueden ser usados para señalar aquellas áreas donde sería conveniente que el país realizara investigación a largo plazo, que actualmente cuentan con una baja factibilidad en el corto plazo pero con un gran número de años de vida saludables perdidos (como por ejemplo las heridas accidentales).

El uso de los DALYs como medida de las necesidades de salud pública de nuestro país trae como consecuencia una clasificación de enfermedades particular. Una limitante es la inevitable necesidad de agrupar los padecimientos para su análisis, ya que los grupos resultantes contienen enfermedades diversas cuyas diferencias en ocasiones impedirían la implementación de una prueba para todas ellas simultáneamente. Sin embargo, aprovechando la capacidad masiva de producción de información de las nuevas tecnologías de análisis genómico será posible desarrollar pruebas cada vez más generales, utilizando como base las características en común que comparten las enfermedades dentro de cada grupo.

Por otro lado, los DALYs permiten tomar en consideración las consecuencias de las enfermedades que impactan no sólo en la mortalidad, proporcionando una visión más real del impacto en la calidad de vida de la población. Además, gracias al interés de la OMS el uso de los DALYs seguirá extendiéndose y refinándose, lo que podría permitir que esta evaluación se realice para un mayor número de patologías, en grupos de enfermedad menos extensos, e incluso utilizando información a nivel subnacional⁶⁵.

Como se mencionó anteriormente, el valor de la factibilidad en el corto plazo es especialmente susceptible al paso del tiempo, debido al continuo descubrimiento de nuevas variantes genéticas que afectan el manejo clínico de los pacientes (posibilitado en gran medida por el dramático abaratamiento en las técnicas de análisis genómico).

Mientras tanto, el principal obstáculo para la evaluación en cuanto al efecto esperado es la falta de consenso acerca de la magnitud del impacto de las variantes genéticas en los fenotipos patológicos y la gran cantidad de asociaciones que no han podido reproducirse en estudios independientes.

Los resultados de la interacción con el área clínica dependen de las características específicas de la entidad con la cual se haya entablado comunicación. Por ejemplo, el uso de microarreglos y secuenciación que encontramos dentro de la práctica clínica no es exactamente generalizable a

todas las instancias de los servicios de salud del país y se debe más bien, a la amplia experiencia del INER en técnicas moleculares y a su estatus de hospital de alta especialidad.

Sin embargo, el interactuar con este tipo de instituciones resulta ventajoso ya que debido al gran volumen de pacientes (provenientes de toda la República Mexicana), que atienden, se podría aprovechar completamente el poder masivo de las nuevas tecnologías de análisis genómico para analizar gran cantidad de muestras simultáneamente, y por esta misma razón, la inserción de una nueva prueba en su práctica clínica cotidiana ayudaría a distribuir sus beneficios a un mayor número de individuos que si se hubiera interactuado con una entidad del primer nivel de atención.

En el diseño de una nueva prueba que impacte la práctica clínica en nuestro país un punto fundamental es la adecuada elección del método que se utilizará para analizar la presencia de los blancos genéticos buscados ya que, como se señaló previamente, el aspecto técnico del análisis genómico entraña retos particulares en razón a la diversidad de plataformas disponibles y a su rápida evolución. Por esta razón nos parece de gran importancia diseñar plataformas flexibles y evolucionables que se adapten a los nuevos descubrimientos científicos y desarrollos metodológicos para acceder a la mayor y mejor información posible en un momento determinado.

La principal ventaja de la secuenciación sobre otros métodos de análisis genómico es que posibilita la detección de todos los blancos genéticos con relevancia clínica, sin importar de qué tipo sean y aún sin tener conocimiento previo de su existencia. Además, el desarrollo de una nueva generación de tecnologías que anticipa todavía un mayor abaratamiento y versatilidad en el ámbito de la secuenciación⁶⁶, lo que augura una mayor adopción de estas tecnologías en el ámbito clínico.

La estrategia utilizada en este trabajo, a pesar de las limitaciones que presenta, puede servir como primera aproximación para desarrollar pruebas genómicas que sean útiles para la práctica clínica y, a la vez, prioritarias para la salud pública de México. Dicha estrategia puede ser resumida en los siguientes tres pasos:

1. Comparar las enfermedades según los parámetros de priorización para seleccionar un objetivo adecuado para el desarrollo de pruebas genómicas.
2. Entablar diálogo con una entidad representativa de la práctica clínica actual en dicha área y utilizar su retroalimentación como referencia para definir las características que debe tener la prueba.
3. Seleccionar un esquema de proceso particular que cumpla satisfactoriamente con las características definidas.

Sin embargo, estos pasos podrían haberse seguido en diferente orden, por ejemplo, otra forma viable e interesante sería, en primer lugar, identificar las necesidades de la práctica clínica en un área de atención a la salud determinada a través de la interacción con ciertas entidades de ese sector, evaluar después la pertinencia del enfoque genómico en su solución y el probable impacto

en salud pública mediante los parámetros propuestos, para finalizar después con la determinación de un esquema de proceso adecuado.

En nuestra opinión ambas estrategias podrían dar resultados satisfactorios al identificar las necesidades del ámbito clínico que tengan la mayor repercusión en salud pública, sin embargo el analizar en primer lugar la importancia del padecimiento en salud pública y la pertinencia del enfoque genómico en el manejo clínico de la enfermedad, nos ayuda a identificar al mismo tiempo con cuáles entidades clínicas sería conveniente entablar comunicación.

9.2 Propuesta de una prueba genómica

Al utilizar la estrategia propuesta en este trabajo, establecimos el fundamento de una nueva prueba molecular que, mediante el uso de tecnologías de análisis genómico, ayude al manejo clínico de un padecimiento de importancia para el país.

La determinación rápida y confiable de farmacorresistencia en *Mycobacterium tuberculosis* es un área de gran interés clínico y científico de actualidad, tanto en México como en el mundo. De hecho, de manera posterior al planteamiento y ejecución de este trabajo, a nivel internacional surgió la proposición independiente de una prueba basada en secuenciación de nueva generación con el mismo objetivo⁶⁷ y que resulta por lo tanto, muy similar a nuestra propuesta.

En nuestro caso, el tiempo necesario para llevar a cabo el protocolo de esta nueva prueba hasta la entrega de resultados se estima que podría llegar a ser de cuatro días (o cinco días si se incluye el envío de las muestras de DNA a un laboratorio central para su procesamiento), aunque es necesario evaluar este tiempo de forma más rigurosa en estudios posteriores. Sin embargo, es claro que ofrece la obtención de resultados clínicamente relevantes mucho antes que las pruebas de susceptibilidad basadas en cultivo, que se mantienen como estándar en la práctica clínica a pesar de que requieren seis semanas o más antes de poder incidir en el manejo clínico del paciente.

Debido a la preocupación por la detección de la resistencia a fármacos en esta enfermedad se han propuesto recientemente pruebas moleculares con estrategias y resultados variables. Cuando comparamos la prueba propuesta en este trabajo con otras pruebas moleculares para la detección rápida de farmacorresistencia en tuberculosis, la mayor ventaja que presenta es la cantidad de regiones que analiza, ya que esto le permite ofrecer información sobre un mayor número de antibióticos importantes en el tratamiento de tuberculosis.

Por ejemplo, se han propuesto pruebas basadas en PCR-RFLP y PCR-SSCP⁶⁸ (PCR-Restriction Fragment Length Polymorphism y PCR-Single Strand Conformation Polymorphism, respectivamente) que aunque son relativamente más rápidas, sólo analizan regiones limitadas de los genes *gyrA* y *rrs* (320 pb y 460 pb) para la detección de resistencia solamente a amikacina y fluoroquinolonas, y debido a que no todas las mutaciones resultan en ganancia o pérdida de sitios

de restricción no lograron detectar mutaciones asociadas a resistencia en aproximadamente el 60% de las cepas resistentes que analizaron, además de que dicha prueba no ofrece información alguna sobre la resistencia ante los antibióticos de primera línea (rifampicina, isoniazida, etambutol y pirazinamida).

Otro ejemplo es la prueba Gene Xpert⁶⁹, una prueba rápida de PCR en tiempo real cuya importancia entre las pruebas moleculares se debe a que recientemente ha sido recomendada por la Organización Mundial de la Salud para la detección de tuberculosis resistente. Sin embargo, dicha prueba sólo identifica la presencia de resistencia a rifampicina a través de la detección de mutaciones en el gen *rpoB*, dejando de lado el análisis de la resistencia a etambutol, pirazinamida y otros antibióticos importantes.

Esta limitación también se encuentra presente incluso en otras pruebas que utilizan secuenciación masiva, como la mencionada en párrafos anteriores⁶⁷, donde solamente se analizan cinco de los genes asociados a resistencia (*rpoB*, *katG*, *pncA*, *gyrA* y *rrs*), obviando genes como *embB*, *gyrB*, *rpsL* e *inhA* que también se encuentran asociados con la resistencia a fármacos importantes de primera y segunda línea para el tratamiento de la tuberculosis (Anexo 3).

El uso de la secuenciación como método de análisis genómico le otorga a nuestra prueba la flexibilidad de incluir en su interpretación nuevas variantes genéticas conforme sea clarificada la relevancia que poseen en el manejo clínico de los pacientes. La importancia de este aspecto es puesta en evidencia por las dos nuevas mutaciones que encontramos (*rrs* 842 A>G y 931 A>C) y es subrayada por la reciente descripción de otras nuevas mutaciones en los genes *rpsL* y *rrs* en muestras de *Mycobacterium tuberculosis* de nuestro país⁷⁰ cuya asociación a la resistencia a fármacos, también es desconocida por el momento, pero que seguramente será investigada y determinada a la brevedad.

También realizamos la validación analítica del protocolo propuesto para esta nueva prueba y obtuvimos valores de sensibilidad y especificidad para la detección de resistencia ante los principales antibióticos utilizados en el tratamiento de la tuberculosis, que se encuentren entre 0.75 y 1.00 a excepción de los casos que se mencionan a continuación (ver tabla 5).

En el caso de la detección de resistencia a rifampicina se obtuvo un valor de especificidad de cero, este valor tan bajo se debe a que dentro de las 16 muestras analizadas, sólo una de ellas era fenotípicamente susceptible a dicho fármaco (6717/09), y dicha muestra fue clasificada como resistente al presentar dos mutaciones en *rpoB* (1349 C>T y 1438 A>G). Esto nos indica que para poder apreciar los valores de especificidad que podría alcanzar nuestra prueba para la detección de resistencia a este antibiótico sería conveniente analizar un mayor número de muestras susceptibles a rifampicina.

También para isoniazida se obtuvo una especificidad de cero, que se debe igualmente a que dentro de las 16 muestras analizadas, sólo una era fenotípicamente susceptible a este antibiótico

(la misma muestra 6717/09), y fue clasificada como resistente al presentar una mutación en el promotor de los genes *mabA-inhA* (-15 C>T). En este caso también sería conveniente analizar un mayor número de cepas susceptibles al fármaco para poder realizar un mejor cálculo de la especificidad.

Sería conveniente también volver a analizar con nuestro protocolo de secuenciación y realizar nuevamente las pruebas de susceptibilidad a antibióticos para la ya mencionada cepa 6717/09, ya que en ella se detectó una mutación en el promotor de los genes *mabA-inhA* (-15 C>T) que se encuentra bien caracterizada como uno de los principales mecanismos que otorgan resistencia a isoniazida en *Mycobacterium tuberculosis* (Anexo3) y otras mutaciones adicionales que se piensa podrían otorgar resistencia a rifampicina, estreptomina, pirazinamida y etambutol, a pesar de que esta cepa se reportó como susceptible fenotípicamente ante todos estos antibióticos.

Además, la prueba no fue concluyente para resistencia a pirazinamida y rifampicina de las muestras 28367/08 y 10055/11 respectivamente, ya que presentaron variantes respecto a la referencia que no pudieron clasificarse como verdaderas mutaciones o errores particulares de la plataforma de secuenciación (o del procesamiento según nuestro protocolo), al encontrarse en mediana proporción de las lecturas en las regiones correspondientes y en un contexto genómico poco claro. El volver a analizar nuevamente estas muestras con el protocolo de completo de la prueba probablemente nos permitiría determinar si estas variantes se encontraban realmente presentes en la muestra o se presentaron debido a errores de la polimerasa durante el paso de amplificación de nuestro protocolo o si acaso son errores intrínsecos de la plataforma de secuenciación (lo que parece poco probable debido a que cada una de dichas variantes se encontraron únicamente en una muestra cada una).

Por otro lado, la baja sensibilidad obtenida para la detección de resistencia ante amikacina (0.50) se debió a que de las 16 muestras analizadas, sólo tres fueron fenotípicamente resistentes a este antibiótico, una de ellas (la cepa 9071/10) fue correctamente clasificada como resistente debido a que se encontró en ella una mutación previamente asociada a resistencia (*rrs* 1401 A>G), otra cepa (10061/11) fue clasificada como susceptible ya que no presentó mutaciones en la región analizada, y para la última cepa (6738/10) no fue posible concluir acerca de susceptibilidad o resistencia al presentar una nueva mutación (*rrs* 931 A>C) cuyo impacto en la resistencia a amikacina (u otros fármacos como estreptomina) es desconocido. Esto nos sugiere que para poder evaluar de mejor manera la sensibilidad de la prueba al respecto de este fármaco sería conveniente analizar un mayor número de cepas resistentes a dicho antibiótico.

En otros casos, las mutaciones encontradas en las regiones analizadas no logran explicar la resistencia fenotípica, lo que indica que sería deseable incluir regiones adicionales en nuestra prueba. Por ejemplo, para la detección de resistencia a estreptomina se obtuvo una baja sensibilidad (0.44), que se debe principalmente a que cinco de las muestras que fueron reportadas como fenotípicamente resistentes a estreptomina (muestras 5763/10, 8317/04, 10061/11,

282/12 y 10458/11) solamente presentaron una mutación sinónima en *rpsL* (363 A>G) que, por lo tanto, no logra explicar su resistencia al antibiótico. Este hecho implica que la resistencia de estas cinco muestras se encuentra asociada a regiones del genoma de *Mycobacterium tuberculosis* que no fueron incluidas en nuestro análisis y por lo tanto, la inclusión de otras regiones podría ayudar a mejorar la sensibilidad para la detección de farmacorresistencia ante estreptomycinina.

Del mismo modo, sólo se alcanzó una sensibilidad con valor de 0.50 en el caso de etambutol, ya que de las seis muestras que se reportaron fenotípicamente resistentes a etambutol, en tres de ellas (4624/06, 10095/11, 10301/11) no se encontraron mutaciones que lograran explicar dicha resistencia. Lo que también indica que otras regiones no analizadas deben estar jugando un papel igual de importante que *embB* y para obtener mejores resultados deberían ser incluidas también en el protocolo de la prueba.

De forma preliminar se estima para la prueba un costo de alrededor de \$ 2,500 pesos por muestra (tabla 6), incluyendo la extracción del DNA, la amplificación por PCR de los genes de resistencia, su purificación y fragmentación, la preparación de bibliotecas de secuenciación de nueva generación, su secuenciación, determinación de variantes y la determinación de resistencia a los siete antibióticos principales para el tratamiento de la tuberculosis. Sin embargo, la aparición de nuevas tecnologías y variantes metodológicas para realizar el análisis genómico nos permiten desde ahora plantear mejoras al proceso para reducir este costo.

Tabla 6. Costo estimado de la prueba propuesta.

Concepto	Costo por muestra (pesos)
Extracción de DNA	\$ 50.65
PCR	\$ 110.10
Electroforesis	\$ 118.38
Purificación de los productos de PCR	\$ 656.24
Preparación de bibliotecas, secuenciación y determinación de resistencia	\$ 1,625.00
Total	\$ 2,560.36

Por último queremos resaltar que aunque hemos propuesto una nueva prueba molecular para el diagnóstico rápido de resistencia a fármacos en infecciones por *Mycobacterium tuberculosis*, su adopción en el ámbito clínico mexicano dependerá de la realización de rigurosos estudios subsecuentes para su validación en diversos grupos de casos y controles y la determinación de su costo efectividad en el entorno clínico.

10. Referencias

- ¹ Mardis ER. A decade's perspective on DNA sequencing technology. *Nature*. 2011 Feb 10;470(7333):198-203.
- ² Ntzani EE, Liberopoulos G, Manolio TA, Ioannidis JP. Consistency of genome-wide associations across major ancestral groups. *Hum Genet*. 2011 Dec 20. DOI 10.1007/s00439-011-1124-4
- ³ Seputienė V, Vilkoicaitė A, Armalytė J, Pavilionis A, Sužiedėlienė E. Detection of methicillin-resistant *Staphylococcus aureus* using double duplex real-time PCR and dye Syto 9. *Folia Microbiol (Praha)*. 2010 Sep;55(5):502-7. Epub 2010 Oct 13.
- ⁴ Selvaraju SB, Selvarangan R. Evaluation of three influenza A and B real-time reverse transcription-PCR assays and a new 2009 H1N1 assay for detection of influenza viruses. *J Clin Microbiol*. 2010 Nov;48(11):3870-5. Epub 2010 Sep 15.
- ⁵ Bell CJ, Dinwiddie DL, Miller NA, Hateley SL, Ganusova EE, Mudge J, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci Transl Med*. 2011 Jan 12;3(65):65ra4.
- ⁶ Metzgar D, Myers CA, Russell KL, Faix D, Blair PJ, Brown J, et al. Single assay for simultaneous detection and differential identification of human and avian influenza virus types, subtypes, and emergent variants. *PLoS One*. 2010 Feb 3;5(2):e8995.
- ⁷ Wong CW, Heng CL, Wan Yee L, Soh SW, Kartasasmita CB, Simoes EA, et al. Optimization and clinical validation of a pathogen detection microarray. *Genome Biol*. 2007;8(5):R93.
- ⁸ Erlandsson L, Rosenstjerne MW, McLoughlin K, Jaing C, Fomsgaard A. The microbial detection array combined with random Phi29-amplification used as a diagnostic tool for virus detection in clinical samples. *PLoS One*. 2011;6(8):e22631. Epub 2011 Aug 10.
- ⁹ Green ED, Guyer MS. Charting a course for genomic medicine from base pairs to bedside. *Nature*. 2011 Feb 10;470(7333):204-13.
- ¹⁰ World Health Organization. The evolving threat of antimicrobial resistance - Options for action. World Health Organization. Ginebra, Suiza. 2012
- ¹¹ Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011 Mar 24;471(7339):467-72.
- ¹² Ioannidis JP. A roadmap for successful applications of clinical proteomics. *Proteomics Clin Appl*. 2011 Jun;5(5-6):241-7. doi: 10.1002/prca.201000096. Epub 2011 Apr 27
- ¹³ Mertes F, Elsharawy A, Sauer S, van Helvoort JM, van der Zaag PJ, Franke A, et al. Targeted enrichment of genomic DNA regions for next-generation sequencing. *Brief Funct Genomics*. 2011 Nov;10(6):374-86. doi: 10.1093/bfpg/eln033. Epub 2011 Nov 26.
- ¹⁴ Cooper GM, Shendure J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet*. 2011 Aug 18;12(9):628-40. doi: 10.1038/nrg3046.
- ¹⁵ Jain, Kewal K. Textbook of personalized Medicine. Springer Verlag. New York, United States. 2010.
- ¹⁶ Hampikian G, West E, Akselrod O. The genetics of innocence: analysis of 194 U.S. DNA exonerations. *Annu Rev Genomics Hum Genet*. 2011 Sep 22;12:97-120.
- ¹⁷ Lee JH, Jeon JT. Methods to detect and analyze copy number variations at the genome-wide and locus-specific levels. *Cytogenet Genome Res*. 2008;123(1-4):333-42. Epub 2009 Mar 11.
- ¹⁸ Miller MB, Tang YW. Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Rev*. 2009 Oct;22(4):611-33.
- ¹⁹ Fluidigm Corporation. Presentación de ventas para el Instituto Nacional de Medicina Genómica (INMEGEN). (22/09/11). Distrito Federal, México.
- ²⁰ Life Technologies. Ion Torrent products. <http://www.iontorrent.com/products-ion-pgm/> (accesado el 28/11/11).
- ²¹ Illumina. Applications / Genome-Wide Association Studies (GWAS). Omni BeadChips. <http://www.illumina.com/applications/gwas.ilmn> (accesado el 28/11/11).
- ²² Pont-Kingdon G, Gedge F, Wooderchak-Donahue W, Schrijver I, Weck KE, Kant JA, et al. Design and analytical validation of clinical DNA sequencing assays. *Arch Pathol Lab Med*. 2012 Jan;136(1):41-6.
- ²³ National Human Genome Research Institute. DNA sequencing costs. Data from the NHGRI large-scale genome sequencing program. <http://www.genome.gov/sequencingcosts/> (Accesado el 29/09/2011).
- ²⁴ Koboldt DC, Ding L, Mardis ER, Wilson RK. Challenges of sequencing human genomes. *Brief Bioinform*. 2010 Sep;11(5):484-98. Epub 2010 Jun 2.
- ²⁵ Glenn TC. Field guide to next-generation DNA sequencers. *Mol Ecol Resour*. 2011 Sep;11(5):759-69. doi: 10.1111/j.1755-0998.2011.03024.x. Epub 2011 May 19.

-
- ²⁶ Lennon NJ, Lintner RE, Anderson S, Alvarez P, Barry A, Brockman W, et al. A scalable, fully automated process for construction of sequence-ready barcoded libraries for 454. *Genome Biol.* 2010;11(2):R15. Epub 2010 Feb 5.
- ²⁷ Bainbridge MN, Wang M, Burgess DL, Kovar C, Rodesch MJ, D'Ascenzo M, et al. Whole exome capture in solution with 3 Gbp of data. *Genome Biol.* 2010;11(6):R62. Epub 2010 Jun 17.
- ²⁸ Montenegro G, Powell E, Huang J, Speziani F, Edwards YJ, Beecham G, et al. Exome sequencing allows for rapid gene identification in a Charcot-Marie-Tooth family. *Ann Neurol.* 2011 Mar;69(3):464-70. doi: 10.1002/ana.22235. Epub 2011 Jan 20.
- ²⁹ Maxmen A. Exome sequencing deciphers rare diseases. *Cell.* 2011 Mar 4;144(5):635-7.
- ³⁰ Ray T. AssureRx Updates PGx Test for Assessing Best Responders to Psychotropic Drugs. *Pharmacogenomics Reporter.* GenomeWeb LLC. January 18, 2012.
http://www.genomeweb.com//node/1016321?hq_e=el&hq_m=1176331&hq_l=4&hq_v=3c165a225f
- ³¹ Ragoussis J. Genotyping technologies for genetic research. *Annu Rev Genomics Hum Genet.* 2009;10:117-33.
- ³² Sierra C, Sánchez E, Henao S, Saavedra A. Determinación de la susceptibilidad a drogas de primera línea en aislamientos de *M. tuberculosis* por la técnica del tubo indicador de crecimiento micobacteriano. *Rev Fac Med.* 2008;56:11-20.
- ³³ Cuevas-Córdoba B, Zenteno-Cuevas R. Tuberculosis drogoresistente: mecanismos moleculares y métodos diagnósticos. *Enferm Infecc Microbiol Clin.* 2010;28(9):621-628.
- ³⁴ Palomino JC. Molecular detection, identification and drug resistance detection in *Mycobacterium tuberculosis*. *FEMS Immunol Med Microbiol.* 2009 Jul;56(2):103-11.
- ³⁵ Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ.* 1994;72(3):429-45.
- ³⁶ World Health Organization. Health statistics and health information systems. Global Burden of Disease. Disease and injury country estimates. http://www.who.int/healthinfo/global_burden_disease/estimates_country/en/index.html
- ³⁷ México: Secretaría de Salud. Diagnóstico, metas de control ambulatorio y referencia oportuna de la diabetes mellitus tipo 2 en el primer nivel de atención. 2008
- ³⁸ Life technologies. Ion Xpress Plus gDNA and Amplicon Library Preparation. Publication Part Number 4471989 Rev. B. 2011
- ³⁹ Babraham bioinformatics. FASTQC A quality control tool for highthroughput sequence data.
<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- ⁴⁰ Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* Apr 19.
- ⁴¹ Reddy TB, Riley R, Wymore F, Montgomery P, Decaprio D, Engels R, et al. TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res* 2009 Jan;37(Database issue):D499-508. Epub 2008 Oct 3.
- ⁴² Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. *PLoS Med.* 2009 Feb 10;6(2):e(2).
- ⁴³ Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073-81.
- ⁴⁴ McCarthy M. Genomics, Type 2 Diabetes and Obesity. *N Engl J Med* 2010;363:2339-50.
- ⁴⁵ Waterfield T, Gloyn AL. Monogenic β -cell dysfunction in children: clinical phenotypes, genetic etiology and mutational pathways. *Pediatr Health* 2008;2:517-32.
- ⁴⁶ Harris TJ, McCormick F. The molecular pathology of cancer. *Nat Rev Clin Oncol.* 2010 May;7(5):251-65. Epub 2010 Mar 30.
- ⁴⁷ Wong KM, Hudson TJ, McPherson JD. Unraveling the genetics of cancer: genome sequencing and beyond. *Annu Rev Genomics Hum Genet.* 2011 Sep 22;12:407-30.
- ⁴⁸ Braggio E, Fonseca R. Unraveling the multiple myeloma genome in the next-generation sequencing era: challenges to translating knowledge into the clinic. *Expert Rev Hematol.* 2011 Dec;4(6):579-81. doi: 10.1586/ehm.11.64.
- ⁴⁹ Technology Strategy Board. Stratified medicines programme: Tumour profiling and data capture to improve cancer care. Competition for collaborative R&D funding. January 2011. www.innovateuk.org (Accesado el 06//07//2011).
- ⁵⁰ Talmud PJ, Hingorani AD, Cooper JA, Marmot MG, Brunner EJ, Kumari M, et al. Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study. *BMJ.* 2010 Jan 14;340:b4838. doi: 10.1136/bmj.b4838.
- ⁵¹ Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, et al. Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med.* 2008 Nov 20;359(21):2208-19.

- ⁵² Distefano JK, Watanabe RM. Pharmacogenetics of Anti-Diabetes Drugs. Pharmaceuticals (Basel). 2010 Aug 1;3(8):2610-2646.
- ⁵³ World Health Organization. 2011/2012 Tuberculosis Global Facts. Disponible en www.who.int/tb Accesado el 13/01/2012.
- ⁵⁴ Abbate E, Vescovo M, Natiello M, Cufre M, García A, Ambroggi M, Poggi S, Símboli N, Ritacco V. Tuberculosis Extensamente Resistente en Argentina Aspectos epidemiológicos, bacteriológicos, terapéuticos y evolutivos. *Revista Argentina de Medicina Respiratoria* 2007 – Nº 1: 19-25.
- ⁵⁵ Luo T, Zhao M, Li X, Xu P, Gui X, Pickerill S, DeRiemer K, Mei J, Gao Q. Selection of mutations to detect multidrug-resistant *Mycobacterium tuberculosis* strains in Shanghai, China. *Antimicrob Agents Chemother*. 2010 Mar;54(3):1075-81. Epub 2009 Dec 14.
- ⁵⁶ Dirección General de epidemiología, Instituto de Diagnóstico y Referencia Epidemiológicos, Dirección General adjunta de Programas Preventivos, Instituto nacional de Salud Pública. Avances de la Encuesta Nacional de Farmacorresistencia en Tuberculosis. 2010. Disponible en: http://www.cenave.gob.mx/tuberculosis/CURSO%20HG/2010/LUNES/Encuesta%20Nacional%20TB%20MDR_20mar10.pdf
- ⁵⁷ Järvinen A, Laakso S, Piiparinen P, Aittakorpi A, Lindfors M, Huopaniemi L, et al. Rapid identification of bacterial pathogens using a PCR- and microarray-based assay. *BMC Microbiol*. 2009; 9: 161. Published online 2009 August 10. doi: 10.1186/1471-2180-9-161
- ⁵⁸ Melnikov A, Galinsky K, Rogov P, Fennell T, Tyne DV, Russ C, et al. Hybrid selection for sequencing pathogen genomes from clinical samples. *Genome Biol*. 2011 Aug 11;12(8):R73. [Epub ahead of print]
- ⁵⁹ Greninger AL, Chen EC, Sittler T, Scheinerman A, Roubinian N, Yu G, et al. A metagenomic analysis of pandemic influenza A (2009 H1N1) infection in patients from North America. *PLoS One*. 2010 Oct 18;5(10):e13381.
- ⁶⁰ Kozhamkulov U, Akhmetova A, Rakhimova S, Belova E, Alenova A, Bismilda V, et al. Molecular characterization of rifampicin- and isoniazid-resistant *Mycobacterium tuberculosis* strains isolated in Kazakhstan. *Jpn J Infect Dis*. 2011;64(3):253-5.
- ⁶¹ Campbell PJ, Morlock GP, Sikes RD, Dalton TL, Metchock B, Starks AM, Hooks DP, Cowan LS, Plikaytis BB, Posey JE. Molecular detection of mutations associated with first- and second-line drug resistance compared with conventional drug susceptibility testing of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2011 May;55(5):2032-41. Epub 2011 Feb 7.
- ⁶² Cui Z, Wang J, Lu J, Huang X, Hu Z. Association patterns in *gyrA/B* genes and ofloxacin resistance levels in *Mycobacterium tuberculosis* isolates from East China in 2009. *BMC Infect Dis*. 2011 Mar 29;11:78. Doi: 10.1186/1471-2334-11-78.
- ⁶³ Sekiguchi J, Miyoshi-Akiyama T, Augustynowicz-Kopec E, Zwolska Z, Kirikae F, Toyota E, Kobayashi I, Morita K, Kudo K, Kato S, Kuratsuji T, Mri T, Kirikae T. Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol*. 2007 Jan;45(1):179-92. Epub 2006 Nov 15.
- ⁶⁴ Shi D, Li L, Zhao Y, Jia Q, Li H, Coulter C, jin Q, Zhu G. Characteristics of *embB* mutations in multidrug-resistant *Mycobacterium tuberculosis* isolates in Henan, China. *J Antimicrob Chemother* 2011; 66: 2240-2247.
- ⁶⁵ Stevens G, Dias RH, Thomas KJ, Rivera JA, Carvalho N, Barquera S, et al. Characterizing the epidemiological transition in Mexico: national and subnational burden of diseases, injuries, and risk factors. *PLoS Med*. 2008 Jun 17;5(6):e125.
- ⁶⁶ Pareek CS, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet*. 2011 Nov;52(4):413-35. doi: 10.1007/s13353-011-0057-x. Epub 2011 Jun 23.
- ⁶⁷ Daum LT, Rodríguez JD, Worthy SA, Ismail NA, Omar SV, Dreyer AW, Fourie PB, Hoosen AA, Chambers JP, Fischer GW. Next-Generation Ion Torrent Sequencing of Drug Resistance Mutations in *Mycobacterium tuberculosis* Strains. *J Clin Microbiol*. 2012 Dec;50(12):3831-7.
- ⁶⁸ Tahmasebi P, Farnia P, Sheikholeslami F, Velavati A. Rapid Identification of extremely drug resistant tuberculosis from multidrug resistant strains; using PCR-RFLP and PCR-SSCP. *Iran J Microbiol*. 2012 Dec;4(4):165-70.
- ⁶⁹ Steingart KR, Sohn H, Schiller I, Kloda LA, Boehme CC, Pai M, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev*. 2013. Jan 31;1:CD009593. doi: 10.1002/14651858.CD009593.pub2
- ⁷⁰ Cuevas-Córdoba B, Cuellar-Sánchez A, Pasissi-Crivelli A, Santana-Álvarez CA, Hernández-Illezcas J, Zenteno-Cuevas R. *rrs* and *rpsL* mutations in streptomycin-resistant isolates of *Mycobacterium tuberculosis* from Mexico. *J Microbiol Immunol Infect*. 2012 Oct 2. doi:pii: S1684-1182(12)00181-8. 10.1016/j.jmii.2012.08.020. [Epub ahead of print]

11. ANEXOS

ANEXO 1. Resultados de la evaluación por enfermedad.

A continuación se presentan, para cada enfermedad evaluada, los correspondientes valores de DALYs por cien mil habitantes en 2004, aquellos valores que hemos asignado a la factibilidad en el corto plazo y al efecto esperado y también un pequeño resumen de las características de las enfermedades que llevaron a la asignación de estos últimos valores.

Heridas no intencionales.

DALYs: 1356 DALYs/100,000 hab.

Factibilidad en el corto plazo: No identificada.

Efecto esperado sobre la enfermedad: Impacto medio.

Características: Podría darse el caso de que existan asociaciones indirectas a blancos genéticos mediante la presencia de enfermedades en algunos subtipos de estas heridas, por ejemplo entre heridas resultantes de caídas y algunos padecimientos del oído, pero este tipo de asociaciones no han sido estudiadas a profundidad. No se han encontrado estudios que hayan demostrado asociación entre factores genéticos y heridas no intencionales. El tratamiento de las heridas no intencionales del mismo tipo es el mismo en prácticamente todos los casos aunque la respuesta al tratamiento podría ser diferente según el individuo del que se trate. Como medida preventiva se tendría solamente el cuidado con el que se realicen las actividades, que, en teoría, debiera ser alto por sentido común. Una prueba para identificar el riesgo, el pronóstico y/o efectividad del tratamiento de estos padecimientos tendría que aplicarse a una gran parte de la población, ya que prácticamente todos están expuestos a este tipo de heridas de una forma u otra.

Diabetes Mellitus subtipos poliagénicos.

DALYs: 789 DALYs/100,00 hab. (considerando un 98.5% del total de los DALYs de diabetes)

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se han identificado blancos genéticos que parecen aumentar el riesgo de presentar diabetes mellitus. El tratamiento es en base a dieta, insulina e hipoglucemiantes orales. Existen medidas preventivas, como podrían ser cambio de dieta y estilo de vida para las personas

con alto riesgo a presentar estos padecimientos. Los pacientes se benefician debido a un diagnóstico más temprano al evitar daño macrovascular, neuropatía, nefropatía o retinopatía.

Existe gran controversia acerca de la utilidad de pruebas genéticas sobre los factores de riesgo utilizados actualmente en la práctica clínica (índice de masa corporal, parientes de primer grado con diabetes, etc.).^{1,2,3,4}

Neoplasias malignas

DALYs: 789 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Debido a la heterogeneidad que existe entre los diversos tipos de tumores y a que las mutaciones que dan lugar a ellos son somáticas en su mayoría, es decir, que no se presentan en células de la línea germinal es difícil pensar en la utilización de herramientas de análisis genómico para el diagnóstico de este tipo de enfermedades. Por otro lado se han encontrado blancos genéticos que ayudan a establecer una mayor predisposición para presentar ciertos tipos de cáncer (por ejemplo, mutaciones en BRCA1 y BRCA2 se asocian a un mayor riesgo de presentar cáncer cervicouterino y de mama). También se cree que puede haber un gran potencial en observar el pronóstico y respuesta a tratamiento de pacientes que compartan genes con mutaciones deletéreas para así poder determinar en un futuro los tratamientos más efectivos para ellos (entre algunos genes con mutaciones que ya han comenzado a investigarse se encuentran EGFR, KRAS, ALK, BRCA1, BRCA2, etc.). Los pacientes se benefician de un diagnóstico más temprano así como de la selección del tratamiento más efectivo para ellos, también existen medidas preventivas que van desde cambios en el estilo de vida hasta cirugía preventiva.^{5,6,7,8}

Trastornos depresivos unipolares

DALYs: 772 DALYs/100,000 hab.

Factibilidad en el corto plazo: Difícil.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se ha encontrado evidencia de que existe cierta asociación entre factores genéticos y la presencia de trastornos depresivos pero no se ha logrado determinar blancos genéticos que modifiquen de manera importante el riesgo de presentar estos padecimientos. La

razón principal por la cual no se han identificado todavía dichos blancos es la heterogeneidad de estos síndromes y el hecho de que la mayor parte tienen orígenes desconocidos. Un área importante de estudio es la farmacogenómica ya que se han encontrado algunos polimorfismos que afectan en cierta medida la respuesta a fármacos antidepresivos. Existen tratamientos más efectivos para unos pacientes que para otros y los pacientes se benefician de un diagnóstico más temprano.^{9,10}

Anomalías congénitas

DALYs: 713 DALYs/100,000 hab.

Factibilidad en el corto plazo: Segura.

Efecto esperado sobre la enfermedad: Impacto medio.

Características: Es un grupo heterogéneo de padecimientos y es común que se presenten varios de ellos en un mismo individuo, lo que aumenta en gran medida la mortalidad. Para varios de ellos se han encontrado blancos genéticos asociados fuertemente a su presencia. Muchos conllevan como tratamiento cirugía en todos sus pacientes. Las medidas preventivas se enfocarían principalmente en la toma de decisiones reproductivas y en la nutrición y estilo de vida de la madre antes y durante el embarazo. Es difícil que se pueda tener un diagnóstico más temprano ya que actualmente éste se realiza durante el embarazo o poco después del parto, pero podría evaluarse el riesgo de los padres de concebir un hijo afectado. El diagnóstico molecular es considerado todavía demasiado invasivo como para realizarlo a la totalidad de los embarazos y actualmente se utilizan varias pruebas para determinar los casos donde deban llevarse a cabo pruebas bioquímicas. Próximos avances tecnológicos quizás ayuden a remediar esta situación, por ejemplo la detección de células del producto en el torrente sanguíneo materno, pero mientras tanto este hecho es de las principales desventajas para la adopción del diagnóstico genómico en estos padecimientos.^{11,12,13,14,15,16}

Enfermedad isquémica del corazón

DALYs: 538 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Ya se han identificado algunas mutaciones de genes específicos que causan enfermedad arterial coronaria e infarto al miocardio. También existe gran información acerca de

blancos genéticos que aumentan el riesgo de presentar este tipo de enfermedades. Es grande la variabilidad en eficacia del tratamiento para estos padecimientos que pueden ser a través de fármacos (anticoagulantes, vasodilatadores, etc.) o quirúrgicos si la ocasión lo amerita. Es difícil pensar en un diagnóstico más temprano de estos padecimientos a través de una prueba genómica, pero mediante este enfoque se podrían identificar las personas con alto riesgo de presentarlos, e indicar las medidas preventivas pertinentes.^{17,18,19}

Heridas intencionales

DALYs: 519 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto medio.

Características: La relación que existe entre heridas intencionales y factores genéticos no ha sido estudiada a profundidad. Se han encontrado algunos blancos genéticos asociados a comportamiento antisocial, rasgos violentos y agresividad sólo en individuos con determinada influencia ambiental, por ejemplo maltrato durante la infancia. Necesitan realizarse estudios para comprobar estas asociaciones debido a que no se han comprobado en poblaciones diferentes a la caucásica, quizás por falta de poder estadístico. Es difícil pensar en un diagnóstico más temprano y en la selección de un tratamiento más adecuado para las víctimas, pero podrían existir medidas preventivas para evitar la conductas de individuos con alto riesgo de ocasionarlas (como podrían ser intervenciones psicosociales preventivas donde se desarrollarían opciones no agresivas para que los individuos puedan lidiar con los estímulos ambientales que, en otro caso, causarían estos comportamientos).^{20,21}

Cirrosis hepática

DALYs: 443 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se han identificado blancos genéticos asociados a este padecimiento, algunos aumentan el riesgo de presentarlo mientras que otros más lo disminuyen, pero todavía se investiga acerca de su utilidad clínica. Para la mayoría de los blancos genéticos se ha encontrado una asociación débil, probablemente debido a que los factores ambientales juegan un papel muy importante en el desarrollo de estos padecimientos. Existe gran variabilidad en la eficacia del

tratamiento, que puede ser a base de cambio de estilo de vida, uso de fármacos e incluso quirúrgico. Existen medidas preventivas para personas susceptibles, como el evitar factores de riesgo adicionales, y los pacientes se benefician de un diagnóstico más temprano. En México la cirrosis hepática se debe principalmente a dos causas: al consumo de alcohol y a secuelas crónicas de infecciones virales. De éstas últimas los principales virus que se relacionan con este padecimiento son los virus de hepatitis B y C. En los pacientes con hepatitis crónica se presenta un proceso de fibrosis que puede durar hasta 20 o 30 años en las primeras etapas y después de este tiempo (al presentarse un disparador) progresa rápidamente aumentando el daño al hígado. Existe un blanco genético (SNP conocido como interleucina 28) que aumenta en gran medida la probabilidad, en pacientes con hepatitis C crónica, de que el tratamiento con interferón pegilado alfa 2B y rivabirina evite el progreso de la enfermedad a cirrosis hepática (aproximadamente de un 40% de probabilidad de respuesta al tratamiento a cerca de un 80%).^{22,23,24}

Prematurez y bajo peso al nacer

DALYs: 376 DALYs/100,000 hab.

Factibilidad en el corto plazo: No identificada.

Efecto esperado sobre la enfermedad: Impacto medio.

Aunque hay evidencias de que existen factores genéticos asociados con prematurez y bajo peso al nacer, no se ha estudiado a fondo dicha asociación y sólo se han propuesto algunos genes que podrían intervenir. La razón principal por la cual se ha dificultado el definir claramente blancos genéticos que determinen la presencia de estos padecimientos es debido a un gran componente ambiental que actúan incluso mediante las condiciones intrauterinas maternas. Existen medidas preventivas para evitar estos padecimientos, sobre todo en base a suplementos nutricionales y estilo de vida antes y durante el embarazo. Sería difícil pensar en diagnosticar más temprano estos padecimientos, pero sería de utilidad la determinación *a priori* del riesgo de los padres de concebir un hijo afectado.

Infecciones de las vías respiratorias inferiores

DALYs: 354 DALYs/100,000 hab.

Factibilidad en el corto plazo: Segura.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se han encontrado algunos blancos genéticos que aumentan el riesgo a presentar este tipo de padecimientos y existe una gran gama de microorganismos (tanto virus, bacterias y hongos) que los causan. Existen antibióticos recomendados para patógenos específicos pero quizás es de mayor utilidad distinguir aquellos casos con causas virales de los demás, así como identificar resistencia a ciertos fármacos en los microorganismos presentes para conocer exactamente el tratamiento que debe utilizarse. También existen medidas preventivas que podrían tomar las personas en riesgo, como vacunación y hábitos de exposición a factores ambientales de riesgo. Los pacientes se beneficiarían de un diagnóstico más temprano. Actualmente son muy utilizados los métodos basados en PCR para identificar los agentes causantes de estos padecimientos.^{25,26,27,28,29,30,31,32,33,34,35}

Asma

DALYs: 328 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se ha encontrado asociación entre la presencia de asma y varios blancos genéticos, por ejemplo algunos polimorfismos en genes que intervienen en los procesos inflamatorios, sin embargo, este padecimiento está asociado de manera importante a factores ambientales. Existen medidas preventivas que podrían tomar los individuos en riesgo, tales como disminuir la exposición a factores ambientales y los pacientes se benefician de un diagnóstico más temprano.^{36,37}

Asfixia y trauma del nacimiento

DALYs: 327 DALYs/100,000 hab.

Factibilidad en el corto plazo: No identificada.

Efecto esperado sobre la enfermedad: Impacto medio.

Características: La relación que existe entre asfixia y trauma del nacimiento con factores genéticos no ha sido estudiada a profundidad. El estudio de estos padecimientos se ha enfocado en cómo afectan el riesgo de presentar otras enfermedades, por ejemplo problemas del desarrollo. Es difícil pensar en un diagnóstico más temprano y en escoger un tratamiento que resulte más efectivo en un caso particular. Por otro lado, existen ciertas medidas preventivas que podrían tomarse si se determina una mayor probabilidad de presentar estos padecimientos, en tales casos se podría

ordenar el monitoreo continuo y más riguroso del embarazo y preparar de antemano las acciones a tomar en el momento del parto, o incluso decidir la realización de cesárea.

Infecciones neonatales.

DALYs: 322 DALYs/100,000 hab.

Factibilidad en el corto plazo: Segura.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Es el mismo caso que las enfermedades infecciosas de las vías respiratorias inferiores: Se han encontrado algunos blancos genéticos que aumentan el riesgo a presentar este tipo de padecimientos y existe una gran gama de microorganismos (tanto virus, bacterias y hongos) que los causan. Es de gran utilidad distinguir aquellos casos con causas virales de los demás, así como identificar resistencia a ciertos fármacos en los microorganismos presentes para conocer exactamente el tratamiento que debe utilizarse. También existen medidas preventivas que podrían tomar las personas en riesgo, como vacunación y hábitos de exposición a factores ambientales de riesgo. Además, los pacientes se beneficiarían de un diagnóstico más temprano.^{25,26,27,28,29,30,31,32,33}

Desórdenes endocrinos

DALYs: 293 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Hay mutaciones causantes y polimorfismos que aumentan el riesgo de presentar varios padecimientos endócrinos autoinmunes. Estos padecimientos afectan a una pequeña proporción de la población y su tratamiento puede constar de cirugía, suplemento de hormonas, isótopos radiactivos y fármacos de diferentes tipos para controlar los síntomas con diferentes niveles de eficacia. Estos padecimientos tienen una alta variabilidad en sus fenotipos, lo que hace difícil la identificación de los blancos genéticos que las causan. Además, tienen un componente ambiental muy importante, ya que es incluso más relevante que el componente genético en algunos casos. Los individuos en riesgo podrían tomar medidas preventivas y los pacientes se benefician de un diagnóstico más temprano.^{38,39,40}

Esquizofrenia

DALYs: 277 DALYs/100,000 hab.

Factibilidad en el corto plazo: Probable.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se han encontrado blancos genéticos que aumentan en gran medida el riesgo de presentar esquizofrenia. Este padecimiento es muy heterogéneo en cuanto a riesgo, síntomas y respuesta al tratamiento, y todavía es necesario más estudio para identificar los blancos genéticos que determinan estas diferencias. Por otro lado, los pacientes sí reciben beneficios de un diagnóstico más temprano en el curso de la enfermedad.^{41,42}

Diabetes Mellitus subtipos monogénicos.

DALYs: 12 DALYs/100,000 hab. (considerando un 1.5% del total de los DALYs de diabetes)

Factibilidad en el corto plazo: Segura.

Efecto esperado sobre la enfermedad: Impacto drástico.

Características: Se han identificado blancos genéticos que determinan directamente la presencia de los subtipos monogénicos de diabetes. Éstos son mutaciones en diversos genes para cada uno de los diferentes subtipos de diabetes monogénicas (MODY: NEUROD1, GCK, HNF1A, IPF1, HNF1B, HNF4A, diabetes mitocondrial y diabetes neonatal: LMNA, ALMS1, EIF2AK3, PPARG, WFS1, WFS2, HFE, RFX6, ZAC, CAV1, FXN, CEL, AGPAT2, INS, KCNJ11, ABCC8, BSCL2, LMNB2, INSR, AKT2, mitocondrial MTTL1). Los pacientes se benefician de un diagnóstico temprano. Debido a que estas mutaciones se transmiten en forma mendeliana las medidas preventivas se enfocarían a la toma de decisiones reproductivas. Los subtipos monogénicos de diabetes representan aproximadamente el 1-2 % de todos los casos de diabetes.^{3,1}

ANEXO 2. Secuencias de los cebadores utilizados para la amplificación de las regiones analizadas.

A continuación se presenta una tabla con las secuencias de los cebadores que se utilizaron para la amplificación por PCR de las regiones de resistencia de *Mycobacterium tuberculosis* analizadas en la prueba.

Tabla 7. Secuencias de “primers” utilizados para la amplificación por PCR de cada región.

Región	Forward Primer (5'→3')	Reverse Primer (5'→3')	Tamaño de amplicón (bp)
katG ⁴³	GCGAACTCGTCGGCCAATTC	AACGACGTCGAAACAGCGGC	455
mabA/inhA (promotor) ⁴³	TGCCCAGAAAGGGATCCGTCATG	ATGAGGAATGCGTCCGCGGA	455
rpoB ⁴³	CTTGACGAGGGTCAGACCA	ATCTCGTCGCTAACCACGCC	543
embB ⁴³	CTGACCGACGCCGTGGTGATAT	TGAATGCGGCGGTAACGACG	490
pncA ⁴³	ACAGTTCATCCCGTTTCGGC	GGTCATGTTTCGCGATCGTCG	689
gyrA ⁴³	CCCTGCGTTCGATTGCAAAC	CTTCGGTGTACCTCATCGCC	423
gyrB ⁴⁴	GCGCTGACGTCGGTGGTGAA	ATTCCGGGTCCTGCGCTGC	772
rpsL ⁴⁵	CCAACCATCCAGCAGCTGGT	GTCGAGAGCCCGCTTGAGGG	572
rrs ⁴⁵	AAACCTCTTTCACCATCGAC	GTATCCATTGATGCTCGCAA	1328

ANEXO 3. Antibióticos para el tratamiento de tuberculosis y mecanismos de resistencia considerados en este trabajo.

El tratamiento de la tuberculosis implica una terapia combinatoria con duración mínima de 6 meses. Los cuatro antibióticos de primera línea que son utilizados para tratar todas las formas de tuberculosis excepto aquellas causadas por organismos que se sabe son resistentes a ellos son: isoniacida, rifampicina, pirazinamida y etambutol.

La isoniacida (INH) es un potente fármaco antituberculoso y bactericida que actúa específicamente contra bacterias que se encuentran en fase de replicación activa, ya que inhibe la síntesis lipídica necesaria para la formación de la pared celular.⁴⁶ Una vez administrada la isoniacida es transformada *in vivo* a su principio activo, mediante la acción de la enzima catalasa peroxidasa. Después, al actuar sobre su blanco InhA (una enoil acil protein reductasa dependiente de NADH), la isoniacida inhibe la síntesis de ácido micólico de la pared celular bacteriana, lo cual permite que el microorganismo sea susceptible a la acción de radicales de oxígeno reactivos y a otros elementos externos de respuesta del huésped.⁴⁷

La gran mayoría de la resistencia a isoniacida se debe a la presencia de mutaciones o deleciones en dos regiones genómicas. La primera de ellas es el gen *katG* (que codifica la catalasa peroxidasa responsable de la activación del fármaco), donde la resistencia a isoniacida surge como consecuencia de la pérdida de la actividad de la proteína y la segunda es la región promotora del operón *mabA-inhA* donde se encuentra codificada la proteína diana del fármaco (InhA) cuya sobreexpresión confiere resistencia a través de la titulación del fármaco activo.⁴⁸

La rifampicina (RIF) es un fármaco antituberculoso que presenta actividad contra organismos que se encuentran en fase de replicación activa, así como también contra poblaciones semilantes, debido a su acción a través de la inhibición de la RNA polimerasa de los organismos infectantes.⁴⁶ Este fármaco se une a la subunidad beta de la RNA polimerasa, bloqueando físicamente la elongación de la cadena de RNA y, por lo tanto, inhibiendo la transcripción. En cuanto a las regiones genómicas involucradas en el desarrollo de resistencia, es bien conocido que las mutaciones en el gen *rpoB* (que codifica para dicha subunidad de la RNA polimerasa), pueden producir cambios conformacionales que disminuyen la afinidad por rifampicina y, por lo tanto, conferir diferentes niveles de resistencia al fármaco.^{47,48}

Por otro lado, la pirazinamida (PZA) ejerce su mayor efecto sobre las poblaciones bacterianas latentes o semilantes que difícilmente son afectadas por otros medicamentos antituberculosos⁴⁶ y una de sus principales ventajas es que presenta sinergia con isoniazida y rifampicina, lo que disminuye el tiempo de tratamiento necesario.⁴⁷

Respecto a su modo de acción, la pirazinamida también es una prodroga y, por lo tanto, requiere ser transformada a su forma activa (ácido pirazinoico) el cual es responsable del efecto antimicrobiano, probablemente a través de la disrupción del potencial de membrana.⁴⁸ El principal mecanismo por el cual se genera resistencia a pirazinamida es el surgimiento de mutaciones, inserciones y deleciones en el gen *pncA* que causan la pérdida de función de la pirazinamidasa, enzima responsable de la activación del fármaco que se encuentra codificada en dicho gen.

Otro fármaco que se incluye en la pauta básica de tratamiento primario de la tuberculosis es el etambutol (EMB), aunque normalmente no se recomienda su uso en pacientes pediátricos a menos que se sospeche que los organismos causantes son resistentes a izoniazida o rifampicina.⁴⁶

Específicamente, el etambutol interfiere en la síntesis de arabinogalactanos mediante la inhibición de la arabinosil transferasa (evitando la formación del complejo micolil-arabinogalactano-peptidoglicano cuya

función es incrementar la permeabilidad de la pared celular). El principal mecanismo por el cual se genera resistencia a etambutol es a través de la aparición de mutaciones en el gen *embB* que causan la alteración estructural de la arabinosil transferasa codificada en dicho gen.⁴⁸

Además de los cuatro agentes de primera línea mencionados en los párrafos anteriores, existen otras alternativas de tratamiento que son utilizadas principalmente en aquellos casos donde no es posible utilizar alguno de los fármacos de primera línea, ya sea por aparición de resistencia a ellos o intolerancia por parte del paciente.

Por ejemplo, la estreptomina (SM) fue el primer fármaco con actividad antituberculosa demostrada, sin embargo, actualmente es considerado una opción de segunda línea utilizada comúnmente en el tratamiento de tuberculosis resistente.⁴⁷

De acuerdo a su mecanismo de acción, la estreptomina se une al 16S rRNA e interfiere con el ribosoma, inhibiendo la síntesis de proteínas en bacterias susceptibles. Mientras tanto, la resistencia a estreptomina se ha asociado a la alteración de la afinidad del 16S rRNA por el fármaco a través de la presencia de mutaciones en los genes *rrs*, que codifica para el 16S rRNA, y *rpsL* que codifica la proteína ribosomal S12 cuya función es la estabilización de la estructura de dicho rRNA.^{48,47}

La amikacina es un fármaco inyectable de segunda línea que se utiliza para el tratamiento de tuberculosis causada por organismos resistentes a los fármacos de primera línea.⁴⁶ Al igual que en el caso de estreptomina, la amikacina inhibe la síntesis proteica bacteriana mediante su unión al 16S rRNA y la resistencia a este antibiótico se ha asociado también a mutaciones en el gen *rrs* que codifica para él.⁴⁸

Las fluoroquinolonas también son fármacos de segunda línea que juegan un papel importante en el tratamiento de tuberculosis resistente a antibióticos debido a sus relativamente menores efectos adversos y a su alta actividad bactericida.⁴⁸ Todos los antibióticos de esta clase actúan inhibiendo la actividad de la DNA girasa, impidiendo de esta manera la replicación de los microorganismos, sin embargo, los que presentan mayor actividad contra *Mycobacterium tuberculosis* son levofloxacin, moxifloxacin y gatifloxacin.

De la misma manera en que estos fármacos comparten un mismo mecanismo de acción, el surgimiento de resistencia es también un efecto de clase y las mutaciones en los genes *gyrA* y *gyrB* que codifican, respectivamente, para las subunidades A y B de la DNA girasa otorgan resistencia a la categoría completa de fluoroquinolonas.⁴⁸

Debido a que la resistencia a fármacos en *Mycobacterium tuberculosis* no se encuentra asociada a elementos genéticos móviles, como sucede en otras especies bacterianas, los mecanismos mencionados en esta sección explican la abrumadora mayoría de las cepas resistentes a los principales antibióticos antituberculosos. Sin embargo, deben existir otros mecanismos por los cuales *Mycobacterium tuberculosis* puede adquirir resistencia a estos antibióticos, pero por el momento quedarán fuera del alcance de este trabajo ya que su relevancia es todavía incierta.

ANEXO 4. Protocolo detallado de la prueba genómica propuesta para detección de farmacorresistencia en *Mycobacterium tuberculosis*.

Extracción de ácidos nucleicos

La extracción del DNA genómico a partir de cultivos puros de *Mycobacterium tuberculosis*, esputo o lavado bronco-alveolar se lleva a cabo según los protocolos estandarizados del INER, utilizando el QIAamp DNA mini kit (QIAGEN).

Amplificación por PCR.

Para llevar a cabo la amplificación de cada una de las nueve regiones del genoma de *Mycobacterium tuberculosis* se preparan de manera independiente mezclas de reacción de 25 µL con las siguientes concentraciones de reactivos (tabla 8) y “primers” específicos correspondientes a cada región (la secuencia de los “primers” se presenta en el Anexo 2).

Tabla 8. Concentraciones de reactivos para amplificación por PCR.

Reactivo	Concentración final para PCR
Buffer	20 mM Tris-HCl [pH 8.4], 50 mM KCl
MgCl ₂	1.5 mM
dNTPs	0.2 mM
Polimerasa Taq	0.625 u (0.025 u/µL)
Forward primer correspondiente	0.5 µM
Reverse primer correspondiente	0.5 µM
ADN molde	>6 ng ADN

Después las mezclas de reacción se colocan en un termociclador (GeneAmp PCR System 9700 Applied Biosystems) y se lleva a cabo la amplificación con el siguiente protocolo de termociclado:

1. Desnaturalización inicial 95 °C, 15 minutos
2. 35 ciclos de:
 - Desnaturalización 95 °C por 30s
 - Alineamiento 60 °C por 30 s
 - Elongación 72 °C por 30 s
3. Elongación final 72 °C durante 7 minutos.

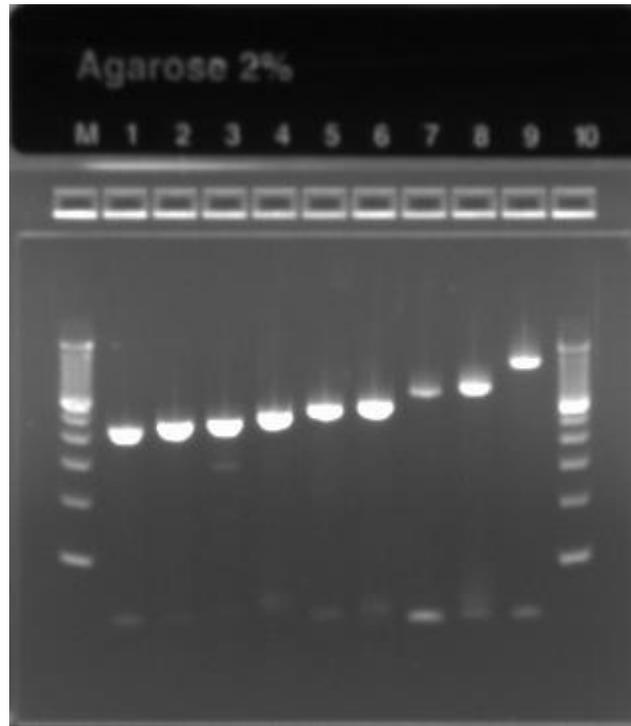
Para verificar la correcta amplificación de las reacciones es posible realizar una electroforesis en gel de agarosa al 2% (Figura 25).

En caso de utilizar el sistema E-gel EX o SizeSelect (Invitrogen) se pueden seguir los pasos descritos a continuación:

1. Conectar el transiluminador Safelmager a la unidad E-gel iBase Power System.
2. Colocar un cartucho E-gel EX en el sistema.

3. Cargar 5 μL de muestra y 15 μL de agua libre de nucleasas en cada pozo
4. Llenar con 20 μL de agua libre de nucleasas todos los pozos vacíos que queden
5. Realizar una electroforesis de 10 min.

Figura 25. Verificación por electroforesis de la amplificación de las nueve regiones asociadas a farmacorresistencia.



Carriles M y 10: marcador de peso molecular 100 pb; Carril 1: amplicón *gyrA* (423 pb); Carril 2: amplicón *katG* (455 pb); Carril 3: amplicón promotor de *mabA/inhA* (455 pb); Carril 4: amplicón *embB* (490 pb); Carril 5: amplicón *rpoB* (543 pb); Carril 6: amplicón *rpsL* (572 pb); Carril 7: amplicón *pncA* (689 pb); Carril 8: amplicón *gyrB* (772 pb); Carril 9: amplicón *rrs* (1329 pb).

Purificación de los productos de PCR

Para la purificación de los productos de PCR se recomienda utilizar el sistema AMPure XP listados a continuación:

Nota: Es necesario que las perlas magnéticas AMPure se encuentren a temperatura ambiente antes de iniciar, por lo que es necesario incubarlas previamente a temperatura ambiente durante aproximadamente 30 min. Además es necesario preparar por muestra un volumen de 400 μL de etanol al 80% que se utilizará en el paso 6.

1. Tomar 20 μL del producto de PCR.
2. Agregar 25 μL de perlas Agencourt AMPure XP.
3. Mezclar con pipeta 10 veces e incubar 15 min a temperatura ambiente.
4. Colocar en un soporte magnético por 5 min.
5. Sin quitar del soporte magnético remover el sobrenadante.

6. Sin quitar del soporte magnético lavar 2 veces con etanol al 80%:
 - a. Agregar 200 μL de etanol al 80%, incubar por 30 s a temperatura ambiente rotando dos veces el tubo en el soporte.
 - b. Aspirar y desechar el sobrenadante.
7. Dejar secar 10 min a temperatura ambiente.
8. Retirar del soporte magnético y eluir en 25 μL de agua (resuspender con vortex 10 s).
9. Colocar en el soporte magnético por 5 min para separar perlas.
10. Recuperar el sobrenadante transfiriéndolo a un nuevo tubo.

Cuantificación de los productos de PCR y preparación de una mezcla equimolar de amplicones

Es necesario hacer una mezcla equimolar de todos los amplicones antes de fragmentarlos enzimáticamente, por lo que resulta indispensable cuantificarlos previamente. Se recomienda realizar esta cuantificación utilizando el fluorómetro Qubit y un kit dsDNA High Sensitivity (Invitrogen), siguiendo los pasos descritos a continuación:

Nota: Es necesario que los reactivos se encuentren a temperatura ambiente antes de iniciar, por lo que previamente deben incubarse a temperatura ambiente durante aproximadamente 30 min. Además se debe tener cuidado de utilizar sólo tubos que no interfieran con la medición del fluorómetro para mezclar la solución de trabajo con las muestras y los estándares.

1. Preparar una mezcla de trabajo que contenga 199 μL de buffer y 1 μL de colorante Qubit reagent por cada una de las muestras y los dos estándares.
2. Tomar 190 μL de la mezcla de trabajo y agregar 10 μL del estándar 1.
3. Tomar 190 μL de la mezcla de trabajo y agregar 10 μL del estándar 2.
4. Tomar 199 μL de la mezcla de trabajo y agregar 1 μL de una de las muestras a analizar.
5. Repetir el paso 4 para cada una de las muestras.
6. Incubar por dos minutos a temperatura ambiente.
7. Tomar en el aparato la medición correspondiente al estándar 1, al estándar 2 y a cada una de las muestras.

Utilizando estos valores de concentración, se calculan los volúmenes necesarios para obtener una mezcla equimolar de todos los amplicones (conteniendo mínimo 0.0005 ngmol de cada amplicón) y se prepara dicha mezcla.

Fragmentación enzimática de la mezcla equimolar de amplicones

Debido a que para la reacción de fragmentación enzimática se necesita una cantidad de 100 ng totales de DNA es conveniente cuantificar la mezcla equimolar con Qubit para obtener la concentración total de DNA. Si se realiza esta cuantificación utilizando el fluorómetro Qubit y un kit dsDNA High Sensitivity (Invitrogen), se pueden seguir los pasos listados a continuación:

Nota: Es necesario que los reactivos se encuentren a temperatura ambiente antes de iniciar, por lo que previamente deben incubarse a temperatura ambiente durante aproximadamente 30 min. Además se debe tener cuidado de utilizar sólo tubos que no interfieran con la medición del fluorómetro para mezclar la solución de trabajo con las muestras y los estándares.

1. Preparar una mezcla de trabajo que contenga 199 μL de buffer y 1 μL de colorante Qubit reagent por cada una de las muestras y los dos estándares.
2. Tomar 190 μL de la mezcla de trabajo y agregar 10 μL del estándar 1.
3. Tomar 190 μL de la mezcla de trabajo y agregar 10 μL del estándar 2.
4. Tomar 199 μL de la mezcla de trabajo y agregar 1 μL de una de las muestras a analizar.
5. Repetir el paso 4 para cada una de las muestras.
6. Incubar por dos minutos a temperatura ambiente.
7. Tomar en el aparato la medición correspondiente al estándar 1, al estándar 2 y a cada una de las muestras.

Para realizar la fragmentación enzimática de la mezcla equimolar de amplicones se utilizan los reactivos Ion Shear Plus (Life Technologies), siguiendo los pasos descritos a continuación:

1. Utilizando la concentración obtenida previamente tomar un volumen de mezcla equimolar de amplicones correspondiente a 100 ng totales de DNA.
2. Agregar 5 μL de Ion Shear Plus 10X Reaction Buffer.
3. Añadir agua libre de nucleasas hasta alcanzar un volumen de 40 μL .
4. Añadir 10 μL de Ion Shear Plus Enzyme Mix y mezclar con la pipeta 10 veces.
5. Incubar la mezcla de reacción a 37 $^{\circ}\text{C}$ por 30 min.
6. Añadir 5 μL de Ion Shear Stop Buffer y mezclar con vortex por 5 s.
7. Mantener el tubo en hielo y proceder inmediatamente a la purificación del producto de fragmentación enzimática.

Purificación del producto de la fragmentación enzimática

La purificación se realiza inmediatamente después de la reacción de fragmentación enzimática siguiendo los pasos mostrados a continuación:

Nota: Es necesario que las perlas magnéticas AMPure se encuentren a temperatura ambiente antes de iniciar, por lo que es necesario incubarlas previamente a temperatura ambiente durante aproximadamente 30 min. Además es necesario preparar por muestra un volumen de 1 mL de etanol al 70% que se utilizará en el paso 5.

1. Agregar 99 μL de perlas Agencourt AMPure XP (1.8 veces el volumen de reacción de la fragmentación enzimática [$1.8 \times 55 \mu\text{L} = 99 \mu\text{L}$]).
2. Mezclar con pipeta 10 veces e incubar 5 min a temperatura ambiente.
3. Colocar en un soporte magnético por 5 min.
4. Sin quitar del soporte magnético remover el sobrenadante.
5. Sin quitar del soporte magnético lavar 2 veces con etanol al 70%:
 - a. Agregar 500 μL de etanol al 70%, dar dos vueltas al tubo en el soporte e incubar por 30 s a temperatura ambiente.
 - b. Aspirar y desechar el sobrenadante.
6. Dejar secar 10 min a temperatura ambiente.
7. Retirar del soporte magnético y eluir en 25 μL de Low TE buffer (resuspender con vortex 10 s).
8. Colocar en placa magnética por 1 min para separar perlas.
9. Transferir el sobrenadante a un nuevo tubo.

Ligación de adaptadores a los productos de fragmentación

En un tubo de PCR de 0.2 mL mezclar los reactivos listados a continuación para completar un volumen total de 100 μ L

Tabla 9. Mezcla de reacción para la ligación de adaptadores.

Reactivo	Cantidad a agregar
DNA fragmentado enzimáticamente	25 μ L
10X Ligase Buffer	10 μ L
Adaptador P1	2 μ L
Adaptador Xpress Barcode X	2 μ L
Mezcla de dNTPs	2 μ L
Agua libre de nucleasas	49 μ L
DNA ligasa	2 μ L
Polimerasa Nick Repair	8 μ L

1. Colocar la mezcla de reacción en un termociclador e incubar 15 minutos a 25°C y después 5 minutos a 72°C.
2. Transferir todo el contenido a un tubo LoBind de 1.5 mL y pasar inmediatamente a la purificación del producto de ligación.

Purificación del producto de ligación

La purificación se realiza inmediatamente después de la reacción de ligación de adaptadores, siguiendo los pasos mostrados a continuación:

Nota: Es necesario que las perlas magnéticas AMPure se encuentren a temperatura ambiente antes de iniciar, por lo que es necesario incubarlas previamente a temperatura ambiente durante aproximadamente 30 min. Además es necesario preparar por muestra un volumen de 1 mL de etanol al 70% que se utilizará en el paso 5.

1. Para una biblioteca de lecturas de 200 pb se deben añadir 140 μ L de perlas Agencourt AMPure XP (1.4 veces el volumen de reacción de la fragmentación enzimática [1.4 x 100 μ L = 140 μ L]).
2. Mezclar con pipeta 5 veces e incubar 5 min a temperatura ambiente.
3. Colocar en un soporte magnético por 3 min.
4. Sin quitar del soporte magnético remover el sobrenadante.
5. Sin quitar del soporte magnético lavar 2 veces con etanol al 70%:
 - a. Agregar 500 μ L de etanol al 70%, dar dos vueltas al tubo en el soporte e incubar por 30 s a temperatura ambiente.

- b. Aspirar y desechar el sobrenadante.
6. Dejar secar 5 min a temperatura ambiente.
7. Retirar del soporte magnético y eluir en 20 μL de Low TE buffer (resuspender con vortex 10 s).
8. Colocar en placa magnética por 1 min para separar perlas.
9. Transferir el sobrenadante a un nuevo tubo LoBind de 1.5 mL.

Selección de tamaño de la biblioteca

Para seleccionar por tamaño la biblioteca es posible realizar una electroforesis en gel de agarosa al 2%.

En caso de utilizar el sistema E-gel SizeSelect (Invitrogen) se pueden seguir los pasos descritos a continuación:

1. Conectar el transiluminador Safelmager a la unidad E-gel iBase Power System.
2. Colocar un cartucho E-gel Size Select en el sistema.
3. Cargar 20 μL de muestra en cada pozo, evitando utilizar los pozos de los extremos
4. Cargar 10 μL de marcador de peso molecular de 50 pb diluido en buffer Low TE a una concentración de 25 ng/ μL (dilución 1:40)
5. Llenar con 20 μL de agua libre de nucleasas todos los pozos grandes que queden vacíos y con 10 μL el pozo pequeño vacío.
6. Realizar una electroforesis hasta que los fragmentos de 330 pb alcancen la línea de referencia (aproximadamente 13 min).
7. Detener la electroforesis una vez que los fragmentos de 330 pb hayan llegado a la línea de referencia.
8. Rellenar los pozos de recolección con agua libre de nucleasas.
9. Reanudar la electroforesis hasta que la banda de 330 pb haya llegado al pozo de recolección (aproximadamente 1 minuto).
10. Detener la electroforesis y recolectar la muestra de los pozos de recolección:
 - a. Recolectar aproximadamente 20 μL de muestra del pozo utilizando una pipeta.
 - b. Añadir 10 μL de agua libre de nucleasas al pozo y recolectarlos nuevamente.

Amplificación de la biblioteca.

En un tubo de PCR de 0.2 mL mezclar los reactivos listados a continuación para completar un volumen total de 130 μL

Tabla 10. Mezcla de reacción para la amplificación de la biblioteca.

Reactivo	Cantidad a agregar
Platinum PCR superMix High Fidelity	100 μL
Library Amplification Primer Mix	5 μL
Biblioteca sin amplificar	25 μL

Después las mezclas de reacción se colocan en un termociclador (GeneAmp PCR System 9700 Applied Biosystems) y se lleva a cabo la amplificación con el siguiente protocolo de termociclado:

1. Desnaturalización inicial 95 °C, 5 minutos
2. 8 ciclos de:
 - Desnaturalización 95 °C por 15 s
 - Alineamiento 58 °C por 15 s
 - Elongación 70 °C por 1 min

Purificación de la biblioteca amplificada

La purificación se realiza inmediatamente después de la reacción de amplificación, siguiendo los pasos mostrados a continuación:

Nota: Es necesario que las perlas magnéticas AMPure se encuentren a temperatura ambiente antes de iniciar, por lo que es necesario incubarlas previamente a temperatura ambiente durante aproximadamente 30 min. Además es necesario preparar por muestra un volumen de 1 mL de etanol al 70% que se utilizará en el paso 5.

1. Añadir 195 µL de perlas Agencourt AMPure XP (1.5 veces el volumen de reacción de la fragmentación enzimática [1.5 x 130 µL = 195 µL]).
2. Mezclar con pipeta 5 veces e incubar 5 min a temperatura ambiente.
3. Colocar en un soporte magnético por 3 min.
4. Sin quitar del soporte magnético remover el sobrenadante.
5. Sin quitar del soporte magnético lavar 2 veces con etanol al 70%:
 - a. Agregar 500 µL de etanol al 70%, dar dos vueltas al tubo en el soporte e incubar por 30 s a temperatura ambiente.
 - b. Aspirar y desechar el sobrenadante.
6. Dejar secar 5 min a temperatura ambiente.
7. Retirar del soporte magnético y eluir en 20 µL de Low TE buffer (resuspender con vortex 10 s).
8. Colocar en placa magnética por 1 min para separar perlas.
9. Transferir el sobrenadante a un nuevo tubo LoBind de 1.5 mL.

Cuantificación de la biblioteca amplificada mediante qPCR

Nota: para la cuantificación de la biblioteca por qPCR es necesario contar con una curva estándar de 5 puntos de una biblioteca control.

De manera previa a su cuantificación deben prepararse diluciones 1:2,000 y 1:20,000 de la biblioteca, siguiendo los pasos descritos a continuación:

1. Preparar una dilución inicial de la biblioteca 1:20.
2. Tomar 1 µL de la dilución inicial 1:20 y añadir 99 µL de agua para preparar la dilución 1:2,000.
3. Tomar 5 µL de la dilución 1:2,000 recién preparada y añadir 45 µL de agua para obtener la dilución 1:20,000.

Como siguiente paso se preparan mezclas de reacción para cada una de las diluciones agregando las siguientes cantidades de reactivos:

Tabla 11. Mezcla de reacción para la cuantificación por qPCR de la biblioteca.

Reactivo	Cantidad a agregar
Ion Library TaqMan qPCR Mix, 2x	10 µL
Ion Library TaqMan Quantitation assay, 20x	1 µL
Agua libre de nucleasas	4 µL
Dilución de la biblioteca a cuantificar	5 µL

Las reacciones preparadas se corren en un termociclador de PCR en tiempo real según las instrucciones del fabricante y se utiliza la curva estándar para calcular la concentración de las diluciones 1:2,000 o 1:20,000 de la biblioteca.

Como siguiente paso se calcula el parámetro “Template Dilution Factor” utilizando la siguiente fórmula:

$$\text{Template Dilution Factor} = \frac{(\text{Cantidad obtenida por qPCR}) * (\text{Factor de dilución de la biblioteca})}{0.32}$$

Una vez se ha calculado el valor de este parámetro se procede a diluir la muestra a un factor 1:(Template Dilution Factor) utilizando el volumen adecuado de buffer Low TE.

Preparación de una mezcla equimolar de bibliotecas

Para obtener una mezcla equimolar de las bibliotecas que se desea secuenciar simultáneamente, se toman en un mismo tubo 5 µL de cada biblioteca previamente diluida de acuerdo a su Template Dilution Factor.

Preparación de templado (Amplificación clonal)

A continuación se describe el procedimiento para llevar a cabo la amplificación clonal del DNA en las Ion Spheres utilizando el sistema automatizado Ion OneTouch, aunque este paso puede realizarse de forma manual siguiendo las instrucciones del fabricante.

Como primer paso se prepara el equipo Ion OneTouch.

1. Colocar una placa nueva en el puerto superior.
2. Colocar “zipers” nuevos en las entradas para aceite y “recovery solution”.
3. Llenar el tubo para aceite hasta la mitad con aceite OneTouch Oil.
 - a. Invertir 10 veces la botella de aceite
 - b. Llenar el tubo OneTouch reagent tube hasta la mitad.
4. Llenar el tubo para “recovery solution” hasta la mitad con solución Ion Onetouch Recovery.

5. Colocar tubos para recuperación de la muestra (Recovery Tubes) y el adaptador OneTouch Recovery Router en la centrífuga.

Después se prepara una mezcla de reacción de amplificación clonal con los pasos siguientes:

1. En tubo LoBind de 1.5 mL agregar:
 - a. 595 μ L de agua libre de nucleasas
 - b. 200 μ L de ion OneTouch Reagent Mix
 - c. 100 μ L de Ion OneTouch Enzyme mix
 - d. 5 μ L de mezcla equimolar de bibliotecas
2. Resuspender las Ion Sphere Particles (ISPs) con vortex durante 1 min.
3. Agregar 100 μ L de Ion OneTouch 200 Ion Sphere particles y mezclar con pipeta.
4. Transferir la solución al Reaction Tube
5. Depositar 1.5 -2 mL de OneTouch Reaction Oil sobre la superficie de la solución
6. Tapar y asegurar el tubo de reacción con un OneTouch Reaction Filter.
7. Invertir lentamente sin mezclar e insertar en los puertos adecuados del ion OneTouch Instrument.
8. Iniciar la corrida del instrumento (que tiene una duración aproximada de 4 horas).

Recuperación de perlas Ion Spheres (ISPs)

1. Remover los tubos de recuperación de muestra del instrumento
2. Lavar las perlas:
 - a. Remover la solución de los tubos de recuperación dejando 50 μ L.
 - b. Resuspender con pipeta y transferir a un nuevo tubo, reuniendo en un tubo LoBind de 1.5 mL.
 - c. Añadir 1 mL de Ion OneTouch Wash Solution
 - d. Centrifugar por 2.5 min a 15,500 x g
 - e. Remover el sobrenadante excepto 100 μ L
 - f. Resuspender con vortex 10 s.
3. Guardar 2 μ L para cuantificación con Qubit.

Cuantificación de perlas Ion Spheres (ISPs)

1. Tomar 2 μ L de perlas previamente lavadas.
2. Agregar 19 μ L de solución Annealing buffer.
3. Añadir 1 μ L de sondas Ion Probes.
4. Mezclar con pipeta.
5. Colocar la mezcla de reacción en un termociclador y llevar a cabo el siguiente protocolo para el alineamiento de las sondas
 - a. 2 min a 95^o C
 - b. 2 min a 37^o C
6. Después del termociclado lavar 3 veces:
 - a. Agregar 200 μ L de solución Quality Control Wash Buffer, mezclar con vortex 5 s
 - b. Centrifugar 1.5 min a 15,500 rcf
 - c. Remover el sobrenadante excepto por 10 μ L

7. Cuantificar con Qubit
8. Los valores arrojados por el tubo se introducen en una hoja de cálculo de Microsoft Excel que determina la proporción de perlas “template-positive”.

Si la proporción de perlas positivas se encuentra entre 10 y 25% se continúa al enriquecimiento de perlas positivas en el Enricher System, en caso contrario, se diluye la muestra hasta obtener un valor en dicho intervalo.

Enriquecimiento de perlas Ion Spheres (ISPs) con templado.

En este paso se enriquece la solución en perlas “template-positive” y se desnaturaliza la hebra complementaria del ADN unido a la perla.

1. Colocar los siguientes reactivos en una tira de tubos para Enricher System:
 - a. En el tubo 1 colocar 100 μ L de perlas ISPs
 - b. En el tubo 2 colocar 130 μ L de perlas magnéticas MyOne Beads, previamente lavadas.
 - c. En cada uno de los tubos 3, 4 y 5 colocar 300 μ L de solución OneTouch Wash Solution
 - d. En el tubo 7 colocar 300 μ L de solución Melt-off Solution.
2. Activar el equipo para llevar a cabo el enriquecimiento.

Lavado de perlas ISPs enriquecidas.

1. Centrifugar 1.5 min a 15,500 rcf
2. Remover el sobrenadante a excepción de 10 μ L
3. Agregar 200 μ L de solución Ion One Touch Wash Solution.
4. Resuspender con pipeta.
5. Centrifugar 1.5 min a 15,500 rcf.
6. Remover el sobrenadante a excepción de 10 μ L.
7. Agregar Ion One touch Wash Solution hasta alcanzar un volume de 100 μ L.
8. Resuspender con pipeta.

Se guardan 10 μ L de estas perlas para analizar con Qubit y obtener la proporción de ISPs “template-positive” después del enriquecimiento, que debe encontrarse al menos a un 50%, de lo contrario debe repetirse el enriquecimiento.

Inicializar el equipo Ion Torrent PGM

1. Se preparan las soluciones W1 W2 y W3 según las instrucciones del fabricante.
2. Se lava el equipo.
3. Se lava el Chip de secuenciación.
4. Se colocan las soluciones y los nucleótidos en los puertos correspondientes.
5. Se verifica el chip de secuenciación en el equipo

Secuenciación con Ion Torrent PGM.

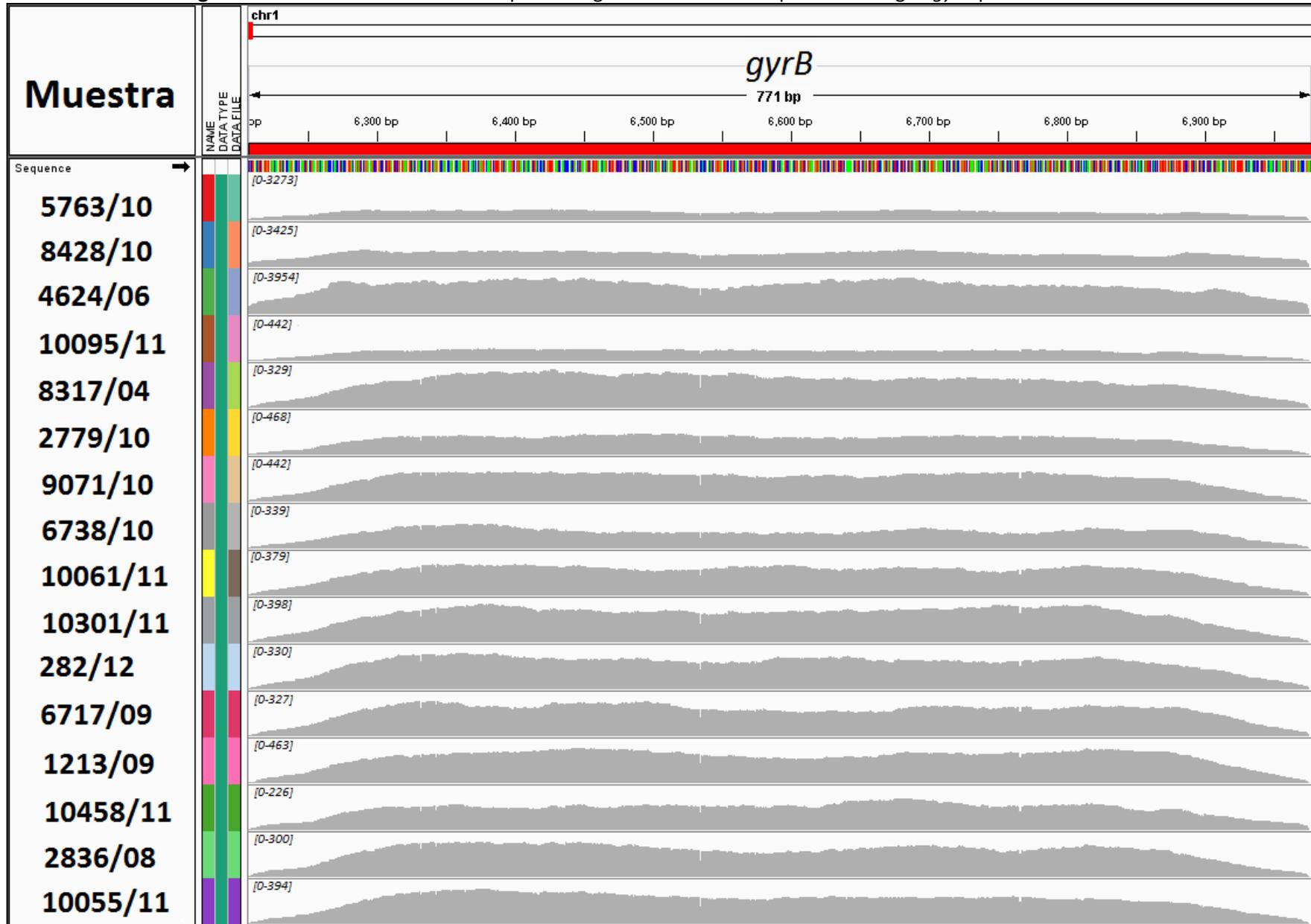
1. Tomar 45 μ L de esferas ISPs enriquecidas.
2. Añadir 5 μ L de esferas control
3. Añadir 150 μ L de Annealing Buffer
4. Centrifugar 2 min a 15,500 rcf
5. Retirar el sobrenadante a excepción de 9 μ L
6. Añadir 5 μ L de Sequencing Primer
7. Colocar en un termociclador y llevar a cabo el siguiente protocolo:
 - a. 2 min a 95^o C
 - b. 2 min a 37^o C
8. Después del termociclado añadir 1 μ L de polimerasa
9. Incubar 5 minutos a temperatura ambiente
10. Cargar el chip con la mezcla de reacción
11. Centrifugar el chip 30 segundos en posición "Tab in" y 30 segundos en posición "Tab out"
12. Iniciar la corrida del PGM y correr el plug-in "Variant Caller" de forma automática o al finalizar la corrida.
13. Al finalizar la corrida descargar los archivos .fastq, .bam, .bai y .vcf para cada una de las muestras.

Análisis de variantes.

1. Realizar el control de calidad de la corrida de secuenciación:
 - a. Analizar los archivos FASTQ generados en cuanto a tamaño y calidad de lecturas con el software FASTQC.
 - b. Visualizar los archivos BAM y BAI en el Integrative Genomics Viewer, analizar cobertura de las regiones y curar manualmente el llamado de variantes presentado en el archivo VCF.
2. Utilizar la hoja de cálculo de Microsoft Excel diseñada para filtrar por posición y frecuencia de variante y resumir las mutaciones encontradas en el archivo vcf
3. Inferir la resistencia a fármacos a partir de la presencia de mutaciones encontradas en muestras previas, buscar en fuentes de información selectas^{49,50} para asociar, si es el caso, la resistencia ante un fármaco particular a las mutaciones encontradas por primera vez en las muestras analizadas y predecir el efecto funcional de las mutaciones⁵¹ para proponer nuevas asociaciones si éstas no han sido reportadas previamente.
4. Generar un reporte de las resistencias inferidas para cada una de las muestras.
5. Por último, mantener un registro de las mutaciones encontradas y de su posible asociación a resistencia para utilizarlo en el análisis de muestras subsecuentes.

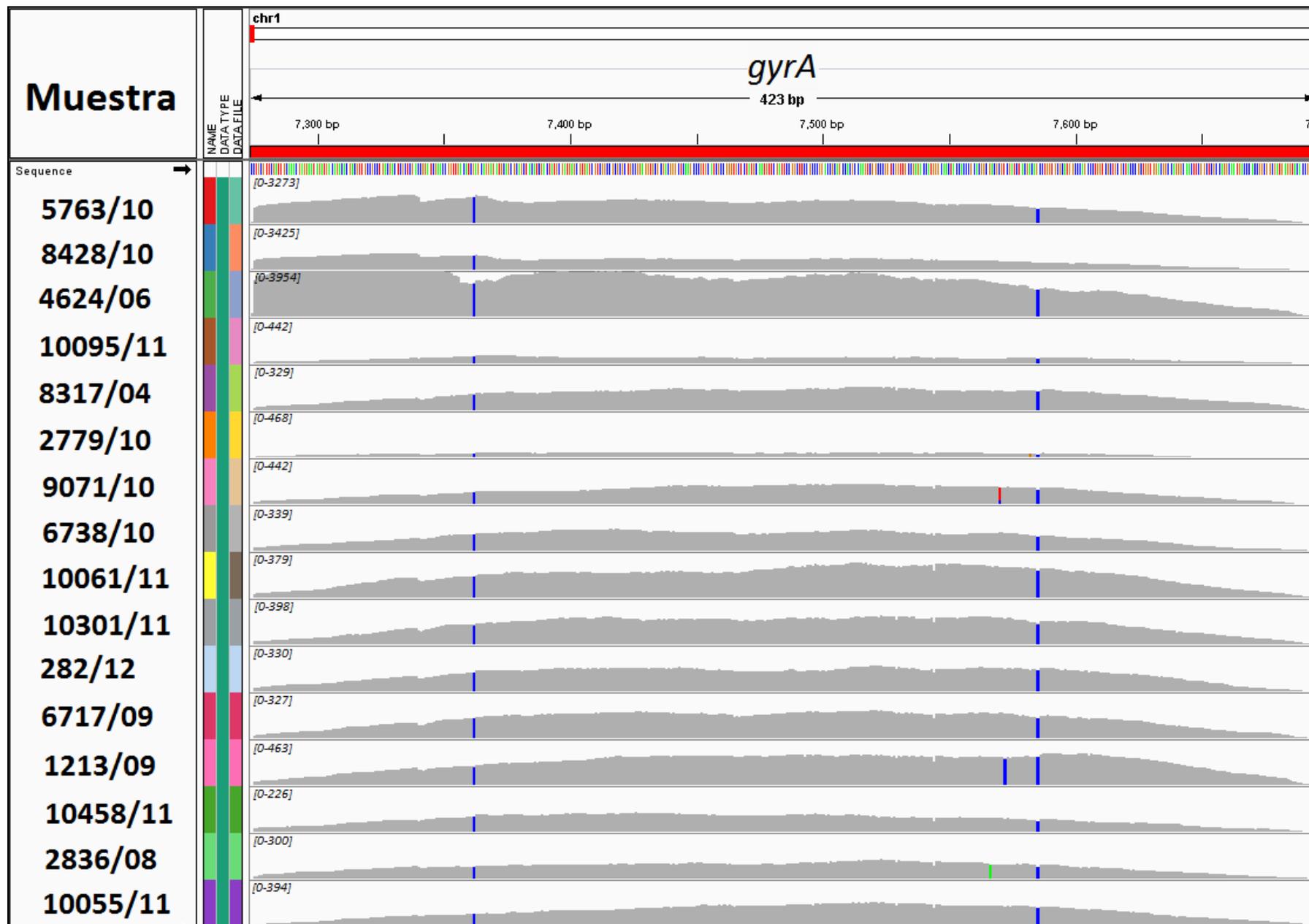
ANEXO 5. Profundidad de cobertura para las regiones del genoma de *Mycobacterium tuberculosis* analizadas.

Figura 26. Profundidad de cobertura para la región analizada correspondiente al gen *gyrB* para cada muestra.



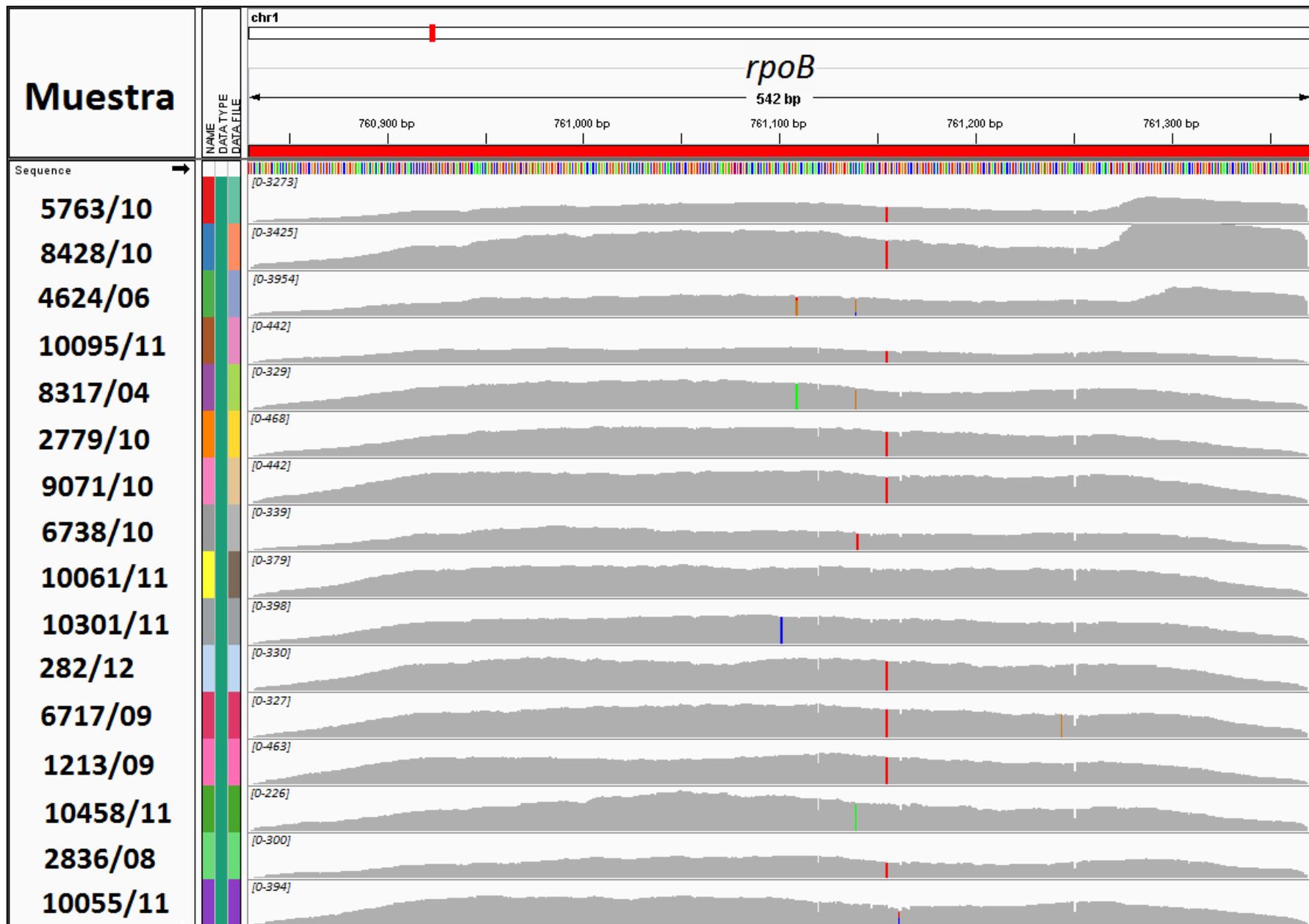
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 27. Profundidad de cobertura para la región analizada correspondiente al gen *gyrA* para cada muestra.



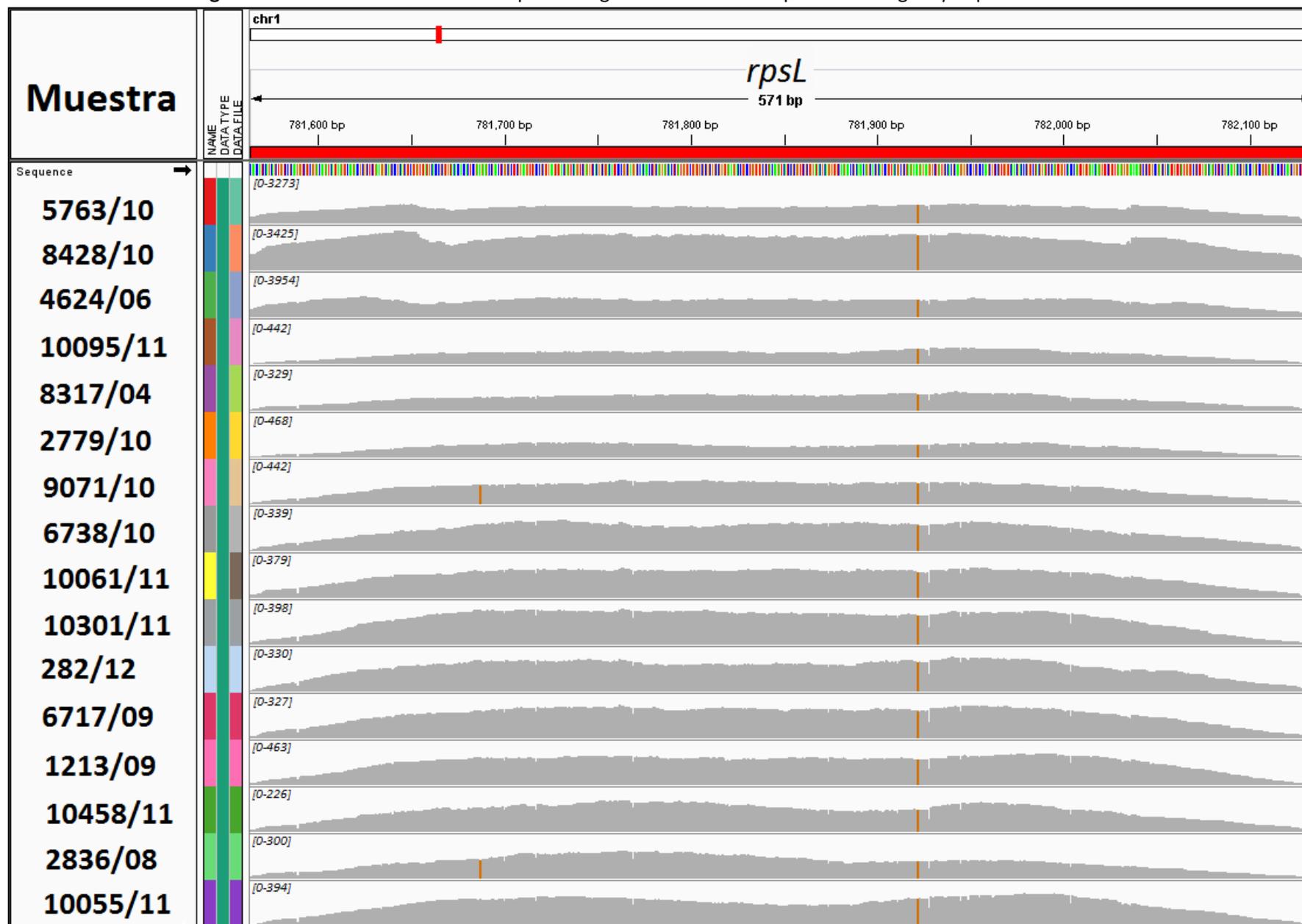
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 28. Profundidad de cobertura para la región analizada correspondiente al gen *rpoB* para cada muestra.



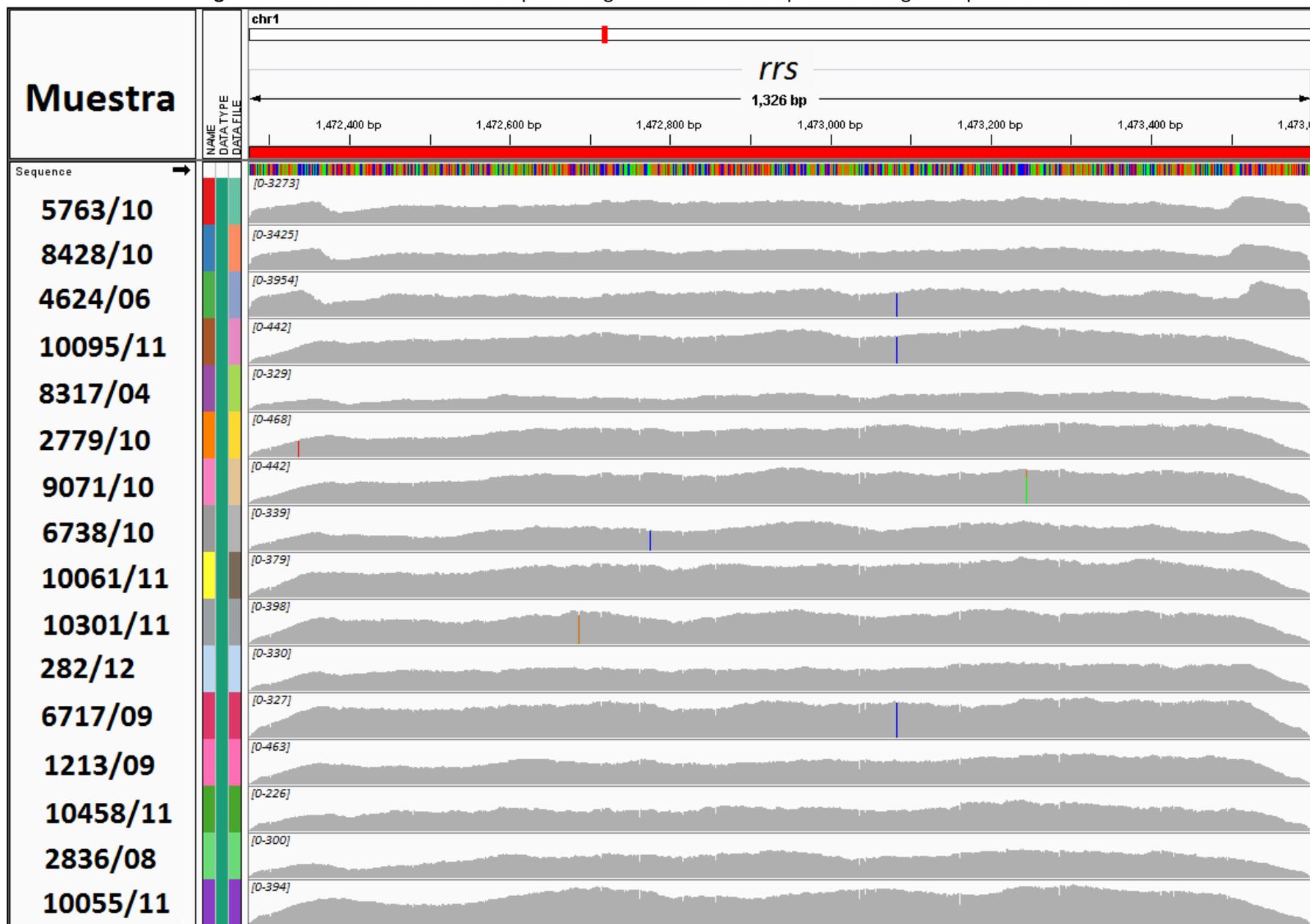
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 29. Profundidad de cobertura para la región analizada correspondiente al gen *rpsL* para cada muestra.



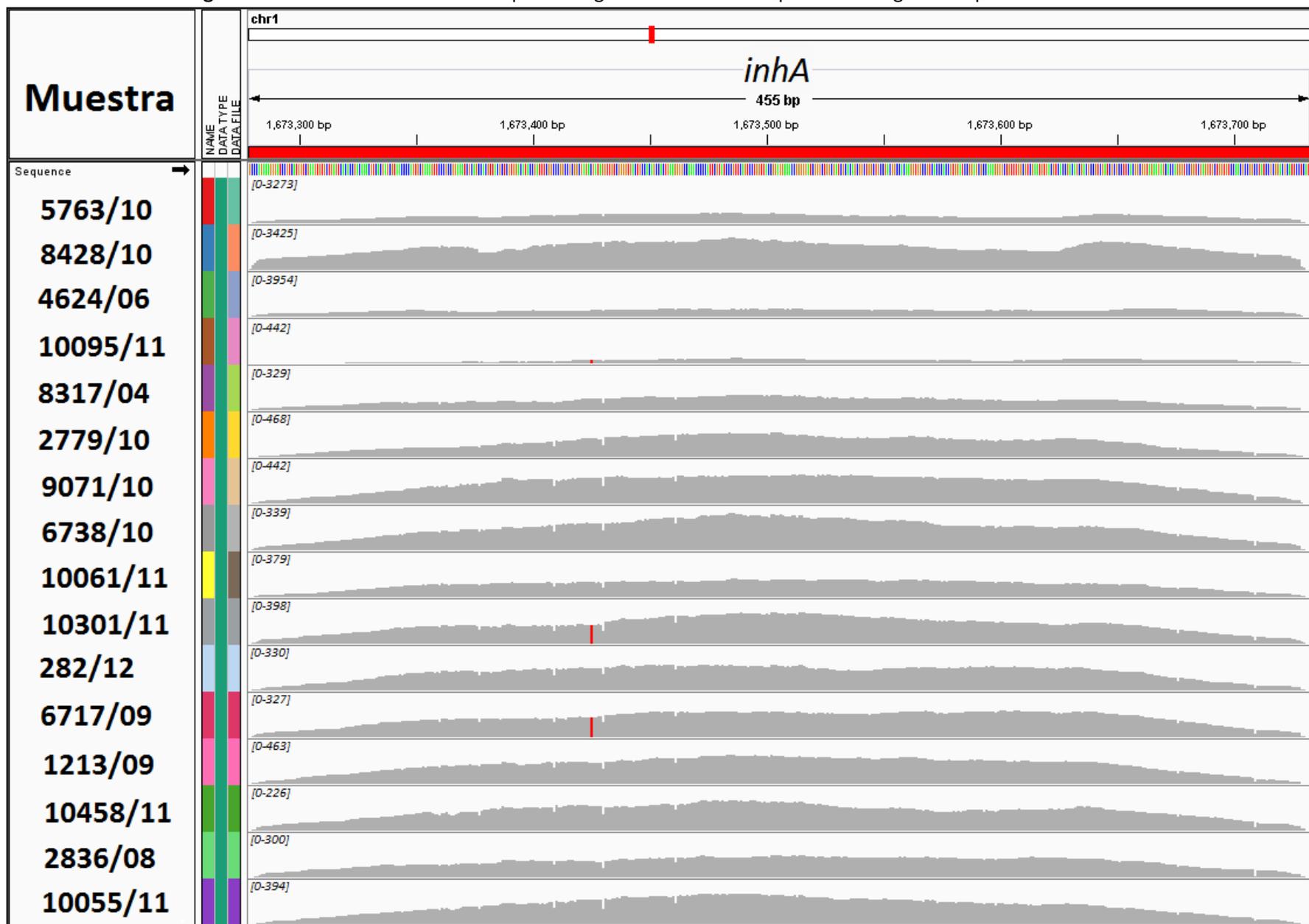
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 30. Profundidad de cobertura para la región analizada correspondiente al gen *rrs* para cada muestra.



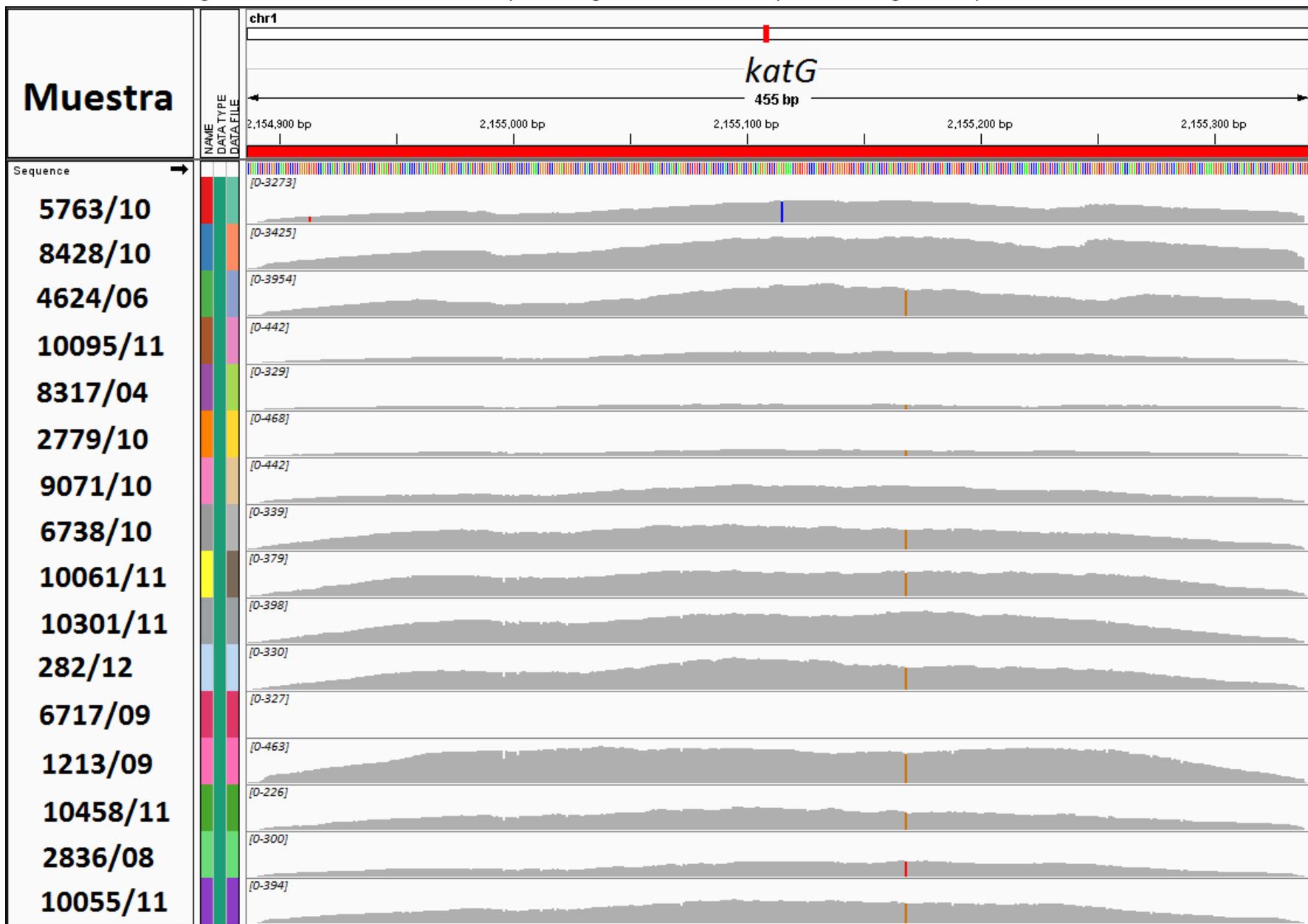
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 31. Profundidad de cobertura para la región analizada correspondiente al gen *inhA* para cada muestra.



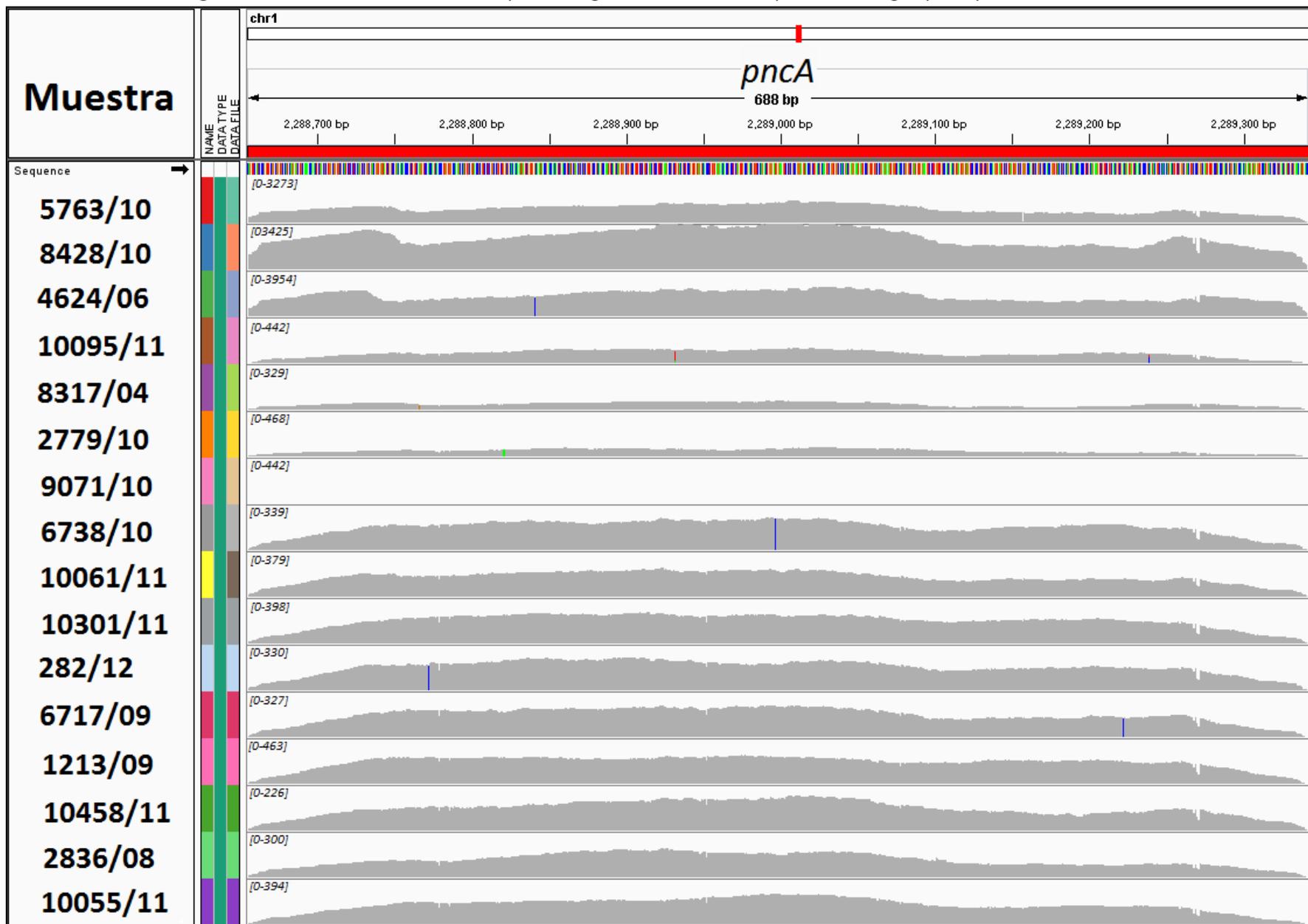
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 32. Profundidad de cobertura para la región analizada correspondiente al gen *katG* para cada muestra.



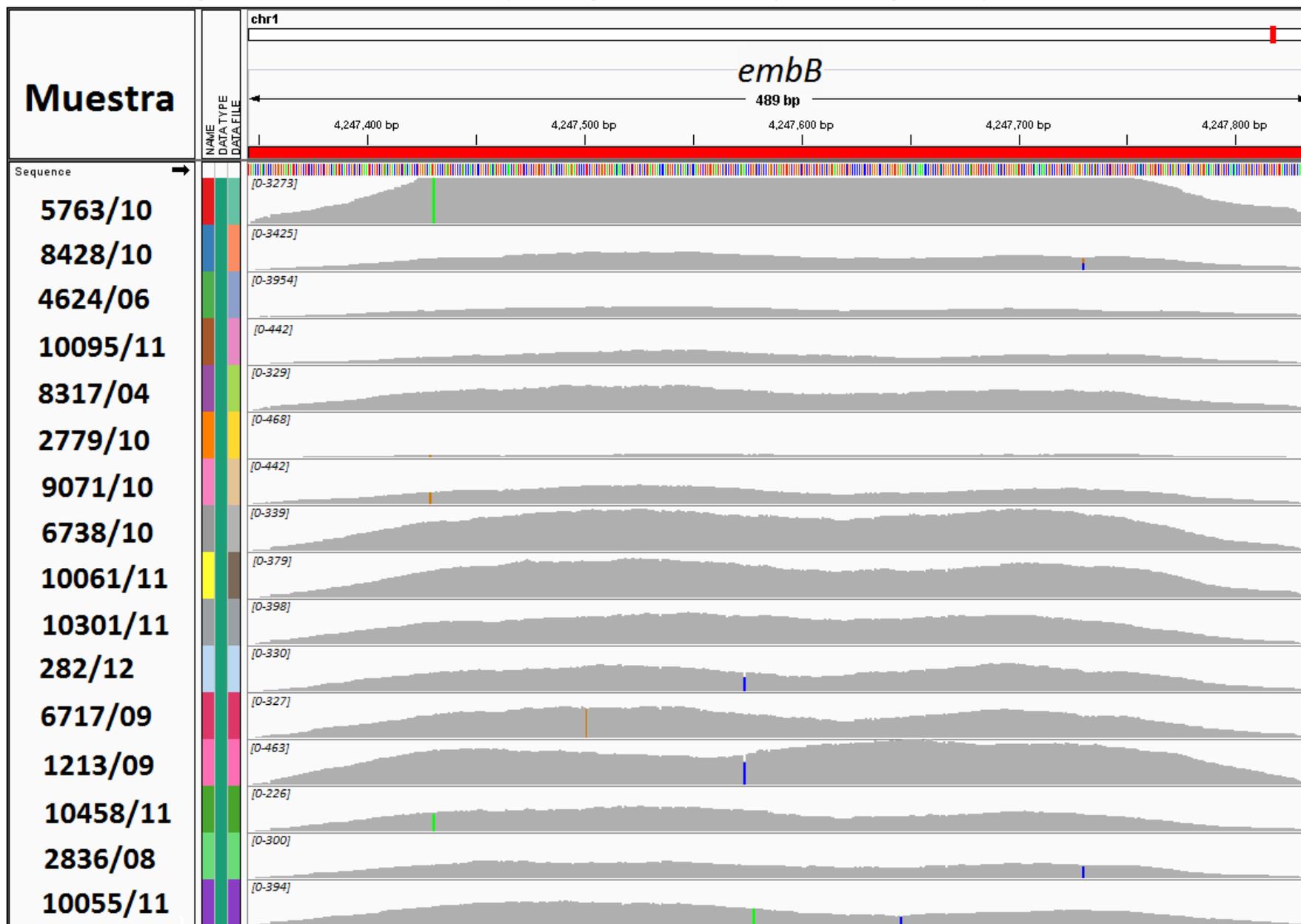
La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 33. Profundidad de cobertura para la región analizada correspondiente al gen *pncA* para cada muestra.



La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

Figura 34. Profundidad de cobertura para la región analizada correspondiente al gen *embB* para cada muestra.



La altura de las barras grises representa el número de lecturas que cubren cada posición, siendo el alto de cada fila la profundidad de cobertura máxima obtenida para cualquier región de una misma muestra. Además se marcan con color las variantes encontradas (rojo: T, azul: C, verde: A, café: G), definidas como aquellas posiciones donde el nucleótido difiere de la referencia en más de un 20% de las lecturas, ponderadas por calidad, que abarcan esa posición.

ANEXO 6. Mutaciones detectadas en las 16 muestras de *Mycobacterium tuberculosis* analizadas con la prueba propuesta.

Tabla 12. Mutaciones detectadas en las muestras de *Mycobacterium tuberculosis*, su clasificación y asociación a resistencia ante fármacos.

Muestra	Gen (producto)	Nucleótido	Mutación	Codón	Cambio de aminoácido	Clasificación	Antibiótico	Resistencia
8428/10	<i>embB</i> (arabinosil transferasa)	1217	<i>embB</i> 1217 G>C	406	Gly>Ala	Predicha	Etambutol	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Ser>Leu	Conocida	Estreptomicina	susceptible
5763/10	<i>embB</i> (arabinosil transferasa)	918	<i>embB</i> 918 G>A	306	Met/Ile	Conocida	Etambutol	EMB
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	997	<i>katG</i> 997 G>C	333	Leu>Val	Predicha	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	86	<i>pncA</i> 86 CGG>CG	29	frameshift (codón 117>STOP)	Predicha	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
4624/06	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	401	<i>pncA</i> 401 G>C	134	Ala>Gly	Conocida	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1303	<i>rpoB</i> 1303 G>T*	435	Asp>Tyr	Conocida	Rifampicina	No concluyente
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1333	<i>rpoB</i> 1333 C>G	445	His>Asp	Conocida	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
	<i>rrs</i> (16S rRNA)	1239	<i>rrs</i> 1239 T>C	413	N/A	Conocida	Amikacina/Estreptomicina	SM
10095/11	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>mabA-inhA</i> (enoil- ACP reductasa)	-14	<i>mabA-inhA</i> -14 C>T	-5	N/A	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	3	<i>pncA</i> 3 C>T	1	Met>Ile	Predicha	Pyrazinamida	PZA
	<i>pncA</i> (pirazinamidasa)	310	<i>pncA</i> 310 T>A*	104	Ser>Cys	Predicha	Pyrazinamida	No concluyente
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
		<i>rrs</i> (16S rRNA)	1239	<i>rrs</i> 1239 T>C	413	N/A	Conocida	Amikacina/Estreptomicina
8317/04	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	476	<i>pncA</i> 476 A>G	159	Leu>Pro	Predicha	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1303	<i>rpoB</i> 1303 G>A	435	Asp>Asn	Conocida	Rifampicina	RIF
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1333	<i>rpoB</i> 1333 C>G	445	His>Asp	Conocida	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible

*Variantes inciertas que no pudieron clasificarse en verdaderas mutaciones o errores de secuenciación

FLQ: fluoroquinolonas, RIF: rifampicina, SM: estreptomicina, INH: isoniazida, PZA: pirazinamida, EMB: etambutol y AMK: amikacina.

Tabla 12. Mutaciones detectadas en las muestras de *Mycobacterium tuberculosis*, su clasificación y asociación a resistencia ante fármacos (continuación).

Muestra	Gen (producto)	Nucleótido	Mutación	Codón	Cambio de aminoácido	Clasificación	Antibiótico	Resistencia
2779/10	<i>embB</i> (arabinosil transferasa)	916	<i>embB</i> 916 A>G	306	Met>Val	Conocida	Etambutol	EMB
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	281	<i>gyrA</i> 281 A>G	94	Asp>Gly	Conocida	Quinolonas	FLQ
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	421	<i>pncA</i> 421 G>A	141	Gln>STOP	Conocida	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
	<i>rrs</i> (16S rRNA)	492	<i>rrs</i> 492 C>T	164	N/A	Conocida	Amikacina/Estreptomicina	SM
9071/10	<i>embB</i> (arabinosil transferasa)	916	<i>embB</i> 916 A>G	306	Met>Val	Conocida	Etambutol	EMB
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	269	<i>gyrA</i> 269 C>T	90	Ala>Val	Conocida	Quinolonas	FLQ
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>pncA</i> (pirazinamidasa)		sin lecturas				Pyrazinamida	No concluyente
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	128	<i>rpsL</i> 128 A>G	43	Lys>Arg	Conocida	Estreptomicina	SM
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
	<i>rrs</i> (16S rRNA)	1401	<i>rrs</i> 1401 A>G	467	N/A	Conocida	Amikacina/Estreptomicina	AMK
6738/10	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	245	<i>pncA</i> 245 T>C	82	His>Arg	Conocida	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1334	<i>rpoB</i> 1334 A>T	445	His>Leu	Conocida	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
		<i>rrs</i> (16S rRNA)	931	<i>rrs</i> 931 A>C	311	N/A	Desconocida	Amikacina/Estreptomicina
10061/11	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
10301/11	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>mabA-inhA</i> (enoil- ACP reductasa)	-14	<i>mabA-inhA</i> -14 C>T	N/A	N/A	Conocida	Isoniazida	INH
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1295	<i>rpoB</i> 1295 A>C	432	Gln>Pro	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
	<i>rrs</i> (16S rRNA)	842	<i>rrs</i> 842 A>G	281	N/A	Desconocida	Amikacina/Estreptomicina	No concluyente

*Variantes inciertas que no pudieron clasificarse en verdaderas mutaciones o errores de secuenciación

FLQ: fluoroquinolonas, RIF: rifampicina, SM: estreptomicina, INH: isoniazida, PZA: pirazinamida, EMB: etambutol y AMK: amikacina.

Tabla 12. Mutaciones detectadas en las muestras de *Mycobacterium tuberculosis*, su clasificación y asociación a resistencia ante fármacos (continuación).

Muestra	Gen (producto)	Nucleótido	Mutación	Codón	Cambio de aminoácido	Clasificación	Antibiótico	Resistencia
282/12	<i>embB</i> (arabinosil transferasa)	1061	<i>embB</i> 1061 A>C	354	Asp>Ala	Conocida	Etambutol	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	470	<i>pncA</i> 470 A>C	157	Val>Gly	Predicha	Pyrazinamida	susceptible
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
6717/09	<i>embB</i> (arabinosil transferasa)	988	<i>embB</i> 988 T>G	330	Phe>Val	Conocida	Etambutol	EMB
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)		sin lecturas				Isoniazida	No concluyente
	<i>mabA-inhA</i> (enoil- ACP reductasa)	-14	<i>mabA-inhA</i> -14 C>T	N/A	N/A	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	20	<i>pncA</i> 20 A>C	7	Val>Gly	Conocida	Pyrazinamida	PZA
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1438	<i>rpoB</i> 1438 A>G	480	Ile>Val	Conocida	Rifampicina	RIF
<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible	
<i>rrs</i> (16S rRNA)	1239	<i>rrs</i> 1239 T>C	413	N/A	Conocida	Amikacina/Estreptomicina	SM	
1213/09	<i>embB</i> (arabinosil transferasa)	1061	<i>embB</i> 1061 A>C	354	Asp>Ala	Conocida	Etambutol	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	271	<i>gyrA</i> 271 T>C	91	Ser>Pro	Conocida	Quinolonas	FLQ
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
10458/11	<i>embB</i> (arabinosil transferasa)	918	<i>embB</i> 918 G>A	306	Met>Ile	Conocida	Etambutol	EMB
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1333	<i>rpoB</i> 1333 C>A	445	His>Gln	Conocida	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible

*Variantes inciertas que no pudieron clasificarse en verdaderas mutaciones o errores de secuenciación

FLQ: fluoroquinolonas, RIF: rifampicina, SM: estreptomicina, INH: isoniazida, PZA: pirazinamida, EMB: etambutol y AMK: amikacina.

Tabla 12. Mutaciones detectadas en las muestras de *Mycobacterium tuberculosis*, su clasificación y asociación a resistencia ante fármacos (continuación).

Muestra	Gen (producto)	Nucleótido	Mutación	Codón	Cambio de aminoácido	Clasificación	Antibiótico	Resistencia
2836/08	<i>embB</i> (arabinosil transferasa)	1217	<i>embB</i> 1217 G>C	406	Gly>Ala	Predicha	Etambutol	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	265	<i>gyrA</i> 265 G>A	89	Asp>Asn	Conocida	Quinolonas	FLQ
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>T	315	Ser>Asn	Conocida	Isoniazida	INH
	<i>pncA</i> (pirazinamidasa)	142	<i>pncA</i> 142 T>TGGTTGCC*	48	frameshift (codón 58>STOP)	Predicha	Pyrazinamida	No concluyente
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1349	<i>rpoB</i> 1349 C>T	450	Ser>Leu	Predicha	Rifampicina	RIF
	<i>rpsL</i> (proteína ribosomal S12)	128	<i>rpsL</i> 128 A>G	43	Lys>Arg	Conocida	Estreptomicina	SM
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible
10055/11	<i>embB</i> (arabinosil transferasa)	1065	<i>embB</i> 1065 G>A	355	Sinónima	Conocida	Etambutol	susceptible
	<i>embB</i> (arabinosil transferasa)	1133	<i>embB</i> 1133 A>C	378	Glu>Ala	Conocida	Etambutol	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	61	<i>gyrA</i> 61 G>C	21	Glu>Gln	Predicha	Quinolonas	susceptible
	<i>gyrA</i> (subunidad A, DNA girasa)	284	<i>gyrA</i> 284 G>C	95	Ser>Thr	Conocida	Quinolonas	susceptible
	<i>katG</i> (catalasa peroxidasa)	944	<i>katG</i> 944 C>G	315	Ser>Thr	Conocida	Isoniazida	INH
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1333	<i>rpoB</i> 1333 C>T	445	His>Tyr	Conocida	Rifampicina	RIF
	<i>rpoB</i> (subunidad β, RNA polimerasa)	1355	<i>rpoB</i> 1355 T>C*	452	Leu>Pro	Conocida	Rifampicina	No concluyente
	<i>rpsL</i> (proteína ribosomal S12)	363	<i>rpsL</i> 363 A>G	121	Sinónima	Conocida	Estreptomicina	susceptible

*Variantes inciertas que no pudieron clasificarse en verdaderas mutaciones o errores de secuenciación

FLQ: fluoroquinolonas, RIF: rifampicina, SM: estreptomicina, INH: isoniazida, PZA: pirazinamida, EMB: etambutol y AMK: amikacina.

12. REFERENCIAS DE LOS ANEXOS:

-
- ¹ Waterfield T, Gloyn AL. Monogenic β -cell dysfunction in children: clinical phenotypes, genetic etiology and mutational pathways. *Pediatr Health* 2008;2:517-32.
- ² Harris TJ, McCormick F. The molecular pathology of cancer. *Nat Rev Clin Oncol*. 2010 May;7(5):251-65. Epub 2010 Mar 30.
- ³ McCarthy M. Genomics, Type 2 Diabetes and Obesity. *N Engl J Med* 2010;363:2339-50.
- ⁴ Instituto Nacional de Medicina Genómica (INMEGEN). Núcleo de discusión académica de enfermedades metabólicas y cardiovasculares. 2011. Distrito federal, México.
- ⁵ Ocaña A, Pandiella A. Personalized therapies in the cancer “omics” era. *Mol Cancer* 2010;9:202
- ⁶ Instituto Nacional de Medicina Genómica (INMEGEN). Núcleo de discusión académica de cáncer. 2011. Distrito federal, México.
- ⁷ Braggio E, Fonseca R. Unraveling the multiple myeloma genome in the next-generation sequencing era: challenges to translating knowledge into the clinic. *Expert Rev Hematol*. 2011 Dec;4(6):579-81. doi: 10.1586/ehm.11.64.
- ⁸ Technology Strategy Board. Stratified medicines programme: Tumour profiling and data capture to improve cancer care. Competition for collaborative R&D funding. January 2011. www.innovateuk.org (Accesado el 06//07//2011).
- ⁹ Krishnan V, Nestler E. The molecular neurobiology of depression. *Nature*. 2008 October 16; 455(7215): 894–902. doi:10.1038/nature07455.
- ¹⁰ Dong C, Wong M L, Licinio J. Sequence variations of ABCB1, SLC6A2, SLC6A3, SLC6A4, CREB1, CRHR1 and NTRK2: association with major depression and antidepressant response in Mexican-Americans. *Mol Psychiatry* (2009) 14, 1105–1118.
- ¹¹ Martin V, Shaw-Smith C. Review of genetic factors in intestinal malrotation. *Pediatr Surg Int* (2010) 26:769–781. DOI 10.1007/s00383-010-2622-5
- ¹² Wieacker P, Steinhard J. The Prenatal Diagnosis of Genetic Diseases. *Dtsch Arztebl Int* 2010; 107(48):857–62
- ¹³ Urban Z et al. Mutations in LTBP4 Cause a Syndrome of Impaired Pulmonary, Gastrointestinal, Genitourinary, Musculoskeletal, and Dermal Development. *Am J Hum Genet*. 85, 593–605, November 13, 2009.
- ¹⁴ Pober B. Overview of Epidemiology, Genetics, Birth Defects, and Chromosome Abnormalities Associated With CDH. *Am J Med Genet C Semin Med Genet*. 2007 May 15; 145C(2): 158–171. doi:10.1002/ajmg.c.30126.
- ¹⁵ Lidral A, Moreno L, Bullard S. Genetic Factors and Orofacial Clefting. *Semin Orthod*. 2008 June; 14(2): 103–114. doi:10.1053/j.sodo.2008.02.002.
- ¹⁶ Weir E. Congenital abdominal wall defects. *CMAJ* OCT. 14, 2003; 169 (8)
- ¹⁷ Wang Q. Molecular genetics of coronary artery disease. *Curr Opin Cardiol*. 2005 May ; 20(3): 182–188.
- ¹⁸ Bentley P, Peck G, Smeeth L, Whittaker J, Sharma P. Causal Relationship of Susceptibility Genes to Ischemic Stroke: Comparison to Ischemic Heart Disease and Biochemical Determinants. *PLoS ONE* 2010 5(2): e9136. doi:10.1371/journal.pone.0009136
- ¹⁹ Velagaleti R, O’Donnell C. Genomics of heart failure. *Heart Fail Clin*. 2010 January ; 6(1): 115–124. doi:10.1016/j.hfc.2009.08.002.
- ²⁰ Ferguson C. Genetic contributions to antisocial personality and behaviour: a meta-analytic review from an evolutionary perspective. *J Soc Psychol*. 2010 Mar-Apr; 150(2):160-80
- ²¹ Mead H, Beauchaine T, Shannon K. Neurobiological adaptations to violence across development. *Dev Psychopathol*. 2010 Winter;22(1):1-22.
- ²² Hohenester S, Oude-Elferink R, Beuers U. Primary biliary cirrhosis. *Semin Immunopathol*. (2009) 31:283–307. DOI 10.1007/s00281-009-0164-5
- ²³ Donaldson P. Genetics of Liver Disease: Immunogenetics and Disease Pathogenesis. *Gut* 2004;53:599–608. doi: 10.1136/gut.2003.031732
- ²⁴ Instituto Nacional de Medicina Genómica (INMEGEN). Núcleo de discusión académica en enfermedades infecciosas. 2011. Distrito federal, México.
- ²⁵ Loeffelholz M, Chonmaitree T. Advances in Diagnosis of Respiratory Virus Infections. *Int J Microbiol*. 2010;2010:126049
- ²⁶ Rello J, Pop-Vicas A. Clinical review: Primary influenza viral pneumonia. *Critical Care* 2009, 13:235 (doi:10.1186/cc8183)
- ²⁷ Kumpf O et al. Influence of genetic variations in TLR4 and TIRAP/Mal on the course of sepsis and pneumonia and cytokine release: an observational study in three cohorts. *Crit Care*. 2010;14(3):R103. Epub 2010 Jun 3.
- ²⁸ Chen Y et al. Toll-like receptor 2 gene polymorphisms, pulmonary tuberculosis, and natural killer cell counts. *BMC Med Genet*. 2010 Jan 30;11:17.

-
- ²⁹ Kasuga I et al. Role of genetic susceptibility to latent adenoviral infection and decreased lung function. *Respir Med*. 2009 Nov;103(11):1672-80. Epub 2009 Jun 6.
- ³⁰ Khoo US et al. Role of polymorphisms of the inflammatory response genes and DC-SIGNR in genetic susceptibility to SARS and other infections. *Hong Kong Med J*. 2008 Aug;14 Suppl 4:31-5.
- ³¹ Roth D, Jones A, Prosser C, Robinson JL, Vohra S. Vitamin D receptor polymorphisms and the risk of acute lower respiratory tract infection in early childhood. *J Infect Dis*. 2008 Mar 1;197(5):676-80.
- ³² Yende S et al. 4G/5G plasminogen activator inhibitor-1 polymorphisms and haplotypes are associated with pneumonia. *Am J Respir Crit Care Med*. 2007 Dec 1;176(11):1129-37. Epub 2007 Aug 29.
- ³³ Baghdadi J et al. An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *J Exp Med*. 2006 Jul 10;203(7):1679-84. Epub 2006 Jun 26.
- ³⁴ Mandell L, et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the Management of Community-acquired Pneumonia in Adults. *Clin Infect Dis*. 2007;44:S27-S72
- ³⁵ McPhee S, Papadakis M. *Current Medical Diagnosis & Treatment*. 2011. Edición 50ª Estados Unidos de Norteamérica. McGraw-Hill.
- ³⁶ Meng J., Rosenwasser L. J. Unraveling the Genetic Basis of Asthma and Allergic Diseases. *Allergy Asthma Immunol Res* 2010;2(4):215-227.
- ³⁷ Kumar A, Ghosh B. Genetics of asthma: a molecular biologist perspective. *Clin and Mol Allergy*. 2009, 7:7
doi:10.1186/1476-7961-7-7
- ³⁸ Rastogi M, LaFranchi S. Congenital hypothyroidism. *Orphanet Journal of Rare Diseases* 2010, 5:17
- ³⁹ Michels A, Eisenbarth G. Immunologic Endocrine Disorders. *J Allergy Clin Immunol*. 2010 February ; 125(2 Suppl 2): S226–S237. doi:10.1016/j.jaci.2009.09.053.
- ⁴⁰ Romero C, Nesi-França S, Radovick S. The Molecular Basis of Hypopituitarism. *Trends Endocrinol Metab*. 2009 December ; 20(10): 506–516. doi:10.1016/j.tem.2009.06.005.
- ⁴¹ Gejman P, Sanders A, Duan J. The Role of Genetics in the Etiology of Schizophrenia. *Psychiatr Clin North Am*. 2010 March ; 33(1): 35–66. doi:10.1016/j.psc.2009.12.003.
- ⁴² MacDonald A, Schulz S. What We Know: Findings That Every Theory of Schizophrenia Should Explain. *Schizophrenia Bulletin* vol. 35 no. 3 pp. 493–508, 2009. doi:10.1093/schbul/sbp017
- ⁴³ Campbell PJ, Morlock GP, Sikes RD, Dalton TL, Metchock B, Starks AM, Hooks DP, Cowan LS, Plikaytis BB, Posey JE. Molecular detection of mutations associated with first- and second-line drug resistance compared with conventional drug susceptibility testing of *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*. 2011 May;55(5):2032-41. Epub 2011 Feb 7.
- ⁴⁴ Cui Z, Wang J, Lu J, Huang X, Hu Z. Association of mutation patterns in *gyrA/B* genes and ofloxacin resistance levels in *Mycobacterium tuberculosis* isolates from East China in 2009. *BMC Infect Dis*. 2011 Mar 29;11:78.
- ⁴⁵ Sekiguchi J, Miyoshi-Akiyama T, Augustynowicz-Kopeć E, Zwolska Z, Kirikae F, Toyota E, et al. Detection of multidrug resistance in *Mycobacterium tuberculosis*. *J Clin Microbiol*. 2007 Jan;45(1):179-92. Epub 2006 Nov 15.
- ⁴⁶ Centers for Disease Control and Prevention. Treatment of Tuberculosis, American Thoracic Society, CDC, and Infectious Diseases Society of America. *Morbidity and Mortality Weekly Report (MMWR)* 2003;52(No. RR-11).
- ⁴⁷ Cuevas-Córdoba B, Zenteno-Cuevas R. Tuberculosis drogoresistente: mecanismos moleculares y métodos diagnósticos. *Enferm Infecc Microbiol Clin*. 2010 Nov;28(9):621-8. Epub 2010 Apr 9.
- ⁴⁸ Laurenzo D, Mousa SA. Mechanisms of drug resistance in *Mycobacterium tuberculosis* and current status of rapid molecular diagnostic testing. *Acta Trop*. 2011 Jul;119(1):5-10. Epub 2011 Apr 16.
- ⁴⁹ Reddy TB, Riley R, Wymore F, Montgomery P, Decaprio D, Engels R, et al. TB database: an integrated platform for tuberculosis research. *Nucleic Acids Res* 2009 Jan;37(Database issue):D499-508. Epub 2008 Oct 3.
- ⁵⁰ Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. *PLoS Med*. 2009 Feb 10;6(2):e(2).
- ⁵¹ Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073-81.