

Bases de datos del proyecto CONACyT: *“Exacerbación de Tuberculosis mediada por Diabetes: Un programa de Investigación clínica y experimental enfocado a la identificación de las huellas moleculares que definen la susceptibilidad de los pacientes con diabetes a desarrollar esta co-morbilidad”*,
Registro: A1-S-48232.

Convocatoria CB 2017-2018.

FICHA TÉCNICA

1. **Objetivo general del proyecto:** Identificar y caracterizar las huellas transcripcionales de monocitos de sangre total de pacientes que presentan la comorbilidad diabetes mellitus tipo 2-tuberculosis pulmonar activa.
2. **Propósito de la base.** La base se construyó para coleccionar la información clínica y de laboratorio de los sujetos de donde se obtuvieron las muestras de RNA total y que fueron seleccionados bajo los criterios de inclusión y exclusión establecidos en el proyecto. Se incluyeron los datos de sujetos sanos, pacientes con tuberculosis, tanto aquellos con como aquellos sin diabetes.

La base de datos se utilizó para recolectar la información clínica y sociodemográfica de la población de estudio en este proyecto. Permitted organizar y estratificar a los pacientes de cada grupo, ya que tiene todos los datos clínicos y de laboratorio de cada sujeto, con todas las variables necesarias para identificar claramente a los sujetos DM2 y con TB de cada grupo. La base sirvió además para el análisis de correlación de las variables clínicas con los resultados del análisis de expresión génica de cada grupo de sujetos, así como para identificar los genes diferencialmente expresados por cada grupo en los ensayos de micorarreglos y para la validación estadística de los datos en los ensayos de qPCR. Eventualmente esta base se complementarí con un mayor número de sujetos para generar un grupo de prueba para los ensayos de

validación de marcadores identificados y se usarian las variables clínicas para analizar estas en relación con la expresión de genes asociados con el binomio DM2-TB, lo cual podría generar pruebas pronosticas y diagnósticas usando estos biomarcadores validados para su uso en el Sector Salud

3. **Investigador principal a cargo.** Dr. Jose Antonio Enciso Moreno
4. **Recolección de la base de datos.** La inclusión de los sujetos y toma de muestras se realizó previa lectura y firma de la carta de consentimiento informado por parte de los sujetos participantes. Los sujetos de estudio se captaron de la consulta de las Unidades de Medicina Familiar y Hospitales de primer y segundo nivel del IMSS, que acudieron por diagnóstico de tuberculosis y/o diabetes. Los médicos establecieron el diagnóstico compatible con DM2 y tuberculosis en base al juicio clínico (antecedentes familiares, signos, síntomas) y exámenes de laboratorio que se realizan de rutina en los derechohabientes. Algunos de esos datos se incluyeron en la base. Particularmente, los datos de sujetos con o sin diabetes, sin tuberculosis, fueron seleccionados de la base de datos de PREVENIMSS y del expediente electrónico de la Unidades Médicas participantes, y considerando que los sujetos cumplieran con los criterios de inclusión del protocolo. Las muestras de los pacientes con tuberculosis (con o sin DM2), se capturaron en la UIBMZ durante la actividad que realiza como Centro de referencia para el Diagnóstico de tuberculosis, proceso que es supervisado por el programa de la Coordinación de vigilancia epidemiológica del IMSS. Los datos de los sujetos de estudio y sus muestras fueron colectados en diversos Estados del Centro y Norte de la Republica Mexicana incluyendo: Durango, San Luis Potosí, Aguascalientes, Jalisco, Tamaulipas, Nuevo León, Coahuila, Sinaloa, Baja California, Chihuahua y Zacatecas. El diagnóstico de tuberculosis se realizó en la UIBMZ con el algoritmo completo (BAAR, PCR por GeneXpert y Cultivo con determinación de multidrogresistencia por MGIT 960) y los datos también incluidos en la base de datos.

En nuestro estudio se incluyeron datos de 5 grupos de acuerdo a la siguiente estratificación:

Grupo Control. Compuesto por personas sanas sin infección y sin diabetes (controles no infectados sin DM2, CNI); Individuos de 18 a 65 años de edad.

Sujetos con diabetes mellitus tipo 2, determinados de acuerdo a los criterios de la Asociación Americana de Diabetes (ADA, por sus siglas en inglés), pero sin tuberculosis. Se consideraron dos subgrupos: a) un grupo con buen control de glucosa: y b) un grupo con pobre control glicémico (Glucemia en ayunas ≥ 126 mg/dl y < 298 mg/dl Diabetes establecida < 5 años de evolución. HbA1c $\geq 6,5\%$ y $< 12\%$. Se incluyeron individuos de 18 a 65 años de edad, con PPD y /o Quantiferon negativos y baciloscopía negativa. Sin síntomas de tuberculosis.

Pacientes con tuberculosis activa diagnosticados por primera vez (TBA) sin DM2 y que nunca hayan recibido tratamiento antituberculoso o lo hayan recibido por menos de treinta días.; Individuos con baciloscopía, GenXpert o cultivo positivo a *Mycobacterium tuberculosis* y sin rasgos clínicos y de laboratorio relacionados con DM2 de acuerdo a los criterios de la ADA. Con edad de 18 a 65 años.

Sujetos con el binomio tuberculosis activa y con DM2 (TBA-DM2) diagnosticados por primera vez sin tratamiento antituberculoso previo, o que lo hayan recibido por menos de treinta días; Individuos con baciloscopía, GenXpert y/o cultivo positivo a *Mycobacterium tuberculosis* y con rasgos clínicos y de laboratorio relacionados con DM2 de acuerdo a los criterios de la ADA. Con edad de 18 a 65 años.

Las siguientes variables antropométricas fueron incluidas en la base: peso del paciente en kg, talla del paciente en metros, datos de biometría hemática (BH), perfil de funcionamiento hepático (PFH), química sanguínea de 6 (glucosa, urea, creatinina, colesterol, triglicéridos, ácido úrico), y examen general de orina (EGO).

A todos los individuos que reunieron los criterios de selección antes descritos se les tomó una muestra de sangre periférica (aprox. 3-4 ml c/tubo) mediante el método de flebotomía convencional, colectando 3 tubos vacutainer de tapa roja y 4 tubos de tapa morada (Becton Dickinson, USA). A cada tubo con EDTA se les agregó 1 ml de RNAlater (Thermo Fisher Scientific, USA); la muestra se homogenizó adecuadamente, se sometió a cadena fría en la UIBMZ y una vez ahí, se colocaron dentro de un tubo Falcón de 50 ml para evitar derrames, almacenando en ultracongelador a -70°C hasta su uso. Estos tubos fueron utilizados para el análisis transcriptómico. En cuanto los tubos de tapa roja, dos fueron procesados para la seroteca y uno se utilizó para los análisis de química sanguínea. El suero para la seroteca se dividió en alícuotas de 0.5 ml utilizando crioviales estériles. Las alícuotas se prepararon para su envío a la UIBMZ en cadena fría, anotando perfectamente los datos del sujeto (nombre completo, tipo de muestra, fecha y las iniciales de quien tomó la muestra). Los tubos de tapa morada restantes fueron

utilizados en el sitio de recolección para la biometría hemática y para medir hemoglobina glicada y para estudios de rutina que se realiza a los derechohabientes del IMSS con DM2.

Descripción detallada de los métodos y técnicas para la recolección de datos.

5. **Periodo de recolección de datos.** Mencionar el inicio y fin de la recolección. En caso de ser una colección continua mencionar cuándo se planea cerrarla.

6. **Variables incluidas.** Mencionar el # de variables total colectadas.

7. Estrategia de aseguramiento para la proyección de datos

sensibles/personales. Desde el principio se estableció un folio unico para cada sujeto incluido en el estudio y solo el responsable de la base y el investigador a cargo del proyecto tenían acceso a la relación de los nombres de los pacientes y el folio asignado. Los accesos a las bases fueron controlados con clave de acceso. En las bases de trabajo solo se manejo el folio asignado **y nunca se incluyo** algun dato como número de seguridad solcial, INE, CURP, etc que pudiera identificar a los sujetos enrolados.

8. Fecha última de actualización. Mencionar la última fecha de captura de las bases que se envían. Desde el principio se estableció un codigo

9. **Mencionar el tipo de controles para la validación y verificación de la captura de los datos.** Participaron dos personas, un capturista y un curador. La base se reviso además por el responsable del proyecto. Los datos se confirmaron dos veces con los expedientes.

10. **Otras plataformas donde se encuentren disponibles las bases. Incluir URL.** Ninguna

11. **Otras fuentes de Financiamiento.** Ninguna más

12. **Seguimiento de la Cohorte en estudio.** No es un estudio longitudinal

13. Publicaciones. 1.- *Potential molecular patterns for tuberculosis susceptibility in diabetic patients with poor glycaemic control: A pilot study, sometido a la revista: Molecular and Cellular Probes, 2022.*

2.- *Transcriptome analysis identifies oxidative stress injury biomarkers for Diabetic Nephropathy in human patients, Sometido a la revista Archives of Medical Research. 2022.*

14. Otros. Cualquier información extra que sea pertinente precisar o agregar. Ninguna adicional.

ESPECIFICACIONES DE CONTENIDO Y FORMATO PARA LAS BASES

1. Entregar de preferencia en archivo Excel
2. La base de datos debe ser exhaustiva y contener todas las variables que se recabaron para las investigaciones.
3. Para proteger la confidencialidad de los participantes se debe eliminar aquella información sensible o que permita la identificación de los sujetos, como el nombre, número de seguridad social, número telefónico o correo electrónico, etc. Se puede cambiar esa información por un número de identificación para la propia base de datos (Folio, ID, etc.)
4. Cada fila debe corresponder a un sólo "registro" y cada columna a un sólo "campo" o variable.
5. Usar únicamente la primera fila para asignar el nombre de las variables, el cuál debe ser escrito con letra minúscula y un máximo de 10 caracteres, sin acentos, espacios ni caracteres especiales (sólo guion o guion bajo).
6. Cada campo debe almacenar las partes lógicas más pequeñas de información.
7. Evitar utilizar las primeras filas del Libro para colocar títulos o notas.
8. Evitar el uso de distintos tipos de letra, resaltados en color o negritas dentro de las celdas.
9. Evitar la inclusión de datos redundantes.

10. Evitar incluir cualquier tipo de análisis estadístico, fórmulas, gráficos o tablas con resultados.
11. Evitar dejar filas o columnas en blanco.
12. Evitar abreviaturas y plurales. Use formas homogéneas y estandarizadas para nomenclaturas.
13. Muy importante: en un segundo Libro del Excel se debe incluir el índice de la base de datos con el nombre corto de la variable (como aparece en la base de datos), el nombre completo, una descripción breve, tipo de variable (clasificación) y su operacionalización (posibles valores).