



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

ESTUDIO IN SILICO DEL
PROCESO DE DORMANCIA DE
Mycobacterium Tuberculosis
MEDIANTE UN MODELO DE
REDES BOOLEANAS

TESIS PARA OPTAR EL GRADO DE
MAESTRO EN BIOQUÍMICA Y
BIOLOGÍA MOLECULAR

DAVID JAMES MONTANO QUIROZ

LIMA – PERÚ

2024

ASESOR:

Dr. Daniel Guerra Giráldez

CO-ASESOR:

Dra. Elisa Domínguez Hüttinger

JURADO EVALUADOR

Dra. PATRICIA SHEEN CORTAVARRIA

PRESIDENTE

Dr. PABLO TSUKAYAMA CISNEROS

VOCAL

Mg. PABLO ALEJANDRO FONSECA ARROYO

SECRETARIO

DEDICATORIA

A mis padres, mi esposa y mis hijos Apolo y Vincent.

AGRADECIMIENTOS

Al Dr. Daniel Guerra por su asesoría durante la elaboración
y la ejecución de este trabajo de investigación.

A la Dra. Elisa Domínguez, por su mentoría y experiencia
compartida, que fueron fundamentales para poder concluir
este trabajo.

FUENTES DE FINANCIAMIENTO

Trabajo de Investigación Autofinanciado.

9% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Coincidencias menores (menos de 10 palabras)

Exclusiones

- N.º de coincidencia excluida

Fuentes principales

- 3%  Fuentes de Internet
- 1%  Publicaciones
- 9%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

-  **Caracteres reemplazados**
32 caracteres sospechosos en N.º de páginas
Las letras son intercambiadas por caracteres similares de otro alfabeto.
-  **Texto oculto**
2 caracteres sospechosos en N.º de página
El texto es alterado para mezclarse con el fondo blanco del documento.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Tabla de contenido

I.	INTRODUCCIÓN.....	1
II.	MARCO TEÓRICO	5
2.1.	TUBERCULOSIS EN INFECCIÓN, LATENCIA Y REACTIVACIÓN	5
2.2.	LA RED DE REGULACIÓN GENÉTICA DE MTB.....	8
2.3.	ESTADO DE DORMANCIA DE MTB.....	14
2.4.	MODULACIÓN DEL SISTEMA INMUNOLÓGICO POR MTB	21
2.4.1.	Modulación de las células fagocíticas por parte de mtb.....	21
2.4.2.	Perfil de citoquinas durante la infección de mtb	22
2.5.	MODELADO DE SISTEMAS COMPLEJOS CON REDES BOOLEANAS.....	25
2.6.	MODELOS BOOLEANOS DE REDES GENÉTICAS.....	27
2.7.	APLICACIONES BIOLÓGICAS DE REDES BOOLEANAS	33
2.7.1.	Robustez y evolucionabilidad.....	33
2.7.2.	Modelos booleanos de sistemas biológicos.....	36
2.7.3.	Modelos booleanos de dormancia de mtb	40
III.	JUSTIFICACIÓN.....	44
IV.	HIPÓTESIS	46

V. OBJETIVOS.....	47
5.1. OBJETIVO GENERAL.....	47
5.2. OBJETIVOS ESPECÍFICOS	47
VI. METODOLOGÍA.....	48
6.1. OBTENCIÓN DE DATOS DE EXPRESIÓN DIFERENCIAL DURANTE LA HIPOXIA	48
6.1.1. Obtención de las lecturas de RNA seq.	48
6.1.2. Alineamiento de las lecturas al genoma de MTB.....	49
6.1.3. Verificación del alineamiento de las lecturas al genoma de MTB usando el programa de visualización IGV.....	50
6.1.4. Obtención de la expresión diferencial para todos los genes de MTB, respecto de la muestra de completa oxigenación (81.62 % de OD).....	52
6.2. BINARIZACIÓN DE LOS NIVELES DE EXPRESIÓN DIFERENCIAL 53	
6.3. OBTENCIÓN DEL AGRUPAMIENTO DE LOS ESTADOS DE HIPOXIA.....	55
6.4. OBTENCIÓN DE LA RED DE DORMANCIA DE MTB.....	55
6.4.1. Verificación y corrección de las funciones booleanas de la red previa de 24 nodos.....	57
6.4.2. Reducción del número de nodos de la Red.....	62
6.4.3. Obtención de los atractores de la red booleana de dormancia.....	63

6.4.4.	Obtención de funciones reguladoras <i>in silico</i>	63
6.5.	ANÁLISIS DE MUTACIONES EN LA RED	64
6.5.1.	Verificación de importancia de los nodos en la estructura general .	64
6.5.2.	Análisis de la trayectoria hacia el atractor.....	64
6.5.3.	Mutaciones de activación	64
6.5.4.	Verificación de mutaciones de inactivación.....	65
6.6.	INTRODUCCIÓN DE NODO RELACIONADO CON LAS CITOQUINAS.....	65
6.6.1.	Búsqueda de genes de MTB relacionados con las Citoquinas.	65
6.6.2.	Análisis de la interacción del nodo relacionado a citoquinas con la red de dormancia.....	65
6.6.3.	Análisis bioinformático de la relación del nodo relacionado a las citoquinas con la red de dormancia.	66
6.6.4.	Análisis de los atractores de la red ampliada con el nodo Rv2031c.	
	66	
6.7.	COMPARACIÓN DE LA RED DE DORMANCIA CON REDES ALEATORIAS DEL MISMO GRADO DE ENTRADA PARA LOS NODOS.	
	66	
6.8.	INTRODUCCIÓN DE NODOS RELACIONADOS AL ESTADO DE REACTIVACIÓN.	69
6.8.1.	Búsqueda de genes de MTB relacionados con la reactivación.....	69

6.8.2.	Análisis de las interacciones de los nodos relacionados a la reactivación con la red de dormancia de MTB.....	69
6.8.3.	Análisis de los atractores de la red ampliada con los nodos Rpf dormancia	70
VII.	RESULTADOS	71
7.1.	EXPRESIÓN DIFERENCIAL DURANTE LA HIPOXIA	71
7.2.	BINARIZACIÓN DE LOS NIVELES DE EXPRESIÓN DIFERENCIAL	74
7.3.	OBTENCIÓN DE AGRUPAMIENTO DE LOS ESTADOS DE HIPOXIA.....	76
7.4.	OBTENCIÓN DE LA RED DE DORMANCIA DE MTB.....	81
7.4.1.	Red de dormancia mínima (RDM)	81
7.4.2.	Obtención de los atractores de la red de dormancia mínima.....	83
7.4.3.	Funciones reguladoras <i>in silico</i>	91
7.5.	ANÁLISIS DE MUTACIONES EN LA RED	95
7.5.1.	Verificación de importancia de los nodos en la estructura general .	95
7.5.2.	Análisis de la trayectoria hacia el atractor	97
7.5.3.	Mutaciones de activación	108
7.5.4.	Verificación de mutaciones de inactivación.....	114
7.6.	INTRODUCCIÓN DE UN NODO RELACIONADO CON LAS CITOQUINAS.....	119

7.6.1. Nodo relacionado a las citoquinas	119
7.6.2. Análisis de las integración del nodo Rv2031c con la red dormancia 119	
7.6.3. Análisis bioinformático de la relación de Rv2031c con la red de dormancia.	119
7.6.4. Análisis de los atractores de la red de dormancia ampliada con el nodo Rv2031c.....	123
7.7. COMPARACIÓN DE LA RED MÍNIMA DE DORMANCIA CON REDES ALEATORIAS DEL MISMO GRADO DE ENTRADA PARA LOS NODOS.	127
7.8. INTRODUCCIÓN DE NODOS RELACIONADOS AL ESTADO DE REACTIVACIÓN.	132
7.8.1. Nodos relacionados al estado de reactivación	132
7.8.2. Análisis de la interacción de los nodos relacionados a la reactivación con la red de dormancia de MTB	133
7.8.3. Análisis de los atractores de la red ampliada con los nodos relacionados a la reactivación.....	135
7.8.4. Análisis de trayectorias hacia el atractor de reactivación.....	138
VIII.DISCUSIÓN.....	143
IX. CONCLUSIONES.....	149
X. DIRECCIONES A FUTURO	151
XI. LIMITACIONES.....	152

XII. BIBLIOGRAFÍA 153

XIII. ANEXOS

ABREVIATURAS

MTB: *Mycobacterium tuberculosis*.

Attr: Atractor.

IL: Interleuquina.

Fn: Función booleana.

RBA: Red booleana aleatoria.

RBS: Red booleana sincrónica.

FT: Factor de transcripción.

FTs: Factores de transcripción.

P(k)in: Distribución de conectividad de entrada.

P(k)out: Distribución de conectividad de salida.

RHT: Genes de respuesta temprana a la hipoxia.

RHD: Genes de respuesta duradera a la hipoxia.

RDM: Red de dormancia mínima de *Mycobacterium tuberculosis*.

PAMP: Patrones moleculares asociados a patógenos.

OD: Oxígeno disuelto.

ILTB: Infección latente de tuberculosis.

RESUMEN

La tuberculosis, cuyo agente causal es *Mycobacterium tuberculosis* (MTB), presenta dos tipos de infección: una activa con una alta tasa de mortalidad, y otra inactiva denominada infección latente de tuberculosis (ILTB), con una elevada prevalencia que alcanza a la cuarta parte de la población mundial.

Durante la ILTB, MTB se encuentra en estado de dormancia, caracterizado por una baja actividad metabólica. Ante un debilitamiento del sistema inmunológico del huésped, el patógeno puede reactivarse y provocar una infección activa. Sin embargo, hasta la fecha se desconocen los mecanismos moleculares involucrados en la transición de la dormancia a la reactivación.

Las redes booleanas permiten modelar sistemas complejos y sus atractores representan estados estables del sistema que pueden asociarse a estados celulares, por lo tanto, pueden ser validados con datos experimentales de expresión genética en dichos estados.

En este estudio, se ha obtenido un modelo booleano de la red de dormancia de MTB, producto de la modificación de un modelo inicial propuesto por Bose y colaboradores e integrando los datos más recientes de ChipSeq y de expresión genética durante la dormancia y la reactivación. Esta red contiene 26 genes de MTB involucrados en los procesos de dormancia y de reactivación como los factores promotores de resucitación (Rpf): RpfC, RpfD, RpfE y el gen *hspX* vinculado a las citoquinas del huésped. Tres de los atractores de esta red, se corresponden con los estados

fundamentales binarizados de MTB: Hipoxia Temprana, Hipoxia Tardía y Reactivación.

Mediante mutaciones en la red, se evaluaron que nodos son necesarios para alcanzar los estados atractores. Así, se observó que los genes Rv0324 y Rv3574 son necesarios para que la red alcance el atractor de Hipoxia Tardía. Por otro lado, los genes RpfD, Rv1033c y Rv2011c son necesarios para alcanzar el atractor de Reactivación. Estos resultados sugieren un posible rol terapéutico de estos últimos genes como dianas terapéuticas para el desarrollo de fármacos que inhiban la reactivación de la tuberculosis.

Palabras clave: *Mycobacterium tuberculosis*, redes booleanas, dormancia, reactivación.

ABSTRACT

Tuberculosis, caused by *Mycobacterium tuberculosis* (MTB), manifests in two forms: active tuberculosis, characterized by high mortality rates, and latent tuberculosis infection (LTBI), prevalent in approximately one-fourth of the global population.

During LTBI, MTB enters a dormant state characterized by low metabolic activity. Upon weakening of the host's immune system, the pathogen can reactivate, leading to active infection. However, the molecular mechanisms underlying the transition from dormancy to reactivation remain poorly understood to date.

Boolean networks enable modeling of complex systems, with their attractors representing stable system states that can correspond to cellular states, and can be validated with experimental gene expression data in those states.

In this study, a Boolean model of MTB dormancy network was developed by modifying an initial model proposed by Bose et al., integrating recent ChipSeq and gene expression data during dormancy and reactivation phases. This network encompasses 26 MTB genes involved in dormancy and reactivation processes, including resuscitation-promoting factors (Rpf) such as RpfC, RpfD, RpfE, and the hspX gene linked to host cytokines. Three attractors of this network correspond to fundamental binary states of MTB: Early Hypoxia, Late Hypoxia, and Reactivation.

Through mutations in the network, it was evaluated which nodes are necessary to reach the attractor states. Thus, it was observed that genes Rv0324 and Rv3574 are required for the network to reach the Late Hypoxia attractor. Furthermore, genes

RpfD, Rv1033c, and Rv2011c are necessary to reach the Reactivation attractor. These findings suggest a potential therapeutic role for these latter genes as therapeutic targets for the development of drugs to inhibit tuberculosis reactivation.

Keywords: Mycobacterium tuberculosis, Boolean networks, dormancy, reactivation.

I. INTRODUCCIÓN

La tuberculosis representa un grave problema de salud mundial. El agente causal es *Mycobacterium tuberculosis* (MTB). Se conocen dos tipos de infección: tuberculosis activa e infección latente de tuberculosis (ILTB). La primera posee una alta letalidad, que es responsable de la muerte de millones de personas todos los años (1). Por otro lado, la ILTB posee una alta prevalencia, como lo demuestra el hecho de que un cuarto de la población mundial sufra de ILTB, siendo los países subdesarrollados los más afectados, algunos con hasta el 80% de su población infectada (2).

La infección por MTB inicia cuando el patógeno ingresa a través de las vías respiratorias de su huésped, una vez dentro de los alveolos, el patógeno infecta a los macrófagos alveolares y origina la fase aguda de la enfermedad. Esta infección puede ser eliminada por completo, dependiendo de la virulencia de la cepa y del estado del sistema inmunológico del huésped. Si el sistema inmunológico del huésped no es capaz de eliminar por completo a los patógenos, estos pueden permanecer en un estado de dormancia, estableciéndose la ILTB. La dormancia de MTB se caracteriza principalmente por una baja tasa de replicación y una alta resistencia a los antibióticos. La ILTB puede durar desde un par de años hasta décadas. La duración de la ILTB y su transición a una infección activa de tuberculosis, depende que se

produzca un eventual debilitamiento del sistema inmunológico del huésped, que puede ser producto de una infección o un tratamiento inmunosupresor (3, 4).

Gracias al uso de modelos *in vitro* del estado de dormancia de MTB, se han identificado dos grandes grupos de genes que se activan en respuesta a la hipoxia: los genes de respuesta temprana a la hipoxia (RHT), que comprende alrededor de 48 genes, y los genes de respuesta tardía a la hipoxia (RHD), que comprenden cerca de 230 genes (5, 6).

La dormancia emerge como resultado de las interacciones de MTB con su entorno. Para estudiar este proceso se han llevado a cabo investigaciones en genómica, proteómica y metabolómica de MTB durante la dormancia. Estos estudios nos han permitido conocer los principales factores del patógeno involucrados en el proceso de dormancia (7). Sin embargo, todavía hacen falta más estudios que permitan revelar sus interacciones con los factores del sistema inmune del huésped, e incluso el efecto de estados patológicos del paciente como el HIV y la Diabetes (8).

Este sistema complejo, en el que MTB regula la expresión de sus genes para evadir la respuesta inmunológica del huésped, se puede entender como un estado de equilibrio en el que el huésped contiene la infección, mientras que el patógeno sobrevive dentro del huésped en estructuras denominadas granulomas.

La biología de sistemas nos permite modelar sistemas complejos y encontrar sus principales propiedades. Las redes booleanas constituyen una herramienta versátil para modelar sistemas biológicos.

Las redes booleanas son una herramienta matemática utilizada para modelar sistemas dinámicos complejos, como las redes de regulación genética. Una ventaja significativa de las redes booleanas es su capacidad para simplificar y comprender la dinámica de los sistemas (9, 10). Además, son fáciles de construir y analizar, lo que las hace útiles para investigar sistemas biológicos complejos. En particular, los atractores de la red booleana, que son los estados estables de la red, pueden ser asociados con estados biológicos específicos (11, 12). Esta característica las hace especialmente útiles para estudiar sistemas biológicos como las redes de regulación genética y para identificar posibles objetivos terapéuticos.

Los datos de expresión genética permiten inferir redes de regulación en diferentes condiciones experimentales (13, 14). La representación dinámica de redes también nos permite simular las perturbaciones producidas en el sistema al introducir mutaciones de inactivación o activación a nodos de la red (15). En el caso del proceso de dormancia de MTB, las redes booleanas pueden ser útiles porque permiten modelar la interacción de los diferentes componentes del sistema sin necesidad de conocer la cinética de interacción entre los componentes. Esta cinética depende de variables difíciles de obtener, como la tasa de consumo de oxígeno del MTB en

estado activo y de dormancia, o la tasa de muerte del MTB en estado de dormancia por día (16).

El enfoque de la biología de sistemas permite comprender las relaciones entre las redes del patógeno el sistema inmune del huésped, lo que abre la posibilidad de simular antimicrobianos e inmunomoduladores de manera conjunta como tratamiento. Esta estrategia, que consiste en combatir no solo al patógeno, sino también en modular la respuesta inmune del huésped, resulta fundamental para reducir el daño tisular que suele acompañar a esta patología incluso después de eliminar al patógeno del huésped. En este sentido, se están llevando a cabo varios estudios en fase preclínica para evaluar la eficacia de inmunomoduladores en combinación con terapia antibiótica para el tratamiento de la tuberculosis (17, 18).

II. MARCO TEÓRICO

2.1. TUBERCULOSIS EN INFECCIÓN, LATENCIA Y REACTIVACIÓN

M. tuberculosis tiene como huésped obligado al ser humano desde hace más de 50 000 años y ha coevolucionado con él, desarrollando mecanismos de evasión del sistema inmunológico (19). Su envoltura es un claro ejemplo de la compleja red de interacciones entre MTB y su huésped, siendo esta cápsula altamente inmunogénica y conteniendo varios patrones moleculares asociados a patógenos (PAMP) (20). Esto configura diferentes vías de internalización de MTB por los macrófagos u otras células fagocíticas, y al mismo tiempo determina diversos mecanismos de modulación de la respuesta inmune (21).

El ciclo de infección de MTB inicia cuando una partícula en suspensión conteniendo al patógeno es expulsada de una persona con tuberculosis activa y es inhalada por otra persona sana. Una vez dentro de los alvéolos, el patógeno es reconocido por las células dendríticas y macrófagos mediante el reconocimiento de sus PAMP y es fagocitado como parte de la respuesta inmune innata. En este momento, MTB modula la respuesta del macrófago a su favor para poder multiplicarse y desarrollar una infección aguda (22). El proceso inflamatorio ocasionado por la lisis de macrófagos desencadena la migración de neutrófilos y de otras células presentadoras de antígenos como las células dendríticas (CD) al lugar de la infección. Estas fagocitan partículas del patógeno, las degradan y exponen los

antígenos en su superficie. Al mismo tiempo, estas células migran a los nódulos linfáticos en busca de linfocitos T y B con receptores para los antígenos de MTB. Cuando encuentran a estos linfocitos, los activan y se monta la respuesta inmune adaptativa, que es más específica y fuerte contra MTB, y que va permitir, dependiendo del estado inmunológico del huésped y de la virulencia de la cepa, la eliminación completa del patógeno (3). Esta respuesta está montada principalmente por linfocitos T CD4+ y linfocitos T CD8+, que no solo pueden activar a los macrófagos, sino que también tienen la capacidad de eliminar células infectadas con MTB (4). Sin embargo, si esta respuesta no es suficiente, lo que se logra es solo contener al patógeno dentro de estructuras denominadas granulomas, que son estructuras conformadas por macrófagos infectados y bacilos libres en su interior, rodeados por una capa de linfocitos, CD y neutrófilos. MTB puede permanecer años e incluso décadas dentro de estas estructuras en un estado denominado de dormancia, en el que es altamente resistente a los antibióticos. Ante un debilitamiento del sistema inmunológico del huésped, debido, por ejemplo, a un tratamiento inmunomodulador, o a enfermedades como HIV o diabetes, o a la senectud, se puede reactivar el patógeno y volver a generar la tuberculosis activa (3). Ver figura 1.

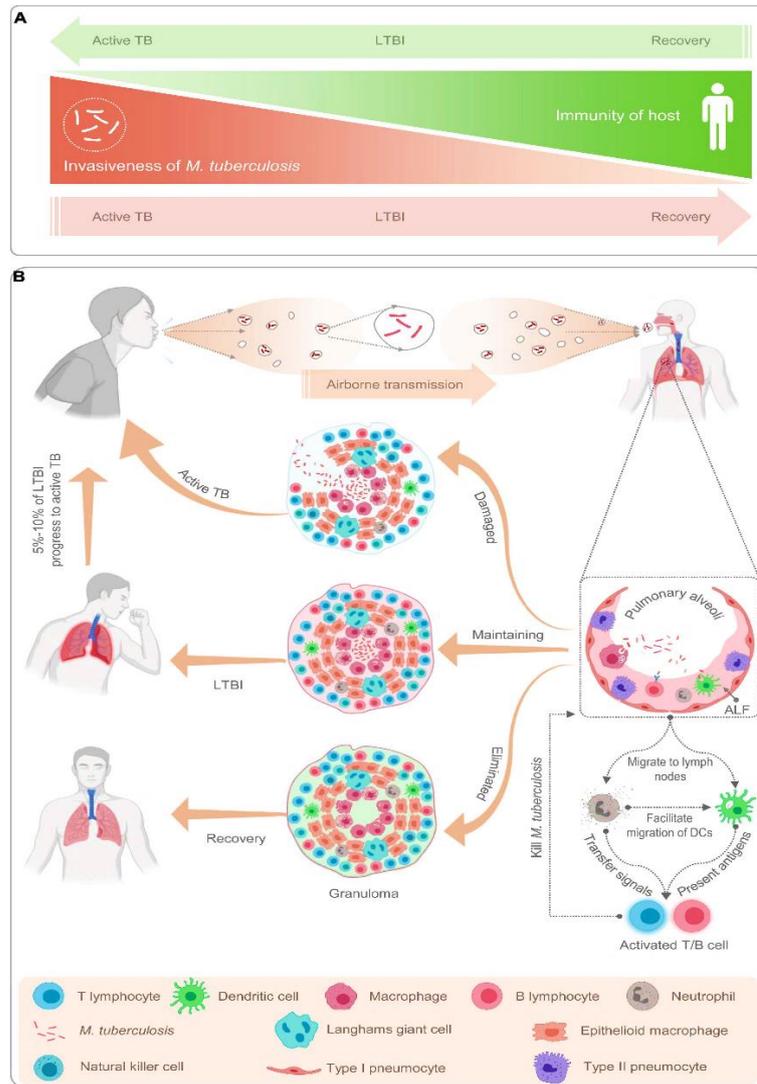


Fig. 1 Diagrama esquemático de la infección latente de tuberculosis. La ILTB representa un equilibrio entre la inmunidad del huésped y la capacidad invasiva de *M. tuberculosis*; cualquier cambio en el estado inmunológico del paciente podría alterar este equilibrio (A). MTB excretada por pacientes con tuberculosis activa es inhalada por personas sanas y, una vez en los alvéolos pulmonares, es fagocitada por células presentadoras de antígenos, como macrófagos, neutrófilos y células dendríticas (B). Posteriormente, estas células presentadoras migran a los nódulos linfáticos para activar células T que, a su vez, se desplazan hacia el sitio de infección por MTB, formando granulomas (B). Si la inmunidad del huésped es suficientemente fuerte, las células del sistema inmunitario eliminarán al patógeno. Si la inmunidad del huésped es débil, MTB se replicará en el tejido granulomatoso, rompiendo la restricción granulomatosa y dando lugar a la tuberculosis activa. Si la inmunidad del huésped y

la capacidad invasiva de *M. tuberculosis* están equilibradas, el huésped desarrollará una infección latente ILTB (A, B). Tomado de: Differential Diagnosis of Latent Tuberculosis Infection and Active Tuberculosis: A Key to a Successful Tuberculosis Control Strategy (3).

Los mecanismos genéticos de la reactivación son objeto de constante estudio. Se sabe que los Factores Promotores de Resucitación (Rpf) son esenciales para este proceso; en MTB, esta familia comprende cinco proteínas: RpfA, RpfB, RpfC, RpfD y RpfE. Estudios estructurales sugieren que estas proteínas tienen actividad enzimática de peptidoglucano hidrolasas, lo que sería relevante para su función fisiológica. Además, la delección de estos genes no afecta su crecimiento normal, pero impide su reactivación desde el estado de dormancia (23). En estudios en modelos murinos de latencia, se ha demostrado cómo la inmunosupresión y la diabetes pueden afectar la expresión de RpfD y RpfB durante la fase de dormancia; en particular, la diabetes puede aumentar la expresión de RpfD durante la dormancia, lo que provocaría la reactivación de MTB (8). A pesar de estos avances, todavía no se conocen los mecanismos moleculares que permiten a las redes genéticas de MTB transitar desde un estado de dormancia a la reactivación.

2.2. LA RED DE REGULACIÓN GENÉTICA DE MTB.

Una red se puede representar mediante un grafo, que consta de un conjunto de nodos conectados por aristas. En las redes de regulación genética, los nodos son los genes y las aristas representan los tipos de interacción, que puede ser activadora o

inhibidora entre los nodos. En una célula, los factores de transcripción (FTs) son los responsables de inhibir o activar a sus genes objetivo, por lo tanto, las aristas siempre van a tener una dirección que va desde el factor de transcripción hacia el gen objetivo, conformándose una red dirigida. Estas conexiones generan dos tipos de distribuciones: **1)** una distribución de conectividad de entrada ($P(k)_{in}$), determinada por la probabilidad de que un determinado gen tenga un número k de FTs que lo regulan y **2)** una distribución de conectividad de salida ($P(k)_{out}$), determinada por la probabilidad de que un FT particular regule k genes.

Existen diversos métodos para construir las redes de regulación genética. Uno de ellos consiste en construir la red directamente a partir de datos experimentales de expresión genética mediante algoritmos de generación de funciones booleanas *in silico* (13, 14, 24). Otro enfoque implica analizar la interacción de los FTs con los promotores de sus genes objetivos mediante estudios ChipSeq, para determinar las conexiones entre los nodos (25). Además, se utilizan datos de expresión genética diferencial en presencia y ausencia del FT, para evaluar si la relación entre un FT y sus genes objetivo es de activación o inhibición para cada conexión. Si la expresión de un gen aumenta significativamente cuando el FT está presente, se considera una relación de activación. Por el contrario, si la expresión de un gen disminuye significativamente en presencia del FT, se considera una relación de inhibición. En la figura 2 se muestra una explicación más detallada del método ChipSeq.

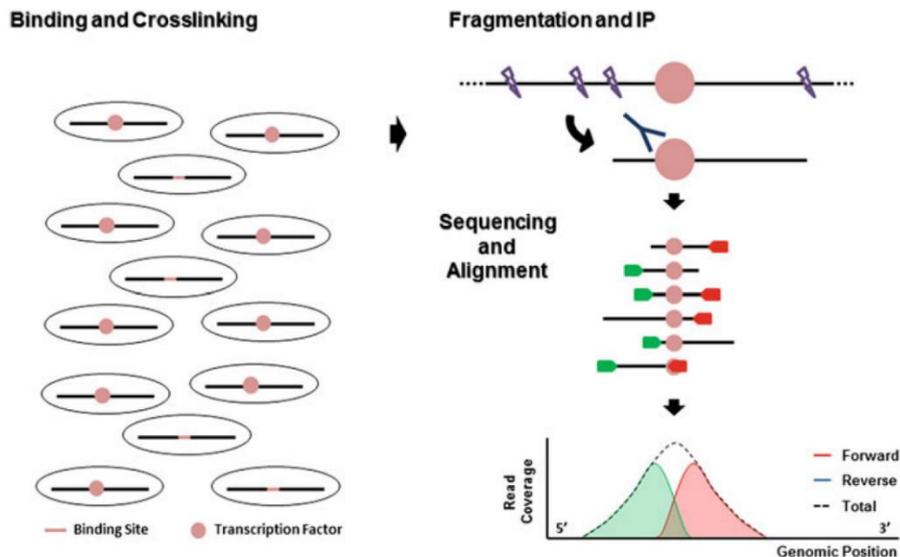


Fig. 2 Ensayo Chip-Seq. A la izquierda, los FTs representados por círculos rosados unidos al ADN se unen a una región específica del ADN y se entrecruzan por la acción del formaldehído, en una siguiente etapa el ADN es fragmentado por sonicación o por un proceso enzimático. Los complejos proteína-ADN son inmunoprecipitados mediante anticuerpos específicos para cada uno. Se separan los fragmentos de ADN y son secuenciados. Mediante alineamiento al genoma del organismo de estudio se identifica la región a la que se une cada FT. Tomado de: ChIP-Seq and the Complexity of Bacterial Transcriptional Regulation (25).

MTB posee un genoma compuesto de 4 411 529 pares de bases (26). En 2008, Balázs y colaboradores construyeron la primera red de regulación genética de MTB, utilizando tres fuentes de información: la plataforma MtbRegList, una base de datos que está dedicada al análisis de expresión génica y datos de regulación en MTB (27); la red de regulación de *E. coli* para inferir regulaciones por ortología con los pares TF-gen objetivo de *E. coli*, y los de resultados de experimentos de microarreglos durante condiciones de normoxia e hipoxia (28). Esta red contenía 45 FTs y se organizó en 18 orígenes, definidos como un FT y sus genes regulados

directa e indirectamente. Uno de estos orígenes comprendía al FT DosR, fundamental en la adaptación de MTB a la hipoxia (29). Ver figura 3.

Al analizar la topología de esta red, los autores encontraron que la red de regulación para MTB presentaba una distribución de tipo exponencial para $P(k)_{in}$, similar a las encontradas para *E. coli* y *S. cerevisiae*. Sin embargo, para $P(k)_{out}$ no se encontró una distribución de ley de potencias, a diferencia de las redes de *E. coli* y *S. cerevisiae* que sí obedecen a una distribución de ley de potencias (29).

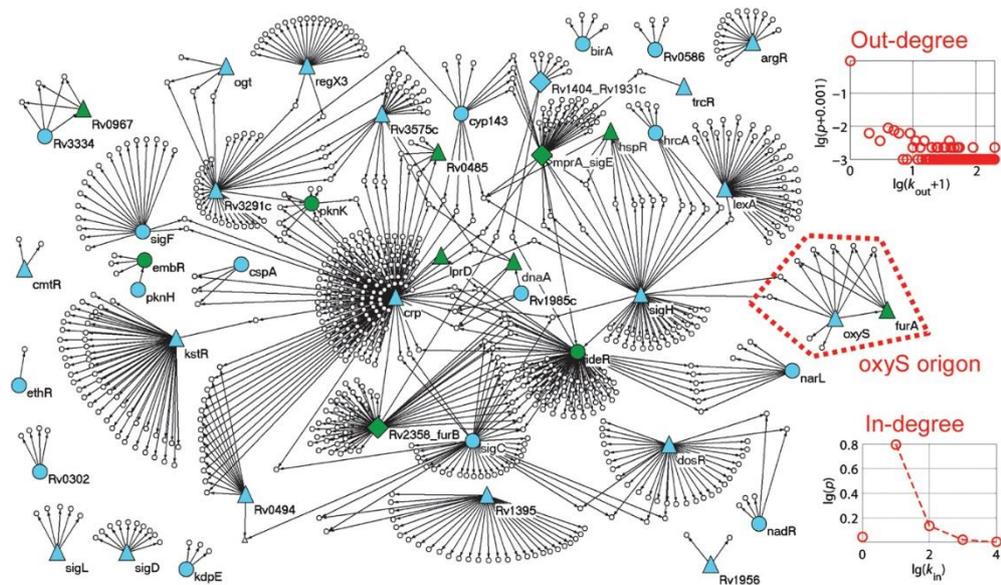


Fig. 3 Red de Regulación genética de MTB. Se muestran la red de 45 FTs conformando 18 orígenes. Los nodos de entrada (genes sin reguladores transcripcionales conocidos) se muestran en azul, mientras que los nodos de tránsito (FT con reguladores transcripcionales conocidos) se muestran en verde. Los nodos blancos representan nodos de salida (genes que codifican proteínas sin actividad de FT). Las aristas muestran la dirección de la regulación. En la gráfica superior derecha se presenta el $P(k)_{out}$ y en la gráfica inferior derecha se muestra el $P(k)_{in}$. Tomado

de: The temporal response of the *Mycobacterium tuberculosis* gene regulatory network during growth arrest (29).

En 2011, Sanz y colaboradores llevaron a cabo un estudio que demostró que la red de regulación génica de MTB sigue una ley de potencias para $P(k)_{out} \approx k^{-\gamma}$, con un exponente γ de aproximadamente $0,99 \pm 0,07$ para la conectividad y una distribución de tipo exponencial para $P(k)_{in}$ (30). Ver figura 4. Una distribución exponencial para $P(k)_{in}$ implica que la mayoría de los genes tienen pocos reguladores, ya que la probabilidad de tener k reguladores disminuye exponencialmente a medida que aumenta k . Por otro lado, $P(k)_{out}$, sigue una distribución de ley de potencias, lo que indica que la probabilidad de tener k reguladores disminuye más lentamente con el aumento de k , lo que significa que es probable encontrar FTs que regulan a muchos genes. Estos FTs, conocidos como nodos Hubs, están implicados en los procesos regulatorios fundamentales, como la adaptación a la hipoxia y el establecimiento del estado de dormancia en MTB. Por ejemplo, los reguladores *dosR* y *phoR* son algunos de estos nodos Hubs (30). Ver figura 5.

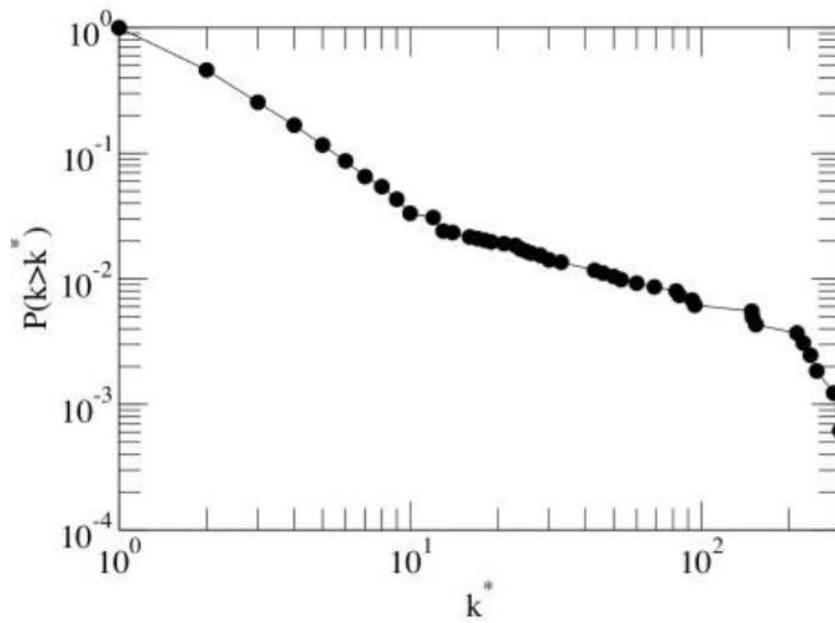


Fig. 4 Distribución de conectividad $P(k)_{out}$ de la red genética de MTB. El eje $P(k > k^*)$ indica la probabilidad de encontrar un nodo con una conectividad mayor a k^* (representado en escala logarítmica). Tomado de: The Transcriptional Regulatory Network of *Mycobacterium tuberculosis* (30).

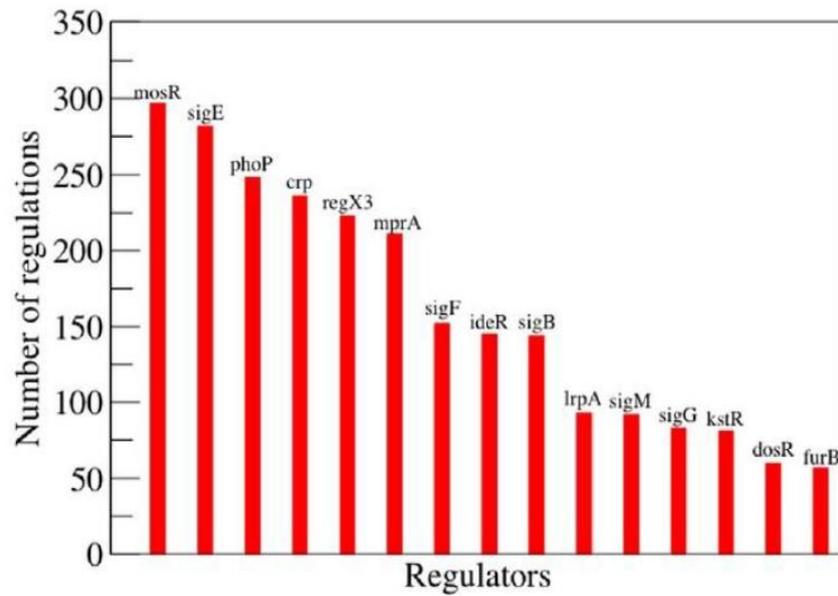


Fig. 5 Reguladores Hubs en la red de regulación genética de MTB. En el eje de ordenadas se representa el número de nodos regulados por los reguladores que forman el eje de las abscisas, donde: mosR, sigE, phoP, crp, regX3, mprA, sigF, ideR, sigB, lrpA, sigM, sigG, kstR, dosR, furB representan FTs de la red de regulación genética de MTB. Tomado de: The Transcriptional Regulatory Network of *Mycobacterium tuberculosis* (30).

2.3. ESTADO DE DORMANCIA DE MTB

El estado de dormancia se caracteriza por la detención del crecimiento de MTB. Este fenotipo se origina como una respuesta adaptativa a condiciones extremas de estrés, como falta de nutrientes (C, N, P, Mg, Fe) o un ambiente adverso que incluye un pH ácido o la presencia de sustancias reactivas de oxígeno y nitrógeno (31). Para estudiar este estado de dormancia, se han desarrollado diversos modelos in vitro

que permiten simular condiciones de estrés específicas, como la limitación de nutrientes o de oxígeno. Uno de los más importantes es el desarrollado por el Dr. Lawrence Wayne, quien demostró cómo una disminución paulatina en la concentración de oxígeno en un cultivo de MTB genera un estado no replicativo que le permite sobrevivir por largos periodos de tiempo. Este estado de dormancia también confiere una resistencia a antibióticos, como la isoniazida y rifampicina (32).

MTB puede sobrevivir en este estado de dormancia, dentro de los granulomas por largos periodos y reingresar a una fase activa cuando el sistema inmune del huésped se debilita. Para mantenerse viable en este estado, MTB ha desarrollado varios mecanismos para modular su actividad. Por ejemplo, disminuye la expresión de componentes de la cadena del transporte de electrones, reduciendo así su producción energética y ralentizando su crecimiento (33). Sin embargo, MTB debe mantener un estado energético mínimo con un potencial de membrana estable que le permita posteriormente reactivarse. Para lograrlo, modifica sus rutas metabólicas, utilizando, por ejemplo, la vía del glioxilato, que le permite utilizar los ácidos grasos como fuente de energía. Además, se produce un incremento en la expresión de la isocitrato liasa, lo que le permite utilizar la producción de isocitrato para mantener su potencial de membrana, especialmente en condiciones de baja concentración de oxígeno, como las que se encuentran en los granulomas (33). Ver figura 6.

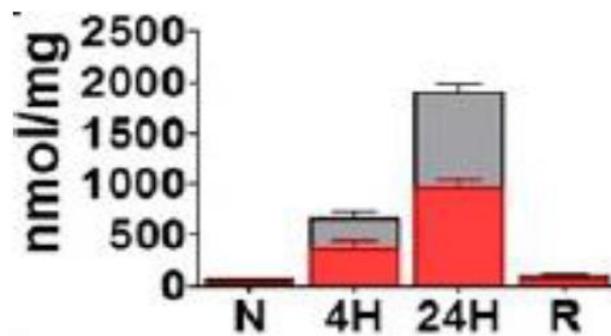


Fig. 6 Expulsión de Isocitrato Vs tiempo de exposición a distintas concentraciones de Oxígeno. N: 24 horas y 21% O₂; 4H: 4 horas y 1% O₂; 24H: 24 horas y 1% O₂ y R: 24 horas de re-aereación a 21% O₂ luego de 48 horas de incubación a 1% O₂. Durante la fase hipóxica, se observa un incremento en la expulsión de isocitrato al medio extracelular para contribuir a mantener el potencial de membrana de MTB. Tomado de: Multifunctional essentiality of succinate metabolism in adaptation to hypoxia in *Mycobacterium tuberculosis* (33).

De acuerdo con el modelo de inducción de la dormancia mediante la hipoxia, se pudo determinar que en la respuesta a la hipoxia existe: una fase inicial (RHT), que comprende la activación de alrededor de 48 genes por parte del sistema de dos componentes Dos R/Dos S; y una segunda fase (RHD), que se activa luego de días en hipoxia y comprende la activación de cerca de 230 genes (5). Para comprender mejor la red genética de dormancia de MTB dentro de la red de regulación genética de MTB, Galagan y colaboradores realizaron experimentos de Chip-Seq en condiciones de normoxia e hipoxia (1% de oxígeno). Analizando los resultados de estos experimentos, no se encontraron diferencias significativas en la topología de la red en estas dos condiciones (34). En esta red se evidenció la presencia de nodos Hub de hipoxia temprana, como Rv0081 y nodos Hub implicados en la respuesta a

la hipoxia tardía, como Rv0324 y Rv0494. Sin embargo, la mayoría de los nodos tienen baja conectividad. Ver figura 7. El análisis de esta red revela que el nodo Rv0081 es fundamental durante la transición de normoxia a hipoxia. Mediante un análisis de una subred regulatoria, que contempla a los FTs implicados en la respuesta a la hipoxia y al metabolismo lipídico, se observa que Rv0081 origina una fuerte inducción hacia el catabolismo de lípidos y degradación del colesterol, mediante la inhibición del Hub KstR (34). Ver figura 8.

Estos resultados son consistentes con una interconexión entre la respuesta a la hipoxia, el metabolismo de lípidos y la producción de lípidos de la pared celular en respuesta a la falta de oxígeno. Estas interacciones permiten regular el metabolismo de MTB para persistir en un estado de dormancia.

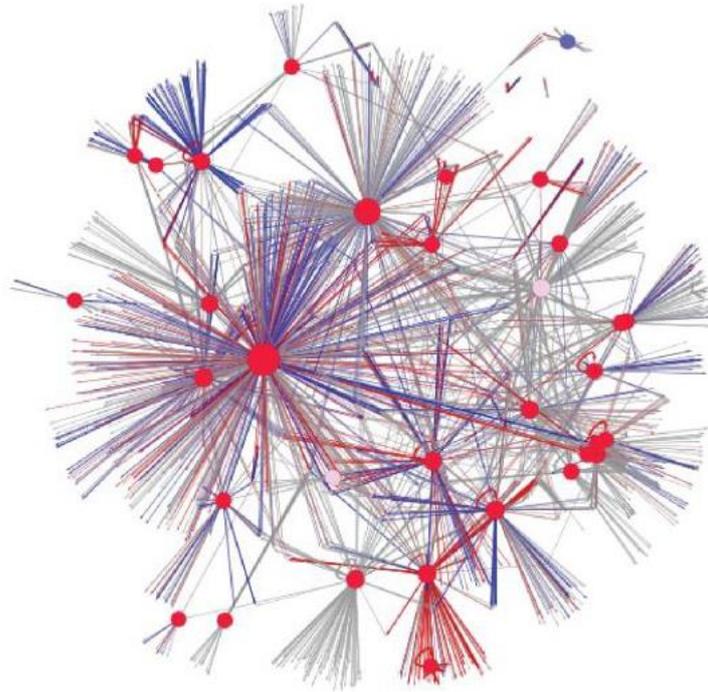


Fig. 7 Red de regulación para 40 FTs de MTB. Las aristas de color rojo representan inhibición y las de color azul activación. Los nodos representan genes y los nodos rojos FTs. El tamaño de los nodos de los FTs es directamente proporcional a su grado de salida. La red solo representa las interacciones en las que el FT se une corriente arriba del gen. Tomado de: The *Mycobacterium tuberculosis* regulatory network and hypoxia (34).

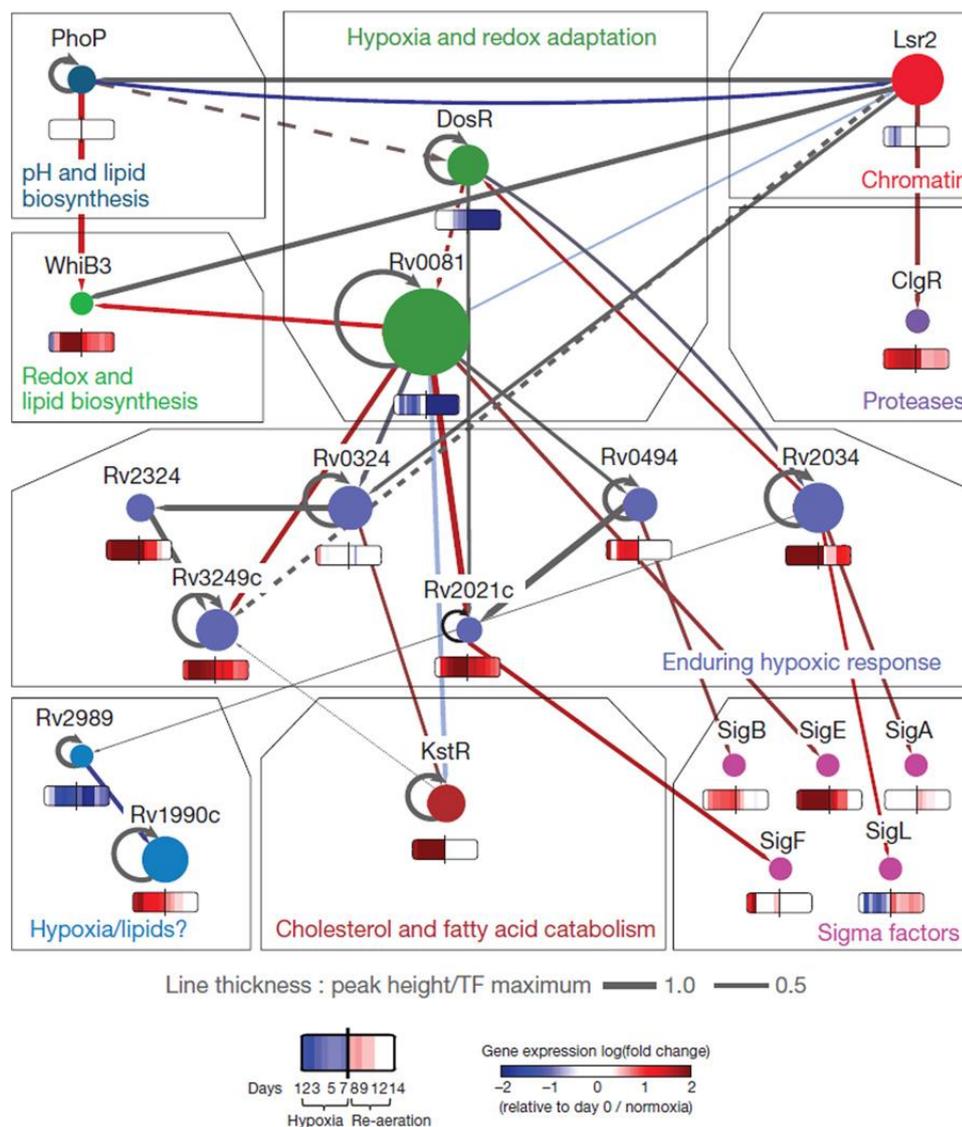


Fig. 8 Subredes de regulación de MTB en hipoxia y normoxia. En esta figura se presentan los principales nodos de las subredes de regulación genética, y el tamaño relativo de los nodos es directamente proporcional a su conectividad. Las aristas de color azul representan inhibición, mientras que las de color rojo indican activación. Tomado de: The *Mycobacterium tuberculosis* regulatory network and hypoxia (34).

El nodo principal de la subred de hipoxia temprana (RHT) es Dos R, que forma parte del sistema de dos componentes Dos R/Dos S (también denominado DevR/S y Rv3133c/Rv3132c). Este sistema está altamente expresado en cepas virulentas de

MTB en comparación con las no virulentas (6), (22). El sistema consta de la proteína Dos S, que se encarga de percibir la concentración de O₂, CO y NO, y de la proteína efectora DevR. En ausencia de O₂, se activa Dos S, y a su vez activa a Dos R, que activará a los genes de respuesta a la hipoxia temprana (22). Si la hipoxia se mantiene por días, MTB iniciaría la vía de ingreso a la fase de dormancia. Ver figura 9.

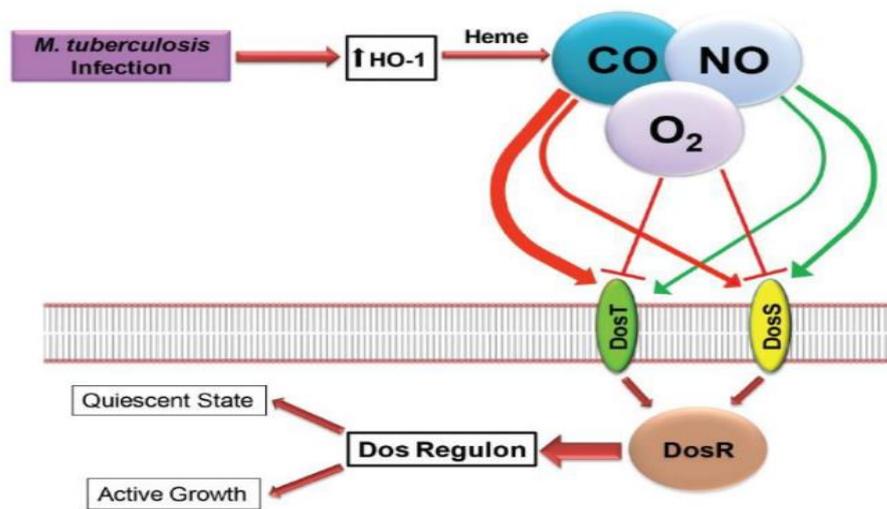


Fig. 9 El rol del CO, NO y O₂ en la activación del sistema DosR/T/S. Las flechas verdes indican la activación del sistema mediante CO o NO, mientras que las flechas rojas señalan su inactivación mediante O₂. Tomado de: *Mycobacterium tuberculosis*: macrophage takeover and modulation of innate effector responses (22).

Una vez activado el sistema DevR/S, se desencadenan eventos como la activación de factores como Rv0081, el cual se expresa en el estado temprano de hipoxia. La transición del estado de hipoxia temprana al de hipoxia tardía está mediada por el

factor Rv3334, que es activado por Rv0081. Rv3334, a su vez, inhibe la expresión de Rv0081 y activa un tercio de los genes involucrados en la respuesta a la hipoxia tardía (35).

2.4. MODULACIÓN DEL SISTEMA INMUNOLÓGICO POR MTB

2.4.1. Modulación de las células fagocíticas por parte de MTB

MTB posee múltiples estrategias para sobrevivir dentro del macrófago. Por ejemplo, la modificación de proteínas de su superficie (36), secretando quinasas como las PtpB (37), o puede realizar modificaciones epigenéticas en el núcleo del macrófago (38).

Otro grupo de estrategias consiste en evitar la maduración del fagosoma (22). Otro mecanismo importante de modulación es la interrupción de la cascada de señalización ERK 1/2 del huésped, lo que inhibe la producción de las citoquinas IL-6 y TNF, y favoreciendo así la supervivencia del patógeno (39). Además, MTB altera la vía de reconocimiento de ADN y ARN exógenos, lo que conlleva a la inhibición de la producción de interferones asociados con una reducción de la resistencia del huésped al patógeno (40).

El tráfico vesicular dentro de las células fagocíticas también es manipulado por MTB y puede escapar del fagosoma en condiciones que favorecen la infección de otras células vecinas (41).

MTB puede modular la autofagia y la apoptosis de las células fagocíticas, evitando la presentación de sus antígenos y una respuesta inflamatoria adaptativa mediada por citoquinas proinflamatorias como INF- γ , que podría eliminar la infección (40). Uno de los mecanismos por los que MTB evita la autofagia del huésped es mediante miR-125a, que bloquea al factor asociado a la resistencia a la radiación UV (UVRAG), inhibiendo así la autofagia (42).

El sistema de ubiquitina también es modulado por MTB con proteínas como la tirosina fosfatasa PtpA, que se une a ubiquitina y puede defosforilar las proteínas Jnk y p38 fosforiladas del huésped, lo que conduce a la supresión de la respuesta innata (43).

2.4.2. Perfil de citoquinas durante la infección de MTB

Además de los antígenos que MTB exhibe en su superficie, posee otras proteínas como Rv3131, que tienen capacidad de desencadenar una respuesta de citoquinas específicas. Esta proteína, perteneciente al regulón DosR, es una nitroreductasa que puede desempeñar una función protectora contra especies reactivas de nitrógeno. Además, es capaz de inducir un perfil de citoquinas proinflamatorias como IFN- γ , TNF- α , IL-1 β y IL-8 en células periféricas mononucleares (44). Este perfil de citoquinas favorece una respuesta pro-inflamatoria que se ha demostrado ser muy efectiva en la erradicación de la infección. Sin embargo, MTB ha co-evolucionado con los seres humanos, y esta coexistencia le ha permitido modular el perfil de ci-

toquinas del huésped a su favor. En principio, MTB sobrevive evitando la activación de las citoquinas proinflamatorias como el TNF, lo cual dificulta su eliminación (45).

La regulación del perfil de citoquinas es fundamental para la modulación de la respuesta adaptativa. En particular, el INF- γ y la IL-6 favorecen la diferenciación de los linfocitos T vírgenes hacia linfocitos ThCD 4+ y Th 17. Estas células favorecen una respuesta inflamatoria y la activación de los macrófagos para eliminar a los patógenos. Sin embargo, esta respuesta debe ser modulada por linfocitos T reguladores (Treg) para evitar el daño tisular que favorecería la diseminación de MTB al resto de tejidos (46).

El perfil de citoquinas también está implicado en la diferenciación de las células Natural Killer (NK) y en su capacidad para lisar monocitos infectados, siendo IL-2 e IL-12 estimuladores de su actividad lítica. Se considera que, en estadios tempranos de la infección, las células NK juegan un rol fundamental porque lisan no solo macrófagos y células dendríticas infectadas, sino también bacterias libres mediante receptores TLR2. Las células NK activadas también expresan diversas citoquinas, entre ellas IFN- γ , TNF- α y GM-CSF. Por lo tanto, no solo tienen actividades líticas, sino también funciones moduladoras de la respuesta inmune adaptativa (47). Ver figura 10.

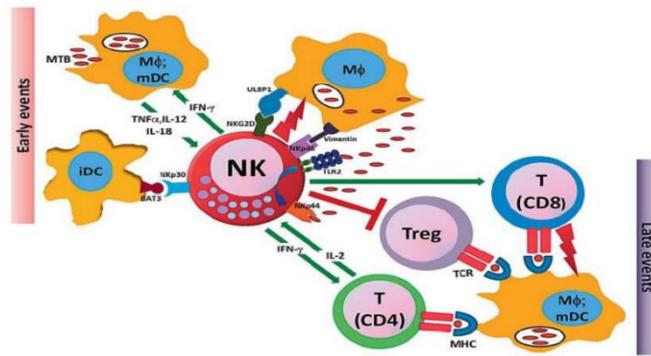


Fig. 10 Efectos de las células NK durante los estadios tempranos y tardíos de la infección. En una fase temprana, las NK pueden ejercer una función lítica y favorecer una acción lítica al reconocer señales de estrés en macrófagos y células dendríticas. En estados tardíos, pueden ejercer funciones regulatorias de células T. Tomado de: Natural Killer Cells: A Coherent Model for Their Functional Role in Mycobacterium tuberculosis Infection (47).

Se ha evaluado el perfil de liberación de citoquinas en personas con infección latente respecto de controles sanos. Así, se encontró que existe una mayor cantidad de citoquinas proinflamatorias (GM-CSF, IFN- γ , IL-1 β , IL-2, IL-6, IL-8, IL-12p70, TNF- α) en personas con tuberculosis latente y también un aumento de citoquinas antiinflamatorias IL-5 y IL-10 asociadas a linfocitos Th2 en dicho grupo (48). Otro marcador importante que se encuentra durante la infección latente es el marcador HLA DC expresado en células T CD4 + específicas de MTB, que se observaron luego de la exposición a antígenos de MTB (49).

En su intento por contener la infección, el sistema inmune promueve la formación de una estructura de contención denominada granuloma, que contiene a los bacilos en estado de dormancia. Sin embargo, durante la reactivación, su interior se vuelve necrótico y tiende a romperse, diseminando la infección.

Se ha encontrado que los macrófagos espumosos están asociados con la progresión de los granulomas hacia un estado activo. En un estudio en modelos murinos, se inhibió la enzima diacilglicerol O-aciltransferasa 1, lo que produce una disminución en la cantidad de triacilglicerol en el interior del granuloma, verificándose una disminución en la carga bacteriana. Este proceso está asociado con una disminución en el reclutamiento de neutrófilos al granuloma y con una disminución en las interleuquinas inflamatorias, como $IL1\alpha$ (50).

2.5. MODELADO DE SISTEMAS COMPLEJOS CON REDES BOOLEANAS

Los sistemas complejos pueden definirse como una red de interacciones de muchos componentes cuyo comportamiento en conjunto es emergente. Es decir, el comportamiento del sistema resulta de la interacción de sus partes y no se puede inferir directamente de las propiedades de sus elementos particulares, por lo tanto, para su estudio se requiere un abordaje necesariamente integrador (51).

El estudio de las redes complejas comprende dos áreas fundamentales: **(1)** el estudio de sus propiedades topológicas, como su distribución de conectividad; **(2)** el estudio de sus propiedades dinámicas, como la sincronización. Ejemplos de esta última propiedad son los procesos físicos como la sincronización de dos péndulos que se sostienen de una misma viga, o la sincronización del brillo de un grupo de luciérnagas (52).

El modelo de Erdős-Rényi (ER) fue la primera aproximación a una red compleja. Este modelo considera una red aleatoria de N nodos con probabilidad P de tener una conexión entre dos nodos y donde el número total de enlaces (E) está dado la siguiente ecuación:

$$E = N \frac{(N-1)}{2}$$

Para un determinado número M de conexiones se obtiene una distribución de grados de probabilidad binomial. En esta distribución, existe una probabilidad P de que un par de nodos elegido al azar esté conectado (éxito) y una probabilidad $1-P$ de que no estén conectados (fracaso). El grado k de un nodo se define como el número de aristas a las que está conectado un nodo en particular. Cuando el número de nodos N tiende a infinito, la distribución de grados de los nodos tiende a la distribución de Poisson (53). Esta distribución cumple la siguiente ecuación:

$$P(x=k) = \frac{\mu^k \cdot e^{-\mu}}{k!}$$

Donde $P(x=k)$ representa la probabilidad de que un nodo tenga grado igual a k ; μ representa el promedio de grado de los nodos de la red $\langle k \rangle$; y k representa el grado de los nodos. Por su naturaleza, las distribuciones binomiales y de Poisson tienen un valor promedio que representa el comportamiento del sistema (Ver anexo 1) (54). Sin embargo, las distribuciones de grados observadas experimentalmente

siguen una ley de potencias, donde el grado promedio $\langle k \rangle$ no representa al sistema. En este tipo de distribución, la mayoría de los nodos tienen un valor bajo de k y, al mismo tiempo, existen algunos nodos con un valor elevado de k .

Muchas redes de sistemas reales tienen distribución de grados que obedece a una ley de potencias con $P(k)_{out} \sim k^{-\gamma}$, como en el caso de la red de regulación genética de MTB (30). En general, las redes con distribución de ley de potencias tienen su exponente γ con valores entre 1 y 3. Ver tabla 1.

Red	Tamaño	$\langle k \rangle$	γ_{out}	γ_{in}
WWW	2×10^8	7,5	2.72	2.1
Coautores de publicaciones en Neurociencia	209 293	11.54	2.1	2.1
Contactos sexuales	2810		3.4	3.4
Metabólica, <i>E. coli</i>	778	7.4	2.2	2.2
Actores del cine	212 250	28.78	2.3	2.3

Tabla 1 Exponentes de redes complejas que siguen una ley de potencias. $\langle k \rangle$ = grado promedio de la red, k = grado del nodo, γ_{in} , γ_{out} son los exponentes de la distribución de grados de entrada y de salida respectivamente. Tomado de: Statistical Mechanics of Complex Networks (53).

2.6. MODELOS BOOLEANOS DE REDES GENÉTICAS

El dogma central de la biología establece que los genes son la fuente de la información hereditaria de todos los organismos vivos. Esta información es transcrita a ARN, para luego -en muchos casos- ser traducida en proteínas. En los

organismos pluricelulares, la información alojada en el o los cromosomas es la misma en cada célula, sin embargo, la expresión diferencial de los genes da lugar a distintos fenotipos. Estos fenotipos son estables, aunque es posible transitar de un fenotipo a otro, por ejemplo, en respuesta a cambios ambientales. Esta variación en la expresión diferencial es robusta y sufre pocas variaciones a través de las generaciones de una especie. No obstante, es importante señalar que también pueden sufrir variaciones y generar cambios evolutivos a escalas de tiempo más largas. El primer paso hacia una comprensión de cómo la interacción de los genes podría originar diversos fenotipos lo proporcionó Waddington, quien propuso un modelo de cómo las variaciones de actividad de los genes, que forman parte de una red, podrían variar el paisaje epigenético y conducir a diferentes fenotipos celulares (55). Ver figura 11.

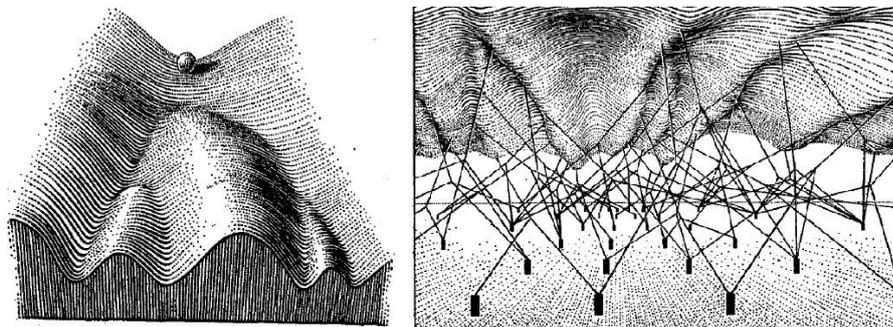


Fig. 11 Paisaje epigenético de Waddington. Los genes son representados como una red que genera el paisaje epigenético. Tomado de: *The Strategy of the Genes* (55).

En este modelo, cada valle del paisaje epigenético conforma un fenotipo celular, donde un estímulo externo podría modificar la forma del paisaje epigenético, favoreciendo un fenotipo Y sobre un fenotipo X. Ver figura 12.



Fig. 12 Variación del paisaje epigenético de Waddington. X e Y representan dos fenotipos. La flecha representa un estímulo externo. Tomado de: *The Strategy of the Genes* (55).

Posteriormente, Kauffman en 1969 utilizó una red booleana aleatoria (RBA) para modelar las configuraciones posibles en redes genéticas (56). En este modelo, los genes conforman los nodos de un grafo dirigido, siendo los vértices las interacciones de activación o represión entre los nodos. Estos nodos solo pueden encontrarse en dos estados posibles: activo o inactivo, representados por 1 y 0 respectivamente. Esta aproximación, que en principio parece simplista, es suficiente para describir el comportamiento de los elementos del sistema, tal como se ha demostrado en los estudios del desarrollo floral, el ciclo celular de la levadura y en el desarrollo embrionario de *D. melanogaster* (11, 12). El estado de cada gen va a depender de fun-

ciones booleanas de las entradas de cada nodo. Estas funciones lógicas están formadas por una combinación de las operaciones lógicas fundamentales. Ver figura 13.

AND	0	0	OR	0	0	NOT	0	1
	0	1		0	1		1	0
	1	0		1	0		1	0
	1	1		1	1		1	0

Fig. 13 Operadores lógicos.

De forma general, el modelo de Kauffman se puede representar como un conjunto de N genes, denotados como “g”. Donde el estado de cada gen “g” al tiempo $t+1$ es el resultado de evaluar al tiempo t una función “Fn”, la cual se define mediante la siguiente ecuación:

$$g(t+1) = Fn(gn_1(t), gn_2(t), \dots, gn_{kn}(t))$$

Donde F_n representa una función booleana de kn reguladores ($gn_1, gn_2, \dots, gn_{kn}$) al tiempo t . Es decir, el estado de los kn genes reguladores al tiempo t determinará el estado del gen regulado al tiempo $t+1$ (56). Ver figura 14.

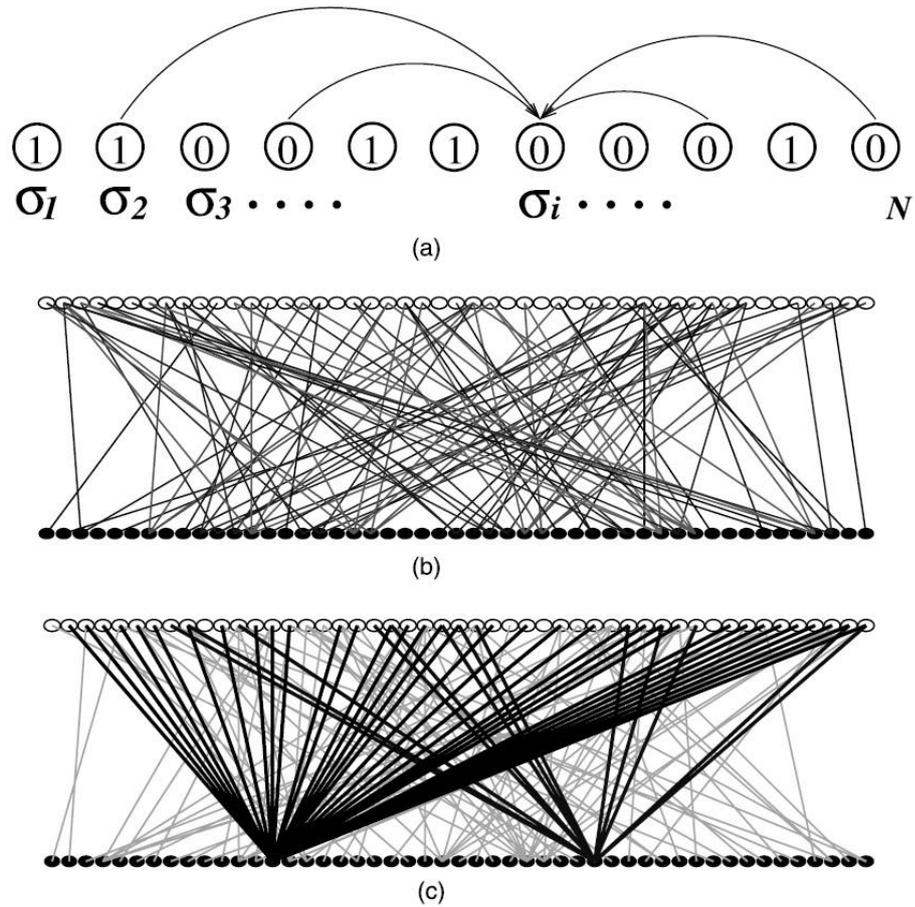


Fig. 14 Arquitectura de redes. a) Red de Kauffman con $k=4$. b) Red con distribución de Poisson para $P(k)$ in, se observa un $\langle k \rangle = 2$. c) Red con distribución libre de escala con exponente $\gamma = 2.5$ para $P(k)$ in y un valor con $\langle k \rangle \sim 2$. Sin embargo, su arquitectura es muy diferente al caso anterior. Tomado de: Boolean dynamics of networks with scale-free topology (57).

En el modelo original de Kauffman, estas funciones “Fn” se asignan de manera aleatoria desde el número total posible de funciones, es decir, $2^{(2^k)}$, donde k es el grado de los nodos. Para $k=2$, habrían $2^{(2^2)}$, es decir, 16 funciones posibles. El grado

de cada gen se establecen como una constante, conformando una red dirigida aleatoria con N nodos. En este tipo de redes, donde cada uno de los N nodos puede tomar solo dos estados posibles 1 y 0, existen 2^N estados posibles del sistema. Por ejemplo, para el caso de una red de 3 nodos y $k=2$, existen 8 estados posibles como entradas al tiempo t . Al tiempo $t+1$, todos los nodos del sistema toman el valor que resulta de evaluar las funciones lógicas de sus entradas. Dependiendo de las funciones lógicas de entrada, el sistema tenderá a estados particulares puntuales o cíclicos en los que permanecerá indefinidamente, sin importar los cambios en t . Estos estados se denominan atractores y se corresponderían con los valles del paisaje epigenético predicho por Waddington (55), como se muestra en la figura 15.

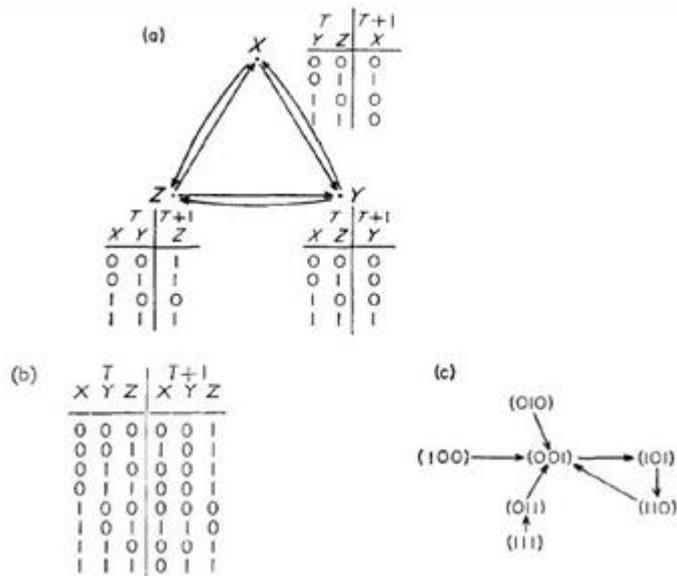


Fig. 15 Ejemplo de una red booleana de 3 variables. a) Red de 3 nodos con $k=2$. Donde se observan 3 funciones booleanas aleatorias para cada uno de los nodos b) Estados de la red al tiempo t y $t+1$. Donde cada nodo alcanza su estado al tiempo $t+1$ al evaluar su función booleana aleatoria de sus dos nodos de entrada c) A medida que evoluciona el sistema a través del tiempo, termina alcanzando un atractor cíclico de longitud 3, constituido por los estados 001-101-110. Tomado de: Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets (56).

Kauffman estudió los estados estables cíclicos de redes con 15, 50, 64, 100, 191, 400, 1024, 4096 y 8191 nodos, para $k=2$, y redes con $N = 50$ para $k=3$, obteniendo atractores cíclicos con pocos estados. Por ejemplo, para $k=2$ y $N=400$, la media del tamaño de los atractores era de solo 8 estados. Esto contrasta con los enormes atractores que se producen si $K=1$ o $K=N$, que originan atractores con millones de estados. Estos resultados llevaron a Kauffman a concluir que estas redes aleatorias con grado 2 y 3 es probable que tengan relevancia biológica (57).

Posteriormente, se descubrió que muchas redes reales tienen grado promedio similar, pero con topología libre de escala (53), como se ilustra en la figura 14.

2.7.APLICACIONES BIOLÓGICAS DE REDES BOOLEANAS

2.7.1. Robustez y evolucionabilidad

Si a un sistema se le aplican perturbaciones puntuales, es decir, se cambia el estado de activación de una fracción de genes particulares, y dichas perturbaciones se

propagan por toda la red con el tiempo, la red entra en un estado denominado “caótico”. Por otro lado, si estas perturbaciones se anulan rápidamente, se dice que la red está en un estado “ordenado”. Por último, si estas perturbaciones se mantienen en el tiempo en zonas restringidas de la red, decimos que el sistema se encuentra en un estado “crítico”, como se muestra en la figura 16. Para que los organismos puedan desarrollarse en interacción con el entorno, es muy beneficioso que sus redes biológicas operen en un estado crítico, de manera que sean estables para mantener el fenotipo y lo suficientemente flexibles para modificarse y evolucionar (58).

Con datos experimentales de expresión genética, se ha podido encontrar las funciones booleanas asociadas a cada gen, para organismos modelo. Específicamente, se han estudiado las perturbaciones a las redes de regulación genética de *S. cerevisiae*, *B. subtilis*, *D. melanogaster* y *A. thaliana*, encontrándose que estarían operando en estado crítico (59), como se muestra en la figura 17.

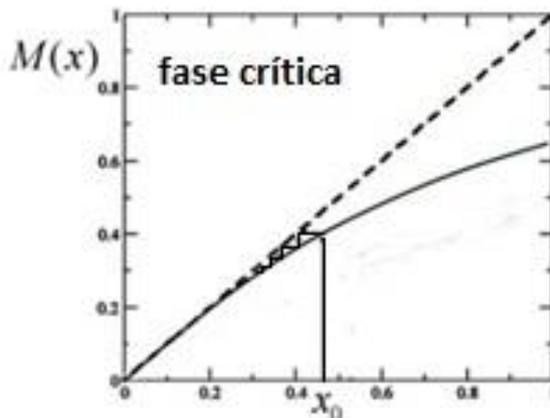


Fig. 16 Gráfica de una red en estado crítico. Se muestra la gráfica de la función recursiva $M(x)$ que describe como evoluciona la red en función del grado de perturbación x . En la gráfica, las iteraciones están representadas por las flechas. Se puede observar que al iterar la función $M(x)$, esta se aproxima lentamente a la recta identidad. Por lo tanto, la perturbación no desaparece, pero tampoco aumenta caóticamente, lo que mantiene la red en un estado crítico. Tomado de: Boolean dynamics of networks with scale-free topology (57).

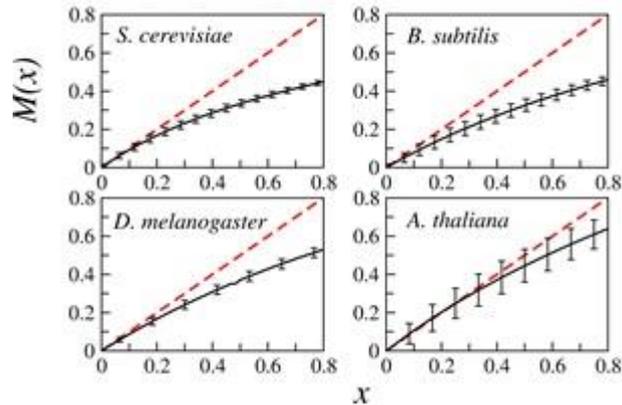


Fig. 17 - Dinámica crítica de organismos modelo. Se muestra la gráfica del mapeo $M(X)$ en función del grado de perturbación x . La línea punteada corresponde a la recta identidad. Como se observa en las gráficas, $M(x)$ se aproxima a la recta identidad en el origen. Por lo tanto, las redes genéticas de los 4 organismos modelo se encuentran en estado crítico. Tomado de: Critical Dynamics in Genetic Regulatory Networks: Examples from Four Kingdoms (59).

Estos resultados hacen suponer que los sistemas biológicos estarían operando en punto crítico. Esto le brindaría la flexibilidad suficiente a los organismos como para poder evolucionar y a la vez tener la robustez suficiente para no perder su organización a pesar de las perturbaciones del medio ambiente. Las redes booleanas, pese a su aparente simplicidad, nos han permitido analizar un proceso

tan complejo como la evolución de los organismos en base al estado en el que se encuentran sus redes biológicas.

2.7.2. Modelos booleanos de sistemas biológicos

Los modelos booleanos se han empleado para describir redes de regulación biológicas en diversos organismos modelo. Es importante destacar que, a diferencia de las redes de Kaufman, que presentan una estructura aleatoria, los modelos booleanos requieren datos experimentales para su construcción. Uno de los primeros organismos en donde se utilizaron redes booleanas para explicar su desarrollo fue *A. thaliana* (11). En este estudio, se implementó un modelo booleano estocástico. En este tipo de modelos, se establece una probabilidad η de que los X_n genes de la red no sigan sus funciones booleanas reguladoras (F_n). Por lo tanto, tenemos dos posibilidades: **(1)** un primer caso, donde se obedecen las funciones reguladoras, el estado de los genes se define como sigue:

$$X_n(t+\tau) = F_n(X_n(t)), \text{ con probabilidad } p = (1 - \eta)$$

Donde X_n representa el estado de activación de un gen en el tiempo t y τ representa el incremento discreto del tiempo t ; **(2)** un segundo caso, en el que se desobedecen las funciones reguladoras, el estado de los genes se define como sigue:

$$X_n(t+\tau) = 1 - F_n(t), \text{ con probabilidad } p = \eta.$$

Usando datos experimentales para construir este modelo, se creó una red booleana de 15 genes reguladores del desarrollo floral. Fueron encontrados diez atractores de

la red, de los cuales cuatro corresponden a las cuatro regiones del meristemo de la inflorescencia (I1, I2, I3 e I4), mientras que los otros seis corresponden a los cuatro tipos de células primordiales de órganos florales dentro del meristemo floral (sépalos, pétalos, estambres y carpelos) (11). Ver figura 18.

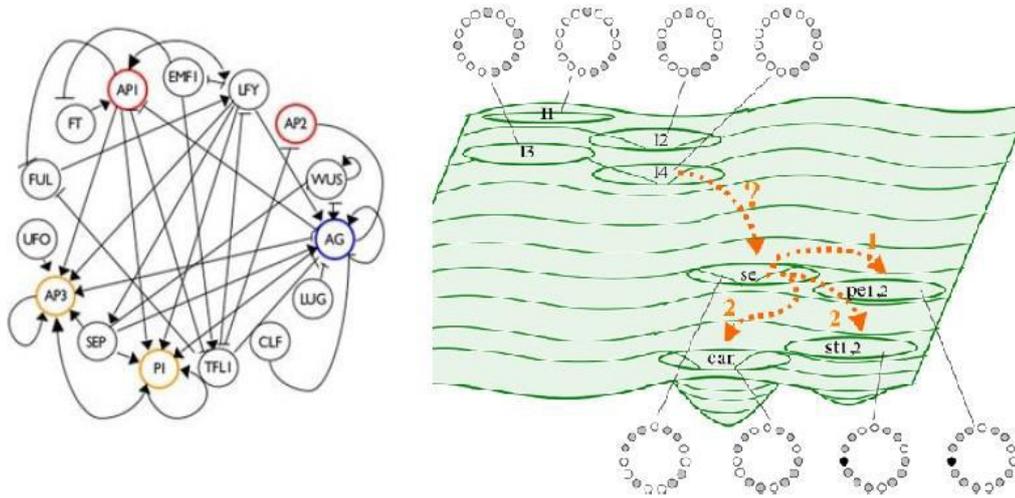


Fig. 18 Red de reguladores con sus correspondientes vasijas de atracción. El modelo permite predecir las transiciones entre distintos atractores, que corresponden a las células de los primordios florales. Tomado de: Stochastic Explorations of a Gene Network Epigenetic Landscape (11).

Otro organismo modelo en el que las redes booleanas se han utilizado para modelar su desarrollo es *D. melanogaster*. En este estudio, el modelo es capaz de reproducir tanto los patrones de expresión genética experimentales de tipos salvaje como los patrones de expresión del tipo mutante. En este modelo, se utilizaron como nodos de la red las proteínas y los RNA mensajeros del organismo en desarrollo (12).

El cáncer es otro sistema que también se está estudiando mediante el uso de redes booleanas. Se ha postulado que las redes genéticas de las células cancerosas se encuentran en atractores atípicos, que surgen debido a mutaciones en nodos de la red genética no cancerosa. Esto abre paso a la terapia de diferenciación, con el objetivo de intentar encausar nuevamente a la red hacia atractores típicos (60).

Las redes booleanas también se han aplicado al estudio de la diferenciación de las células T CD4+. En este estudio, se construyó una red booleana que tenía en cuenta no solo los genes reguladores de la diferenciación de linfocitos, sino también las citoquinas moduladoras de la diferenciación como parte de la red (61). Ver figura 19. Del análisis de esta red se obtuvieron atractores correspondientes a los distintos destinos celulares de las células T CD4+ y una interpretación de su plasticidad. Se identificaron los atractores de los siguientes destinos celulares: Th0, Th1, Th2, Th17, iTreg, Tfh, Th9 y células reguladoras T independientes de Foxp3 (TrFoxp3-) (61). Estos resultados demuestran como las redes booleanas permiten integrar factores de distinta naturaleza en una red, lo que las convierte en una herramienta poderosa y versátil.

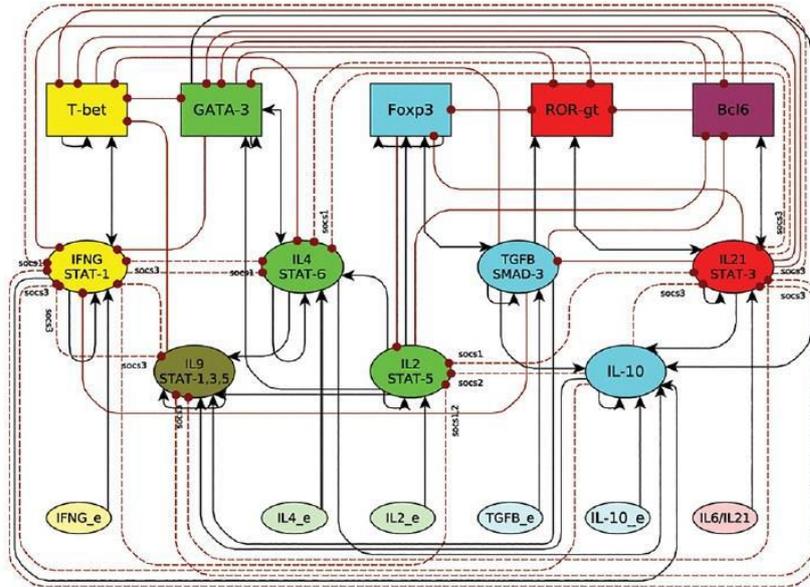


Fig. 19 Red de regulación transcripcional y de señalización de células T CD4+. Tomado de: A Minimal Regulatory Network of Extrinsic and Intrinsic Factors Recovers Observed Patterns of CD4+ T Cell Differentiation and Plasticity (61).

La diferenciación de los macrófagos también ha sido modelada con redes booleanas. Siguiendo la expresión temporal de citoquinas producidas en células de ratones C57BL/6, se determinaron las reglas lógicas de la red de regulación que origina la diferenciación de los tres principales destinos celulares de macrófagos: M1, M2 tipo A y M2 tipo C. En este estudio se encontró que los factores IFN- γ , IL-1 α , Linfotactina, MCP-1, MCP-5 y SCF son fundamentales para la activación de macrófagos tipo M1, mientras que los factores FGF-9, MCP-1, MCP-5, MIP-1 β , MIP-2, SCF y VEGF-A lo son para la activación de macrófagos M2 tipo A (62).

2.7.3. Modelos booleanos de dormancia de MTB

Las redes booleanas también se han utilizado para modelar el estado de dormancia de MTB. El primer modelo booleano de dormancia se construyó utilizando datos de expresión genética recopilados durante la hipoxia. Las conexiones entre los nodos se derivaron de análisis computacionales de interacciones proteína-proteína que se integraron con los datos de expresión genética y las redes de regulación genética (63). Esto permitió generar un módulo con dos grupos de proteínas: un primer grupo de 172 proteínas que interactúan con genes regulados al alza y un segundo grupo de 632 proteínas que interactúan con genes regulados a la baja durante la dormancia (ver figura 20). Mediante el análisis de datos experimentales de expresión genética de MTB en distintos modelos de dormancia, se identificaron dos grupos de genes según su nivel de expresión durante la hipoxia: 50 con expresión aumentada y 34 genes con expresión disminuida, en todos los datos experimentados analizados. De este módulo, se seleccionaron 26 genes que aumentaron su expresión y 13 genes que la disminuyeron para crear el modelo booleano de dormancia. En las simulaciones de trayectorias hacia el atractor, se parte de un estado inicial de la red en el tiempo t con los nodos a evaluar encendidos y el resto de los nodos apagados. Luego, en el tiempo $t+1$, se evalúan las funciones booleanas F_n para obtener el siguiente estado y se continúa hasta llegar a un estado que no cambia en el tiempo, lo que representa un atractor de la red. Al realizar las simulaciones de trayectorias partiendo de nodos individuales activados, no se obtuvieron estados estables que

representaran al sistema. Sin embargo, al iniciar las simulaciones con los factores Rv0081, DevR y DevS en estado activo, se llegó a un estado estable que concordaba bien con los datos experimentales. Esto respalda la hipótesis de que estos tres factores son fundamentales para iniciar la respuesta temprana a la hipoxia de MTB (63).

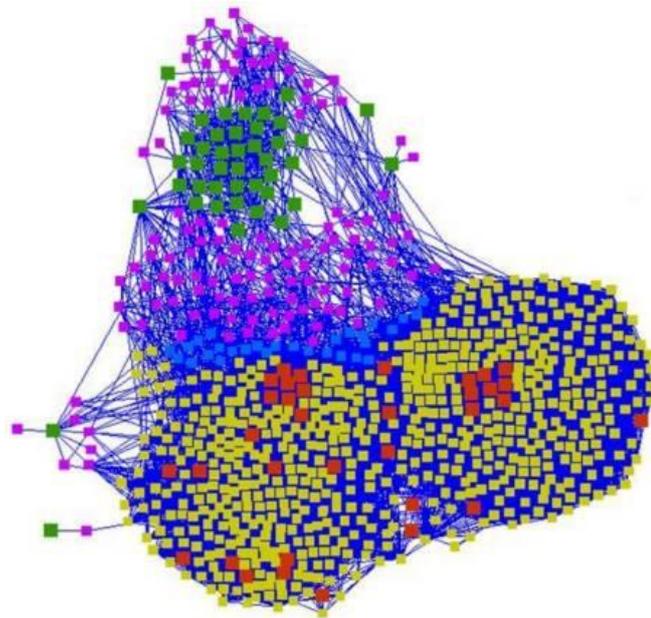


Fig. 20 Agrupamiento de genes de MTB en base a su nivel de expresión durante la hipoxia. En verde se muestran los nodos con expresión aumentada; en rojo se muestran los nodos con expresión disminuida; en púrpura, se muestran los primeros vecinos de los nodos con expresión aumentada; en amarillo, se muestran los primeros vecinos de los nodos de expresión disminuida; en azul, se muestran los nodos que interactúan tanto con nodos de expresión aumentada y disminuida. Tomado de: Understanding Communication Signals during Mycobacterial Latency through Predicted Genome-Wide Protein Interactions and Boolean Modeling (63).

En un segundo modelo de dormancia de MTB, se usaron datos de expresión genética obtenidos durante 7 días de hipoxia. En este modelo, las relaciones entre los nodos se obtuvieron de experimentos de ChipSeq asociados a expresión genética. Al realizar experimentos in silico de mutagénesis puntual se determinó que los nodos: CsoR, Rv0081, Rv2034, Lsr2, PhoP, Rv1985c y Rv0324 son los más críticos en la red (64). Los resultados del modelo booleano se compararon con los datos experimentales de expresión genética durante los siete días de hipoxia como se muestra en la figura 21. El modelo pudo predecir el aumento de expresión de los nodos de hipoxia temprana devR, devS y Rv0081, que posteriormente disminuyen su nivel de expresión durante la hipoxia tardía. De forma similar también predice el comportamiento de los nodos hipoxia tardía como Rv0324, Rv0348, Rv2034 que tienen un bajo nivel de expresión durante la hipoxia temprana y un alto nivel de expresión durante la hipoxia tardía (64).

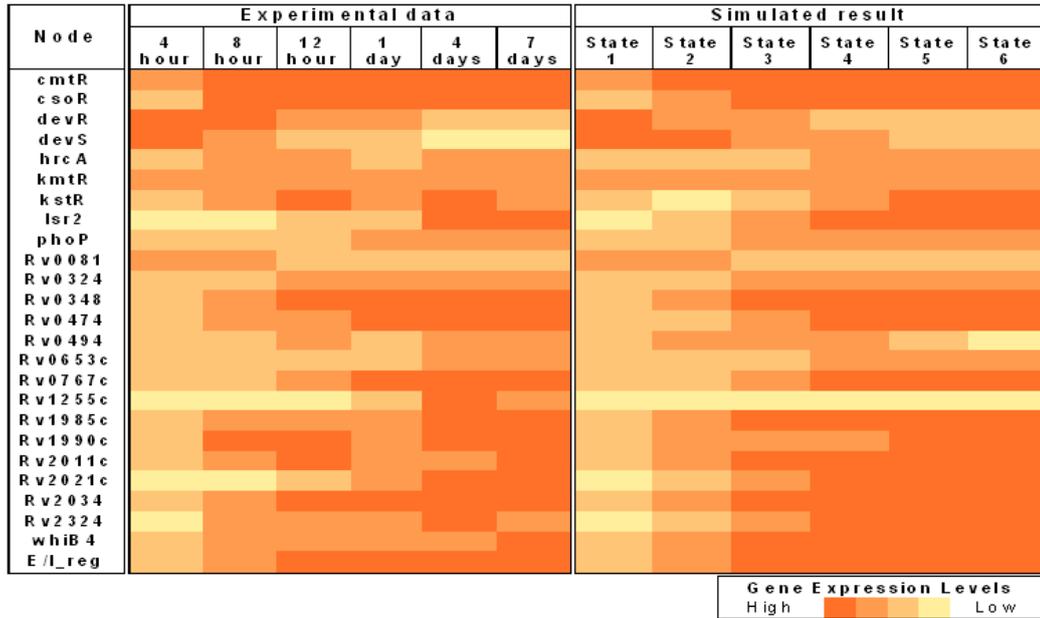


Fig. 21 Comparación de resultados del modelo booleano a los datos experimentales. Los valores de los nodos se agruparon en 4 niveles de expresión genética para 6 estados de hipoxia. Los colores indican los niveles de expresión de los genes de la red, el amarillo indica un bajo nivel de expresión, el rojo un alto nivel de expresión, mientras que las tonalidades naranjas representan un valor intermedio de expresión. A la izquierda se representan los niveles de expresión experimentales para los 25 nodos de la red durante seis estados temporales (4h, 8 h, 12h, 1 día, 4 días y 7 días), mientras que a la derecha se muestran los niveles de expresión predichos por el modelo para los mismos nodos Tomado de: Understanding the role of interactions between host and *Mycobacterium tuberculosis* under hypoxic condition: an *in silico* approach. Genomics (64).

III. JUSTIFICACIÓN

La capacidad de MTB de causar tuberculosis activa a partir de una reactivación desde el estado de latencia mantenido por largo tiempo, incluso décadas, es un serio agravante como problema de salud a nivel mundial (4, 8). Los modelos de redes booleanas han demostrado ser de gran utilidad para simular procesos de interacciones complejas en diversos sistemas biológicos, que abarcan a distintos organismos como *D. melanogaster*, *A. thaliana* y bacterias como *B. subtilis* (12, 11, 59). Dada la complejidad del estadio de dormancia de MTB, de los procesos de su inicio y la reactivación de MTB, se propone un abordaje desde la biología de sistemas para integrar los diversos factores del sistema en una red que nos permita una mejor comprensión de estos procesos (51).

Estudios previos del proceso de dormancia de MTB usando redes booleanas han permitido corroborar la importancia de algunos nodos ‘hub’ importantes como Rv0081, DevR y DevS, que forman parte de la respuesta temprana a la hipoxia y de otros nodos como Rv0324 y Rv1985c, que forman parte de la respuesta tardía a la hipoxia (63, 64). Sin embargo, estos modelos no han demostrado la existencia de atractores que se correspondan con el estadio de reactivación. Por lo tanto, un modelo que genere los atractores de hipoxia temprana, tardía y de la reactivación, basado en datos experimentales de trabajos más recientes como el de Peterson y cola-

boradores (65), que posee datos de expresión durante las etapas de hipoxia temprana, tardía y de reactivación, sería de mucha utilidad para determinar los nodos críticos del sistema y proporcionar candidatos de blancos terapéuticos.

Por otro lado, los estudios publicados solo contemplan factores del patógeno y no tienen algún tipo de vinculación con la respuesta inmune del huésped. Estudios previos han demostrado como la inmunosupresión que se produce durante la co-infección con HIV o el tratamiento inmunosupresor pueden favorecer la reactivación (66, 67). Un modelo que plantee la influencia de factores del sistema inmune del huésped como las citoquinas IFN- γ e IL-10, que han demostrado modificar la expresión del gen HspX de MTB en modelos murinos (68) y que son producto de mecanismos como el censado de IFN- γ por parte de la proteína de membrana MmpL10 de MTB (69), podría contribuir a una mejor comprensión de las redes de interacción entre MTB y su huésped y al estudio de nuevas terapias que integran el uso de antimicrobianos junto a inmunomoduladores para tratar esta patología.

IV. HIPÓTESIS

Una red booleana compuesta por los genes reguladores y efectores del proceso de dormancia de MTB y sus respectivas conectividades, arrojará atractores que representan los estados celulares de hipoxia temprana, hipoxia tardía y reactivación identificados experimentalmente en estudios previos.

V. OBJETIVOS

5.1.OBJETIVO GENERAL

Diseñar una red booleana de los genes involucrados en la regulación del estado de dormancia de MTB que pueda simular los estados de hipoxia temprana, tardía y reactivación previamente caracterizados experimentalmente.

5.2.OBJETIVOS ESPECÍFICOS

- Desarrollar un modelo de redes booleanas para representar el sistema de dormancia de MTB.
- Evaluar la capacidad del modelo de reproducir resultados experimentales en hipoxia temprana, tardía y reactivación, mediante el estudio de los atractores del modelo.
- Expandir el modelo booleano de dormancia de MTB introduciendo nodos relacionados con la acción de las interleuquinas y evaluar sus resultados.
- Expandir el modelo booleano de dormancia de MTB introduciendo nodos relacionados con la reactivación y evaluar sus resultados.
- Realizar simulaciones de mutaciones que inactiven nodos específicos para identificar los nodos críticos de la red booleana de dormancia.

VI. METODOLOGÍA

6.1.OBTENCIÓN DE DATOS DE EXPRESIÓN DIFERENCIAL DURANTE LA HIPOXIA

Se analizaron los datos crudos de lecturas de RNAseq del trabajo de Peterson y colaboradores (65). En este trabajo se tomaron 81 muestras de RNA total de un cultivo de MTB sometido a tres condiciones generales: **(1)** disminución de oxígeno disuelto (OD), desde 81,62 % a 0 % (durante el periodo de 0 a 47 horas); **(2)** mantenimiento en hipoxia (hasta las 114 horas) y **(3)** re-aireación (desde 114 a 120 horas). Las muestras son tomadas por triplicado. Las lecturas de las muestras tienen el identificador GEO: GSE116353. En el Anexo 2 se detallan las condiciones de cada muestra.

6.1.1. Obtención de las lecturas de RNA seq

Para todos los análisis computacionales se usó el lenguaje de programación R en la versión 3.6.3. El código para la obtención de las lecturas de RNA seq, se muestran en el Anexo 6.

a) Usando la plataforma The European Nucleotide Archive, se procedió a obtener los archivos “Fastq1.gz” y “Fastq2.gz”, que contienen las lecturas en dirección 5'→3' y 3'→5', respectivamente, del RNA total de cada una de las 81 muestras.

b) Se crearon las tablas “allFastq₁” y “allFastq₂”, con las lecturas “fastq1” y “fastq2” de las 81 muestras.

6.1.2. Alineamiento de las lecturas al genoma de MTB

a) Se descargó y descomprimió el genoma de MTB usando la base de datos del NCBI: “GCF_000195955.2_ASM19595v2_genomic.fna.gz”. Se procedió a indexar el genoma de MTB.

b) Se generó un banco de datos (data frame) llamado “PROJECT” con todas las lecturas, para poder alinearlas al genoma de MTB.

c) Se generaron los 81 archivos con formato BAM, que contienen las lecturas validadas alineadas a cada región del cromosoma de MTB.

d) Se descargó de la base de datos de NCBI el archivo “sequence.gff3”, que contiene todos los genes de MTB. Luego se generó el archivo “data4”, que contiene todas las lecturas válidas para cada gen de MTB.

Esta tabla "raw_counts.tsv" nos permitió obtener la distribución de lecturas válidas de cada gen de MTB, para cada una de las 81 muestras. Como ejemplo, se muestra la distribución de lecturas para el gen Rv3133c, que es un nodo fundamental de la red de dormancia. Ver figura 22.

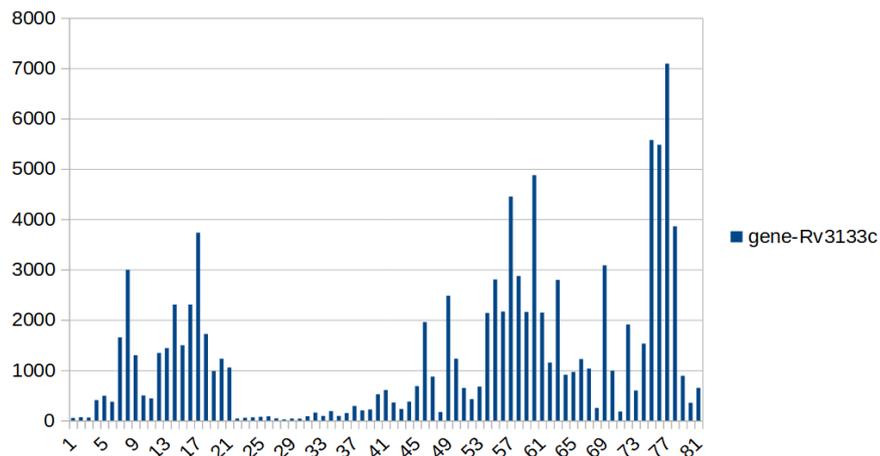


Fig. 22. Lecturas válidas del gen Rv3133c para las 81 muestras. En la parte inferior se muestran los 81 análisis que conforman las 25 muestras generales. En el panel derecho se ubica el nombre del gen evaluado. En el eje Y se encuentra la cantidad de lecturas para cada uno de los 81 análisis realizados.

6.1.3. Verificación del alineamiento de las lecturas al genoma de MTB usando el programa de visualización IGV

Este programa permite cargar todas las lecturas de los 81 archivos BAM, junto con el genoma de MTB, y visualizar la ubicación de las lecturas para cada gen de MTB. De esta manera, se corroboraron los datos de la tabla "raw_counts.tsv". En la figura 23, se observa un aumento significativo de lecturas a medida que disminuye el porcentaje de oxígeno disuelto para el gen Rv3133c (devR). Estas lecturas alcanzan un máximo y luego vuelven a disminuir, lo cual coincide con lo observado en la figura 22, donde se muestran las lecturas de la tabla "raw_counts.tsv" para el gen Rv3133c. Este proceso de comparación se repitió para otros dos genes más de la

red de dormancia: Rv0081y Rv3132c, lo que permitió verificar la correlación entre la cantidad de lecturas según la tabla "raw_counts.tsv" y los alineamientos en IGV para tres muestras: (1) 17 horas (con un 57.6 % de OD); (2) 50 horas (con un 0.0 % de OD) y (3) 112 horas (con un 0.0% de OD).

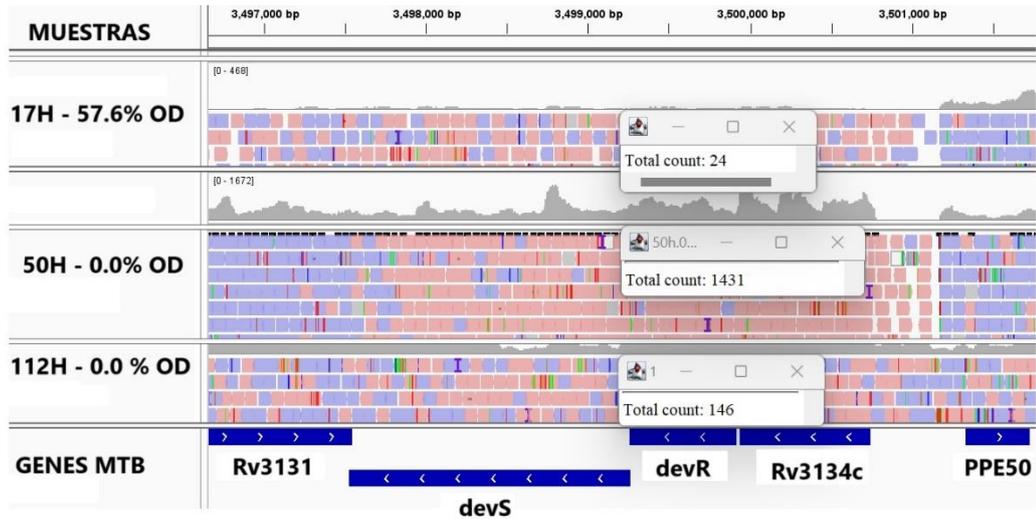


Fig. 23 Comparación de lecturas para las muestras 17h-57.6%, 50h-0.0% y 112h-0.0 % de OD. En la parte inferior de la figura se encuentran los genes de MTB, como el gen Rv3133c (DevR) y Rv3132c (DevS). En el lado izquierdo se observan tres muestras: la primera tomada a las 17 horas (con un 57.6 % de OD), la segunda a las 50 horas (con un 0.0 % de OD) y la tercera a las 112 horas (con un 0% de OD). Las lecturas alineadas a genes orientados en sentido 5' a 3', como PPE50 y Rv3131, se muestran en color azul, mientras que las lecturas alineadas a genes en sentido 3' a 5', como DevR, DevS y Rv3134c, se representan en color naranja. El programa permite el conteo de lecturas alineadas en cada región de cada gen. En particular, se muestran las lecturas alineadas a la parte central del gen DevR: 24 lecturas en la muestra de 17 horas (con un 57.6 % de OD); 1430 lecturas en la muestra de 50 horas (con un 0.0 % de OD) y 146 lecturas en la muestra de 112 horas (con un 0.0% de OD). Estos valores coinciden con los encontrados en la tabla "raw_counts.tsv", lo que corrobora la precisión del proceso bioinformático de alineamiento de las lecturas al genoma.

6.1.4. Obtención de la expresión diferencial para todos los genes de MTB, respecto de la muestra de completa oxigenación (81.62 % de OD).

a) Los datos fueron analizados mediante el algoritmo DESeq (70). Este algoritmo nos permite obtener los valores de expresión diferencial respecto a la muestra con 81.62% de OD. Las 81 muestras, en su mayoría, son análisis por triplicado, a excepción de la muestra de 54 horas que es por sextuplicado, constituyendo en total 25 condiciones diferentes a la inicial de 81.62% de OD. Ver tabla 2.

Muestra N°	Hora	% Oxígeno	Etiqueta
1	17	57.6	17h.57.6
2	23	42.8	23h.42.8
3	40	11.5	40h.11.5
4	41	8.8	41h.8.8
5	42	7.49	42h.7.49
6	43	6.16	43h.6.16
7	44	4.31	44h.4.31
8	45	2.93	45h.2.93
9	46	0.41	46h.0.41
10	47	0	47h.0.0
11	48	0	48h.0.0
12	49	0	49h.0.0
13	50	0	50h.0.0
14	52	0	52h.0.0
15	54	0	54h.0.0
16	64	0	64h.0.0
17	70	0	70h.0.0
18	88	0	88h.0.0
19	94	0	94h.0.0
20	112	0	112h.0.0
21	114	0.45	114h.0.45
22	114.5	3.84	114.5h.3.84
23	115	6	115h.6.0
24	116	10.89	116h.10.89
25	120	46.92	120h.46.92

Tabla 2 Condiciones generales de las muestras. A la derecha se encuentra la hora, a la izquierda el porcentaje de oxígeno disuelto para cada condición y en la última columna se muestran las etiquetas con las que está rotulada cada condición experimental.

El código en R para el análisis de expresión diferencial entre la muestra de condición 17 horas (con 57.62 % OD) y la muestra de 0 horas (con 81.6 % OD) se muestra en el Anexo 6.

b) Este código fue utilizado para cada una de las distintas 25 condiciones y se obtuvo la expresión diferencial para todos los genes de MTB. Todos los resultados se agruparon en una tabla conteniendo los valores de expresión diferencial para cada gen en cada una de las 25 condiciones experimentales (estados_hipoxia.csv). Estos valores son expresados en unidades de “log2ratio”, es decir, expresan el logaritmo en base dos del cociente entre las lecturas a una condición y la condición inicial. Ver Anexo 2.

6.2. BINARIZACIÓN DE LOS NIVELES DE EXPRESIÓN DIFERENCIAL

Para analizar los cambios de expresión genética mediante el uso de una red booleana, es necesario binarizar los datos de expresión diferencial, es decir, asignar un valor de 1 cuando exista un aumento de la expresión diferencial y un valor de 0 cuando exista una disminución. Por lo tanto, debemos agrupar los datos de expresión diferencial en dos categorías, de modo que todos los N datos se dividan en 2 grupos. En general, dividir un conjunto de N datos en k grupos es un problema

complejo, con $k^N/k!$ soluciones posibles. Esto constituye un problema computacionalmente posible solo para valores pequeños de N. Para abordar este problema, se han desarrollado diversas heurísticas. Una de ellas es el algoritmo de agrupamiento k-means, también denominado algoritmo de Lloyd. Este algoritmo permite agrupar los N datos en k conjuntos. Cada punto se asigna al centro más cercano y luego se recalcula cada centro como la media de todos los puntos asignados a él. Estos dos pasos se repiten hasta que se cumple un criterio de finalización predefinido, que generalmente implica minimizar la suma de las distancias de Euclides al centro (71). Para obtener este agrupamiento con k=2, se siguió el siguiente procedimiento:

a) Se procedió a guardar en un vector, con el nombre de cada gen de interés, todos los datos de expresión diferencial sin procesar para cada una de las muestras, obtenidos del archivo “estados_hipoxia.csv”.

En el Anexo 3 se muestra el código en R.

b) Se generó una matriz “serie_tiempo” con todos los vectores de cada nodo.

c) Se binarizaron los datos de expresión diferencial y se guardaron en el archivo "serie_tiempo_25.tsv".

6.3.OBTENCIÓN DEL AGRUPAMIENTO DE LOS ESTADOS DE HIPOXIA.

Para obtener los estados de hipoxia de MTB, en primer lugar, se cargó el archivo `estadod_hipoxia.csv`, que contiene los valores de expresión diferencial para las 25 condiciones experimentales y se realizó el agrupamiento mediante un análisis de componentes principales (72). En el anexo 6 se muestra código en R.

6.4.OBTENCIÓN DE LA RED DE DORMANCIA DE MTB

Para obtener la red de dormancia de MTB, se partió de la red previa inicial del trabajo de Bose y colaboradores (64), sin reguladores externos o internos. Esta red consta de 24 genes con las siguientes funciones booleanas:

GEN	FUNCIONES BOOLEANAS INICIALES
CmtR (Rv1994c)	(Lsr2 OR Rv0081 OR Rv0767c OR CmtR) AND (NOT Rv0324)
CsoR (Rv0967)	(Rv0081 OR CsoR) AND (NOT Rv2324)
DevR (Rv3133c)	DevS OR (PhoP AND Rv2034) OR (DevR AND (DevS OR (PhoP AND Rv2034)))
DevS (Rv3132c)	DevR
HrcA (Rv2374c)	Rv0324
KmtR (Rv0827c)	KmtR AND (NOT Rv2034)
KstR (Rv3574)	(Rv0348 OR Rv2324 OR KstR OR (Lsr2 AND Rv0324)) AND (NOT Rv0081)
Lsr2 (Rv3597c)	(Rv1985 OR CmtR OR Rv2034 OR KstR OR Lsr2) AND (NOT(PhoP OR Rv0324 OR Rv0474))
PhoP (Rv0757)	(WhiB4 OR Rv0081 OR (Lsr2 AND Rv2034)) AND (NOT Rv2011)
Rv0081	DevR AND (NOT (Rv1990c OR Rv0081))
Rv0324	(Rv0324 OR (Lsr2 AND Rv1985c)) AND (NOT Rv0081)
Rv0348	Rv0081 AND (NOT (Rv0324 OR Rv1985c OR Rv2034))
Rv0474	(Rv1985c OR Rv0081 OR Rv0474) AND (NOT CmtR)
Rv0494	(Rv0081 OR Rv0494) AND (NOT Rv0767c)
Rv0653c	Lsr2 OR Rv0081 OR Rv0653c
Rv1255c	NOT Rv1990c
Rv1985c	(Rv0081 OR Rv1985c) AND (NOT (DevR OR CmtR OR Rv0324))
Rv1990c	Lsr2 OR Rv0081 OR Rv1990c
Rv2011c	Rv2011c
Rv2021c	Rv0081
Rv2034	Rv0081 AND (NOT DevR)
WhiB4 (Rv3681c)	Rv0081 AND (Not Rv2011c)

Tabla 3 Funciones booleanas de la red de dormancia inicial con 24 nodos.

En la columna de la izquierda se muestran los nodos de la red y en la columna de la derecha sus funciones reguladoras. Los nodos Rv0767c y Rv2324 no son regulados por otros nodos de la red, por lo tanto, solo figuran en la columna de la izquierda. Tomado de: Understanding the role of interactions between host and *Mycobacterium tuberculosis* under hypoxic condition: an in silico approach. Genomics (64).

Para visualizar la red se crea un archivo redmtbe.24.txt con las funciones booleanas y se procede a obtener su gráfica. En el anexo 6 se muestra código en R.

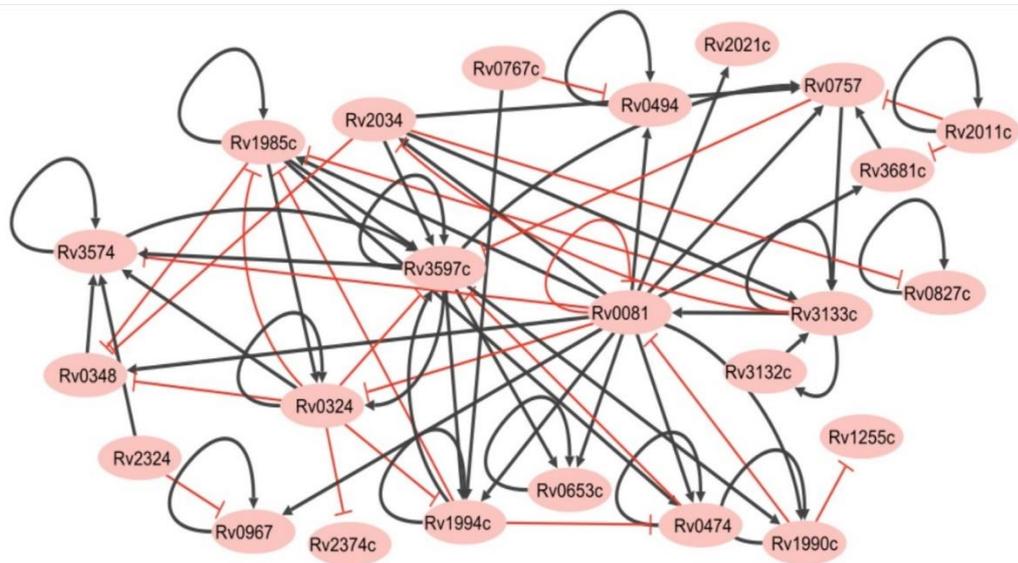


Fig. 24 Estructura de la red de dormancia de 24 nodos inicial. Las aristas negras representan las interacciones de activación, mientras que las aristas rojas con extremos en forma de “T” indican las interacciones inhibitorias.

6.4.1. Verificación y corrección de las funciones booleanas de la red previa de 24 nodos

Para la verificación o corrección, según fuera necesario, de las funciones booleanas de la red previa inicial del trabajo de Bose y colaboradores (64), se utilizaron los resultados experimentales del trabajo de Turkarslan y colaboradores (73), disponibles en la plataforma “MTB Network Portal v2” y los resultados de Galagan y colaboradores (34). El primer trabajo contiene una base de datos con la información de los genes que son activados o inhibidos por los factores de transcripción de MTB, con sus correspondientes valores de *p-value*. El *p-value* es un valor estadístico que indica la probabilidad de obtener resultados tan extremos o más extremos que los observados, si la hipótesis nula (que no hay efecto o diferencia real) es cierta. En este contexto, los valores de *p-value* asociados con la activación o inhibición de

genes por factores de transcripción de MTB indican la significancia estadística de estas interacciones.

Se aplicó el criterio de que cada *p-value* debe ser menor a 0.05 para considerar significativo el efecto del FT en cuestión y, en consecuencia, incluir dicho FT como parte de la función booleana del gen que regula.

Al revisar esta base de datos, se encontraron FTs que no cumplían con el criterio de tener de *p-value* < 0.05. A continuación, en la tabla 4 se muestran los nodos y sus reguladores con sus valores de *p-value* que no cumplen con el criterio de inclusión.

GEN	REGULADORES	p-value
Rv0474	Rv1985c	0.9
	Rv1994c (cmtR)	0.94
Rv1994c (cmtR)	Lsr2(3597c)	0.81
	Rv0767c	0.5
Rv0967	Rv2324	0.94
Rv2374c (hrcA)	Rv0324	0.58
Rv3574 (KstR)	Rv2324	0.92
	Rv3597c	0.6
	Rv0348	0.38
Rv0348	Rv1985c	0.35
	Rv2034	0.9
	Rv0324	0.93
Rv0653c	Rv3597c Lsr2)	0.99
	Rv0081	0.28
Rv1985c	Rv3133c (DevR)	0.75
Rv1990c	Rv3597c (Lsr2)	0.81
Rv2034	Rv3133c (DevR)	0.75
Rv3597c (Lsr2)	Rv1985	0.53
	Rv2034	0.5
	Rv3574 (KstR)	0.74
	Rv0324	0.27
	Rv0474	0.97
Rv0757	Rv0081	0.34
	Rv3597c (Lsr2)	0.74
Rv0081	Rv1990c	0.77
	Rv3597c	0.27

Tabla 4 Nodos reguladores con un p -value > 0.05. En la columna de la izquierda se muestran los genes de la red previa inicial del trabajo de Bose y colaboradores (64), en la columna del centro se muestran sus nodos reguladores, que en principio no serían significativos por tener un valor de p -value > 0.05. Es importante tener en cuenta que un valor de p -value > 0.05 no siempre indica que los nodos no sean relevantes, ya que esto depende del contexto y de la interpretación específica del análisis estadístico realizado.

Una vez verificada la presencia de funciones reguladoras con un p-value mayor a 0.05 en la base de datos de Turkarslan y colaboradores (73), se procedió a verificar todas las funciones booleanas con la segunda base de datos de Galagan y colaboradores (34). En el caso de funciones reguladoras con un p-value > 0.05 o que solo se encontraron en una de las bases de datos, se procedió a corroborar dichas funciones mediante el análisis de expresión diferencial entre el nodo regulador y su nodo regulado, utilizando los datos experimentales obtenidos en la sección 7.1.4. Si se encontraba una correlación positiva entre los nodos analizados, se confirmaba una función activadora, mientras que una correlación negativa confirmaba una función inhibidora tipo “NOT”. Si el análisis de expresión diferencial no confirmaba la función reguladora, se procedió a descartar dicha función. Se consideró que una interacción era insignificante si la correlación de expresión entre los dos miembros estaba entre -0.3 y $+0.3$, y significativa en caso contrario (63).

Las funciones booleanas corregidas son las siguientes:

GEN	FUNCIONES BOOLEANAS CORREGIDAS
Rv1994c (CmtR)	(Rv0081 OR CmtR) AND (NOT Rv0324)
CsoR (Rv0967)	(Rv3133c OR Rv2034)
DevR (Rv3133c)	DevS OR (PhoP AND Rv2034) OR (DevR AND (DevS OR (PhoP AND Rv2034)))
DevS (Rv3132c)	DevR OR Rv2034
KmtR (Rv0827c)	KmtR AND (NOT Rv2034)
KstR (Rv3574)	(KstR OR Rv0324) AND (NOT Rv0081 OR Rv3597c)
Lsr2 (Rv3597c)	(CmtR OR Rv3597c OR Rv2034 OR Rv0474 OR Rv3574) AND (NOT(Rv0324))
PhoP (Rv0757)	(WhiB4 OR Rv0081 OR PhoP OR Rv3597c) AND (NOT Rv2011)
Rv0081	(NOT Rv3574 OR DevR)) AND (NOT (Rv0081))
Rv0324	(Rv0324) OR (NOT (Rv3597c AND Rv1985c)) AND (NOT Rv0081)
Rv0348	Rv0081 AND NOT (Rv0324 OR Rv1985c)
Rv0474	Rv0081 OR Rv0474 AND (NOT Rv1985c)
Rv0494	(Rv0494) AND (NOT Rv0081)
Rv0653c	Rv0653c AND NOT (Rv3597c OR Rv0081)
Rv1255c	Rv1255c AND NOT Rv1990c
Rv1985c	Rv0324 OR Rv0081 OR (Rv1985c) AND (NOT (CmtR OR Rv3133c OR Rv3597c))
Rv1990c	(Rv3597c OR Rv0081 OR Rv1990c) AND (NOT Rv3574)
Rv2011c	Rv2011c OR Rv0324
Rv2021c	Rv0081 OR Rv3597c OR Rv3133c
Rv2034	Rv0081 OR Rv2034 OR CsoR
Rv2324	Rv3597c OR Rv0081
Rv3681c (WhiB4)	Rv3597c OR Rv0081 AND (NOT Rv2011c)

Tabla 5 Funciones booleanas corregidas de la red de dormancia inicial propuesta por Bose y colaboradores con 24 nodos. Después de corregir las funciones reguladoras se eliminaron los nodos Rv0767c y Rv2374c debido a que son nodos de salida y sus funciones reguladoras tienen un valor de *p-value* > 0.05. La función “AND” implica que es necesaria la presencia dos nodos reguladores para producir un efecto en el nodo regulado, la función “NOT” representa un efecto inhibitorio y la función “OR” indica que basta la presencia de uno de los dos reguladores para producir el efecto sobre el nodo regulado.

Se procedió a cargar las nuevas funciones booleanas y a obtener la nueva estructura, mediante el uso del mismo código usado para la red inicial del trabajo de Bose y colaboradores (64). Ver figura 25.

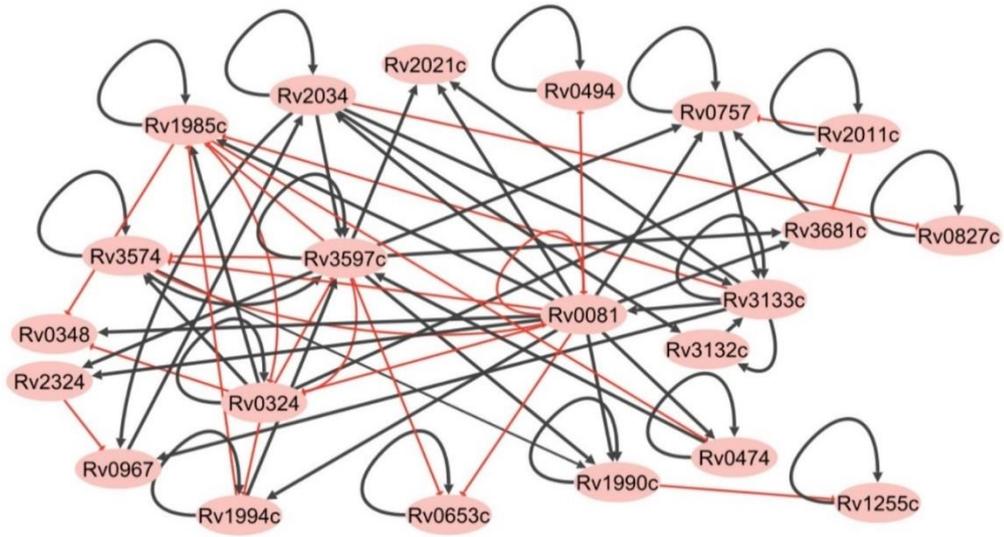


Fig. 25 Estructura de la red de dormancia de 22 nodos corregida.

6.4.2. Reducción del número de nodos de la Red

Con el objetivo de analizar exhaustivamente el comportamiento de los atractores que representan los estados experimentales de MTB, se procede a reducir el tamaño de la red sin alterar sus propiedades fundamentales, tomando en cuenta las siguientes pautas:

- 1) El nodo debe ser un nodo de salida de la red, dado que estos no afectan la dinámica de la red, por lo que pueden ser retirados para visualizar claramente los atractores de la red (74).

2) Además, el nodo en cuestión no debe ser parte de las subredes críticas para la adaptación a la hipoxia, como la degradación y biosíntesis de lípidos, o la degradación de proteínas, tal como se muestra en la figura 8.

6.4.3. Obtención de los atractores de la red booleana de dormancia

Se carga el archivo de texto con las funciones booleanas de la red. En el anexo 6 se muestra código en R.

6.4.4. Obtención de funciones reguladoras *in silico*

Para verificar las reglas lógicas de las funciones reguladoras de cada nodo en la red, se procede a generar funciones reguladoras *in silico* utilizando los datos binarizados de expresión diferencial obtenidos en el punto **7.1.4.**, mediante el uso del algoritmo best-fit. El algoritmo best-fit es un método comúnmente utilizado en bioinformática para encontrar el mejor ajuste entre un conjunto de datos y un modelo (24).

Una vez obtenidos los datos binarizados se procede a obtener las funciones reguladoras *in silico*. En el anexo 6 se muestra código en R.

Así, se obtuvo un conjunto de funciones reguladoras *in silico* con un determinado error para cada nodo.

Estas funciones *in silico* se comparan a las funciones lógicas obtenidas en la sección 7.4.1 de la red de dormancia. En el caso que la función reguladora *in silico*

corrobore la regla lógica evaluada, reforzaría la elección de dicha función reguladora.

6.5. ANÁLISIS DE MUTACIONES EN LA RED

6.5.1. Verificación de importancia de los nodos en la estructura general

Para este análisis se procede a retirar un nodo de la red y luego se evalúan sus atractores. Si se mantiene la estructura del atractor en ausencia de un nodo, se concluye que ese nodo no es esencial en la estructura general de la red.

6.5.2. Análisis de la trayectoria hacia el atractor

Se analiza cuál es la distancia hacia un atractor de hipoxia desde un estado en el que solo se encuentra activo un gen en particular. En el anexo 6 se muestra el código en R para llegar al atractor desde un estado con solo el gen Rv3133c activo.

6.5.3. Mutaciones de activación

Para analizar el efecto de activar un gen durante toda la simulación, se fija en 1 el estado del nodo durante toda la simulación y se analiza el cambio en los atractores de la red. En el anexo 6 se muestra el código en R para la simulación donde el primer gen de red se encuentra activo permanentemente durante la simulación.

6.5.4. Verificación de mutaciones de inactivación

Para analizar el efecto de inactivar un gen durante toda la simulación, se fija en 0 el estado del nodo durante toda la simulación y se analiza el cambio en los atractores de la red. En el anexo 6 se muestra el código en R para la simulación donde el gen Rv0081 se encuentra inactivo.

6.6. INTRODUCCIÓN DE NODO RELACIONADO CON LAS CITOQUINAS

6.6.1. Búsqueda de genes de MTB relacionados con las Citoquinas

Se realizó una búsqueda bibliográfica sistemática de factores de MTB relacionados al perfil de citoquinas del huésped. y se encontró que la expresión del gen HspX (Rv2031c) está relacionada con la concentración relativa de IFN- γ y la IL 10, que son citoquinas con funciones inflamatoria y antiinflamatoria, respectivamente (68)

6.6.2. Análisis de la interacción del nodo relacionado a citoquinas con la red de dormancia

Se realizó una búsqueda en la bibliografía para encontrar interacciones del gen Rv2031c con los genes reguladores de la red de dormancia. Se encontró una relación de activación por parte del gen Rv3133c en experimentos de microarray y modelos de dormancia inducida por hipoxia (73), (75, 76).

6.6.3. Análisis bioinformático de la relación del nodo relacionado a las citoquinas con la red de dormancia

Se realizó un análisis de componentes principales incluyendo al gen Rv2031c y se generó las funciones booleanas *in silico* del nodo Rv2031c mediante el uso de su perfil de expresión diferencial usando el código visto en la sección **7.4.4**.

Se cargaron los perfiles de expresión del nodo Rv2031c en la red mínima de 19 nodos. El algoritmo “bestfit” se usó con $K=2$ (funciones con un máximo de 2 genes reguladores).

6.6.4. Análisis de los atractores de la red ampliada con el nodo Rv2031c.

Para verificar el comportamiento de la red mínima con el nodo Rv2031c, se generan los atractores de la red y se verifica si se ajustan a los datos experimentales.

6.7.COMPARACIÓN DE LA RED DE DORMANCIA CON REDES ALEATORIAS DEL MISMO GRADO DE ENTRADA PARA LOS NODOS.

Para verificar que los atractores generados en la red de dormancia se obtienen únicamente como producto de la arquitectura de la red, se crearon redes aleatorias con el mismo número de nodos y con exactamente el mismo número de nodos de entrada (k), para determinar el grado de similitud entre los atractores generados por estas redes y la red de dormancia de MTB.

Las funciones booleanas de la red mínima de dormancia se muestran en la tabla 6.

GEN	FUNCIONES BOOLEANAS
Rv1994c (CmtR)	(Rv0081 Rv1994c) & (!Rv0324)
CsoR (Rv0967)	Rv3133c & !(Rv1033c) Rv2034
DevR (Rv3133c)	Rv3132c (Rv0757 & Rv2034) (Rv3133c & (Rv3132c (Rv0757 & Rv2034)))
DevS (Rv3132c)	Rv3133c Rv2034
KstR (Rv3574)	(Rv3574 Rv0324) & !(Rv0081 Rv3597c)
Lsr2 (Rv3597c)	(Rv1994c Rv2034 Rv3574 Rv3597c Rv0474) & !(Rv0324)
PhoP (Rv0757)	(Rv3681c Rv0081 Rv0757 Rv3597c) & (! Rv2011c)
Rv0081	!(Rv3574) Rv3133c) & (!Rv0081)
Rv0324	(Rv0324 !(Rv3597c & Rv1985c)) & (!Rv0081)
Rv0474	Rv0081 Rv0474 & (!Rv1985c)
Rv0494	(Rv0494) & (! Rv0081)
Rv0653c	(Rv0653c & !(Rv3597c Rv0081))
Rv1985c	Rv0324 Rv0081 (Rv1985c) & (! (Rv1994c Rv3133c Rv3597c))
Rv1990c	(Rv3597c Rv0081 Rv1990c) & ! (Rv3574)
Rv2011c	(Rv2011c Rv0324)
Rv2021c	Rv0081 Rv3597c Rv3133c
Rv2034	Rv0081 Rv2034 Rv0967
Rv2324	Rv3597c Rv0081
WhiB4 (Rv3681c)	Rv3597c (Rv0081 & !Rv2011c)
Rv2031c	Rv3133c

Tabla 6 Funciones booleanas corregidas de la red de dormancia con 20 nodos. El símbolo “|” representa la función booleana “OR”, el símbolo “&” representa la función booleana “AND”, y el símbolo “!” representa la función booleana “NOT”.

Se genera una tabla con el número k de reguladores para cada nodo de la red mínima de dormancia. Ver tabla 7.

Nodo	GENE	K
1	Rv3133c	4
2	Rv3132c	2
3	Rv3574	4
4	Rv3597c	6
5	Rv0757	5
6	Rv0081	3
7	Rv0324	4
8	Rv2034	3
9	Rv0474	3
10	Rv1990c	4
11	Rv0967	2
12	Rv1994c	3
13	Rv2021c	3
14	Rv3681c	3
15	Rv2324	2
16	Rv0494	2
17	Rv0653c	3
18	Rv1985c	6
19	Rv2011c	2
20	Rv2031c	1

Tabla 7 Valores de k para cada nodo de la red mínima de dormancia. Se puede apreciar que los nodos Rv3597c y Rv1985c son los que tienen la mayor cantidad de reguladores, con 6 nodos reguladores para cada uno.

Se crean redes aleatorias indicando el número de reguladores para cada nodo. En el anexo 6 se muestra el código en R.

6.8.INTRODUCCIÓN DE NODOS RELACIONADOS AL ESTADO DE REACTIVACIÓN.

6.8.1. Búsqueda de genes de MTB relacionados con la reactivación.

Se llevó a cabo una búsqueda bibliográfica sistemática para identificar los genes asociados al proceso de reactivación. Esta búsqueda reveló que además de los factores promotores de resucitación RpfA, RpfB, RpfC, RpfD y RpfE de MTB que desempeñan un papel fundamental en dicho proceso otros factores importantes son los genes FtsZ y ripA (77, 78, 79, 80).

6.8.2. Análisis de las interacciones de los nodos relacionados a la reactivación con la red de dormancia de MTB.

Se determinaron las funciones reguladoras los nodos relacionados a la reactivación utilizando las bases de datos de regulación genética de Turkarslan y colaboradores (74) y de Galagan y colaboradores (34). Estas funciones fueron verificadas usando los datos experimentales obtenidos en la sección 7.1.4, siguiendo el procedimiento descrito en la sección 7.4.1.

6.8.3. Análisis de los atractores de la red ampliada con los nodos Rpf dormancia

Para investigar el comportamiento de la red de dormancia tras la incorporación de los nodos Rpf, se generaron los atractores de la red ampliada y se verificó su correspondencia con los datos obtenidos experimentales en la sección 7.1.4.

VII. RESULTADOS

7.1. EXPRESIÓN DIFERENCIAL DURANTE LA HIPOXIA

Como se describió en la sección 7.1.4 el algoritmo DESeq permite comparar la expresión genética inicial al 81.62 % de oxígeno, respecto de la expresión en las otras 25 condiciones experimentales para todos los genes de MTB. Estos resultados se encuentran en el Anexo 2. A partir de estos resultados se generan gráficos para cada gen en particular que permiten visualizar su perfil de expresión a lo largo de las 25 muestras analizadas. En la figura 26 se muestra como ejemplo la expresión diferencial para nueve genes centrales de la red genética de dormancia.

Con el fin de verificar la significancia estadística de las lecturas de RNAseq de las 25 muestras, se calculó el coeficiente de correlación (COEF.COR) para cada triplicado de muestras (M1, M2 y M3). De esta manera, se obtuvieron los valores de COEF.COR entre las muestras M1 y M2, entre las muestras M2 y M3, y entre las muestras M1 y M3 para los 25 triplicados de muestras. El promedio de todos estos valores fue de 0.956179 y su Desviación Estándar fue de 0.087651, lo que demuestra una buena reproducibilidad entre las muestras. En el anexo 2 se muestra una gráfica con el triplicado de lecturas para la primera muestra y sus valores de los coeficientes de correlación.

Como se puede observar en la figura 26 hay genes como Rv0324 con valores positivos y negativos de expresión diferencial que intuitivamente se pueden asignar a

estados activos e inactivos del gen respectivamente, pero hay otros genes como Rv3133c donde casi la totalidad de sus estados de expresión diferencial son positivos, lo que dificulta la asignación de un estado activo e inactivo. Para salvar esta dificultad y poder captar la dinámica de activación e inactivación de cada gen de la red de dormancia, el valor límite de corte de expresión diferencial (threshold), se determinó con el algoritmo de binarización descrito en la sección 7.2.

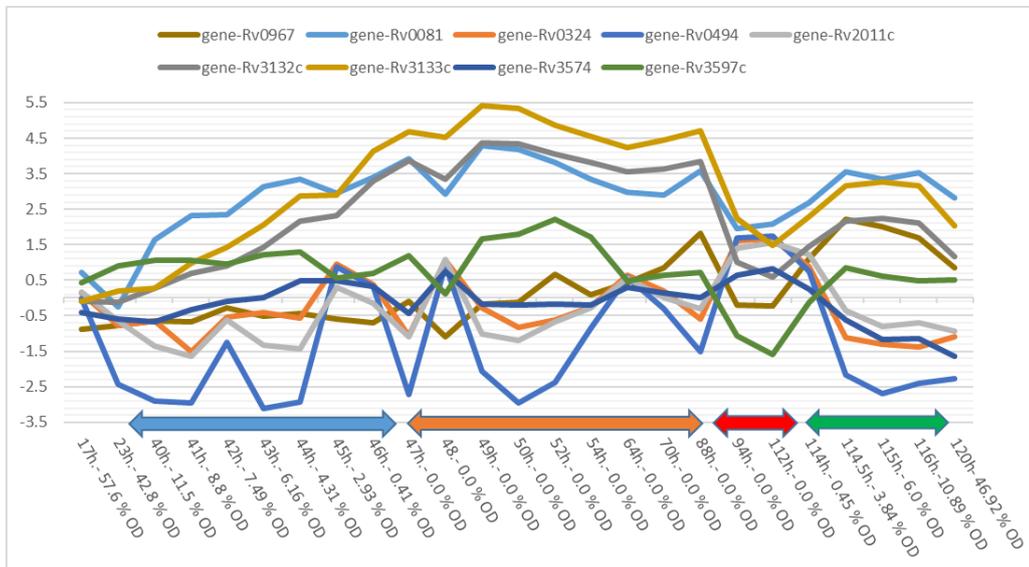


Fig. 26 Comparación de expresión diferencial para nueve genes de dormancia centrales. En la parte superior de la figura se muestran los nombres de nueve genes centrales de la red de dormancia de MTB. En el eje X se representan las 25 condiciones de tiempo y porcentaje de oxígeno disuelto (OD) para cada muestra analizada. La flecha azul representa el período de oxigenación normal, la flecha naranja representa la zona de hipoxia temprana, la flecha roja indica la zona de hipoxia tardía, mientras que la flecha verde indica la zona de reoxigenación. En el eje Y se representa el nivel de expresión diferencial de cada uno de los genes (Rv0967, Rv0081, Rv0324, Rv0494, Rv2011c, Rv3132c, Rv3133c, Rv3574, Rv3597c) expresados en valores de “log₂ratio”, a lo largo de las 25 muestras analizadas. Se pueden evidenciar los diferentes

patrones dinámicos de expresión. Por ejemplo, los genes Rv0081, Rv3132c y Rv3133c muestran un alto nivel de expresión durante la hipoxia temprana, que disminuye durante la hipoxia tardía y vuelve a aumentar durante la reoxigenación. Mientras que, genes como Rv0324, Rv3574 y Rv2011c presentan perfiles de expresión opuestos.

Para mayor detalle se muestra el gráfico de expresión diferencial para el gen Rv0081 en valores de “log2ratio”, que se calcula como el logaritmo en base dos del cociente entre las lecturas de las dos condiciones analizadas y tiene la ventaja de uniformizar los valores de expresión diferencial para los distintos genes y de permitir una mejor interpretación, dado que un valor de 1 implica una duplicación del nivel de expresión en la segunda condición, mientras que un valor de -0.5 que se disminuyó a la mitad la expresión del gen en la segunda condición. Ver figura 27.

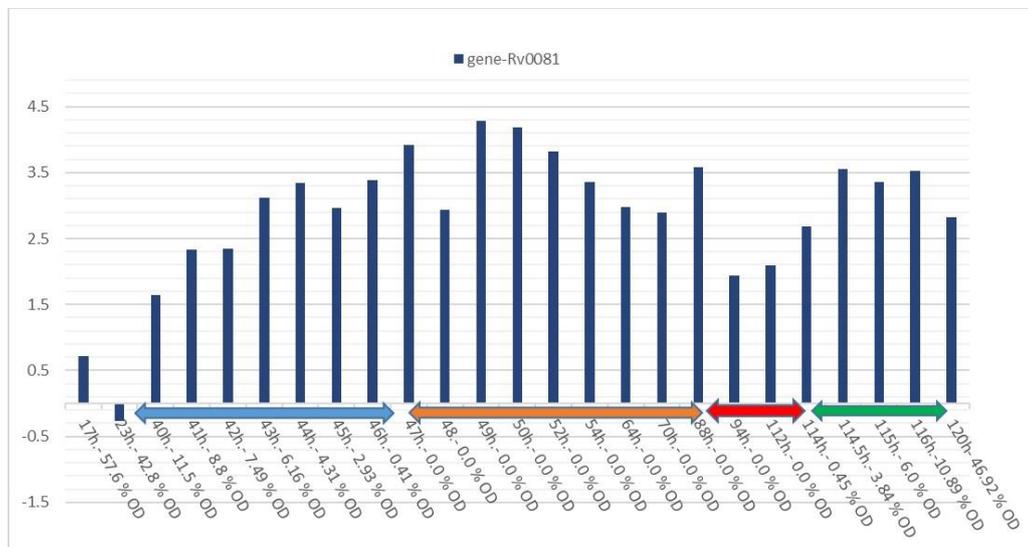


Fig. 27 Expresión diferencial para el gen Rv0081. En el eje X se muestran las 25 muestras analizadas. En el eje Y se muestra el valor de expresión diferencial expresada en valores de “log2ratio”. La flecha naranja representa la zona de hipoxia temprana, la flecha roja indica la zona de hipoxia tardía, mientras que la flecha verde indica la zona de reoxigenación. En la parte inferior del eje X se describen las 25 condiciones de tiempo y porcentaje de oxígeno disuelto (OD) para cada muestra analizada.

7.2. BINARIZACIÓN DE LOS NIVELES DE EXPRESIÓN DIFERENCIAL

En la tabla 8 se muestran los resultados binarizados de expresión diferencial para todos los nodos de la red inicial.

GENES	MUESTRAS																									Umbrales
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
Rv3133	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1.071882336
Rv3132c	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	1.051298420
Rv0081	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	2.357482073
Rv0474	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	0.012028063
Rv0757	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0.178179047
Rv1990c	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0.779452129

Tabla 8 Binarización de la expresión diferencial de 19 genes de la red de dormancia. En el lado derecho se encuentran los genes de dormancia de MTB. Las 25 muestras representan los estados experimentales descritos en la tabla 2. Los valores de expresión diferencial se binarizaron empleando el algoritmo descrito en la sección 7.2. El resultado se representa con un uno si el gen está activo o con un cero si está inactivo, conforme a los criterios definidos por el algoritmo. Los valores de umbral (thresholds) describen el valor límite de “log2ratio” utilizado por el algoritmo para distinguir entre un estado activo y uno inactivo.

7.3. OBTENCIÓN DE AGRUPAMIENTO DE LOS ESTADOS DE HIPOXIA

El resultado del Análisis de Componentes Principales (ACP) para las 25 condiciones experimentales muestra que los dos primeros componentes principales (Dimensión Importante 1 y 2) en la gráfica 27 representan el 92.16 % de la varianza de las muestras. Por lo tanto, al estudiar solo estas dos componentes, se puede evaluar el grado de correlación entre las 25 condiciones experimentales. Para analizar la correlación entre estas 25 condiciones, se grafica el círculo de correlaciones. En este tipo de gráficos, si los vectores se ubican a 180°, implica una correlación negativa; mientras que si tienen la misma orientación, significa que tienen una correlación positiva.

En la figura 28, se observa una correlación negativa entre dos condiciones del estado de normoxia (X23h.42.8 y X17h.57.6) con una condición del estado de reactivación (X114h.0.45), y dos condiciones del estado de hipoxia tardía (X94h.0.0 y X112h.0.0). Estos tres estados (normoxia, reactivación e hipoxia tardía) se diferencian de las condiciones del estado de hipoxia temprana (X47h.0.0

a X70h.0.0), que se orientan principalmente hacia la primera componente principal (Dim1) y tienen un valor negativo en Dim2. Finalmente, las condiciones del estado de disminución de oxígeno (X40h.11.5 a 46h.0.41) también se orientan hacia Dim1, con un valor cercano a uno, pero presentan un valor positivo en Dim2. En conclusión, los resultados de expresión diferencial binarizados de la tabla 8 se pueden dividir en cinco estados experimentales generales de expresión diferencial que se utilizarán para evaluar el modelo booleano de dormancia.

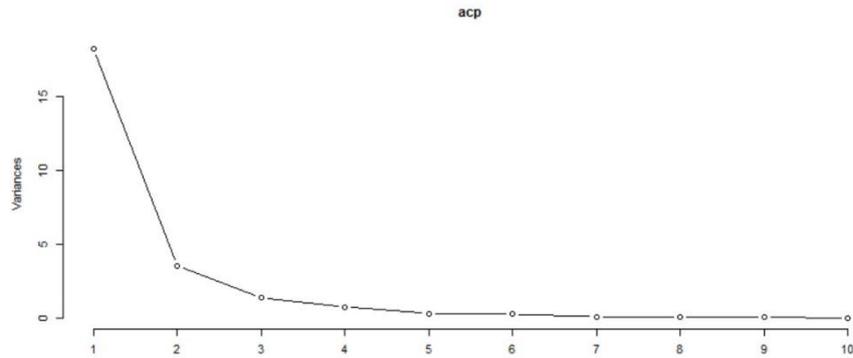


Fig. 28 Varianza de los primeros diez componentes principales. En el eje Y se muestra la varianza de los componentes principales, mientras que en el eje X se presentan los primeros diez componentes principales.

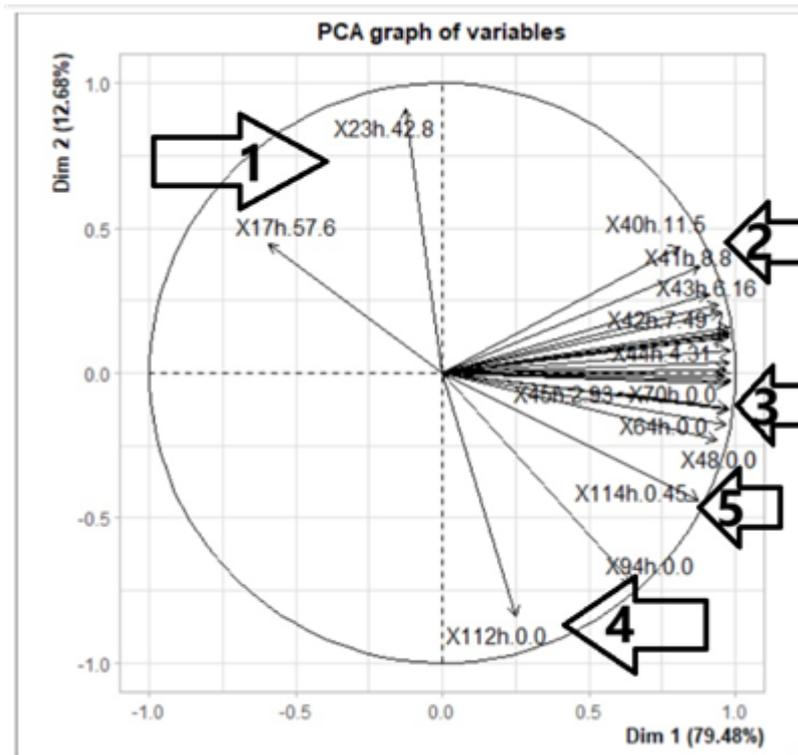


Fig. 29 Círculo de correlaciones para las 25 condiciones experimentales. En el eje “X” se encuentra la primera componente principal (Dim1) que contiene el 79.48 % de la varianza. En el eje “Y” se ubica la segunda componente principal (Dim2) que contiene al 12.68 % de la varianza. Los vectores representan a cada una de las condiciones experimentales. En base a la orientación de los vectores se pueden distinguir 5 estados generales: 1 (Normoxia), 2 (Disminución de oxígeno), 3 (Hipoxia Temprana), 4 (Hipoxia Tardía) y 5 (Reactivación).

Los estados de hipoxia se binarizados de la tabla 8 se agruparon en estados generales en base a los resultados de agrupamiento de la sección anterior. Estos estados deben ser reproducidos por el modelo de dormancia.

Como se muestra en la tabla 9, los diez primeros genes, desde Rv3133c hasta Rv3597c, muestran actividad durante la hipoxia temprana e inactividad durante la

hipoxia tardía, mientras que los siguientes nueve genes muestran actividad durante la hipoxia tardía. Estos estados de activación de los genes en cada estado general deben poder ser reproducidos por los atractores del modelo booleano de dormancia.

GENES	N				D						H. TEMP						H. TAR		R		Umbrales					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		21	22	23	24	25
Rv3133c	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0	1.071882336
Rv3132c	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	1.051298420
Rv0081	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	2.357482073
Rv0474	0	0	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	0	0.012028063
Rv0757	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0.178179047
Rv1990c	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0.779452129

Rv2011c	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0.0460901
Rv3574	0	0	0	0	1	1	1	1	1	0	1	0	0	0	1	1	0	0	0	0	0	0	0.2879582

Tabla 9 Agrupamiento en estados generales para las 25 muestras. Como se describe en la sección 7.1, el porcentaje de oxígeno disuelto es del 57.6 % en la muestra 1 y disminuye hasta el 0.0 % en la muestra 10, para luego volver a aumentar a 0.41 % en la muestra 21, hasta llegar al 46.92 % en la muestra 25. Teniendo en cuenta esto y los resultados de agrupamiento en la sección anterior, se identificaron cinco estados generales: Normoxia (N), Disminución de oxígeno (D), Hipoxia Temprana (H Temprana), Hipoxia Tardía (HT) y Reactivación (R). En las primeras diez filas se ubican los genes que se activan durante la hipoxia temprana y en las siguientes nueve filas los que se activan durante la hipoxia tardía.

7.4.OBTENCIÓN DE LA RED DE DORMANCIA DE MTB

7.4.1. Red de dormancia mínima (RDM)

En base a los criterios descritos en la sección 7.4.2, se retiraron a los siguientes nodos de la red inicial del trabajo de Bose y colaboradores (64): Rv0767c, Rv2374c, Rv0348, Rv0827c y Rv1255c. Todos son nodos de salida como se observa en la figura 8 y ninguno forma parte de las subredes críticas para la adaptación a la hipoxia (34).

A continuación, se muestra la red de dormancia reducida de 19 nodos con sus funciones lógicas optimizadas. Ver figura 30.

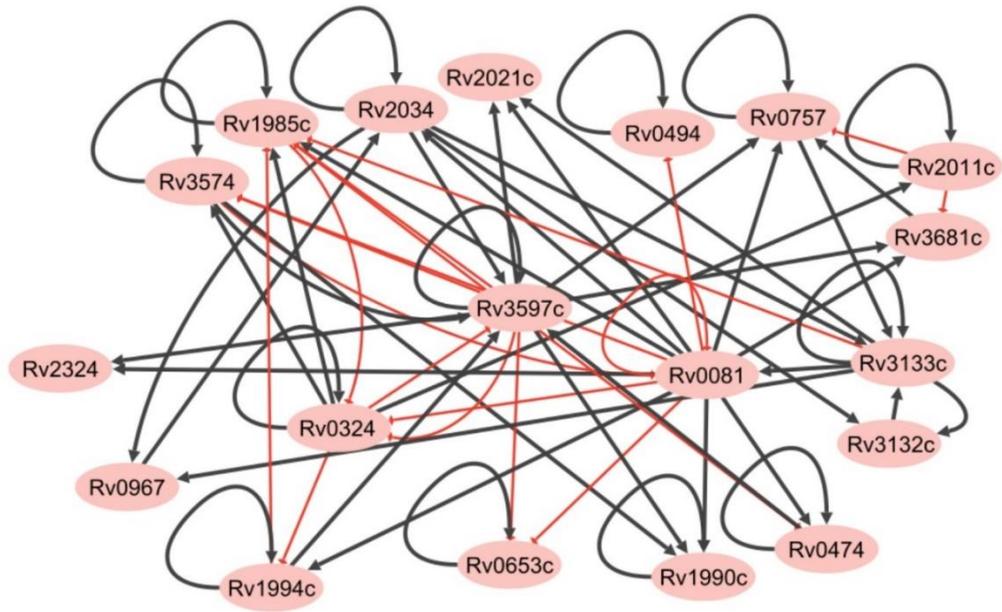


Fig. 30 Red de dormancia reducida de 19 nodos. En color negro se muestran las interacciones dirigidas activadoras y en color rojo las interacciones dirigidas inhibitoras entre los nodos.

GEN	FUNCIONES BOOLEANAS
Rv1994c	$(Rv0081 \mid Rv1994c) \& (! Rv0324)$
Rv3133c	$Rv3132c \mid (Rv0757 \& Rv2034) \mid (Rv3133c \& (Rv3132c \mid (Rv0757 \& Rv2034)))$
Rv3132c	$Rv3133c \mid Rv2034$
Rv3574	$(Rv3574 \mid Rv0324) \& (!(Rv0081 \mid Rv3597c))$
Rv3597c	$(Rv1994c \mid Rv3574 \mid Rv2034 \mid Rv0474 \mid Rv3597c) \& (!(Rv0324)$
Rv0757	$(Rv3681c \mid Rv0081 \mid Rv0757 \mid Rv3597c) \& (!(Rv2011c)$

Rv0081	$(!(Rv3574 Rv3133c)) \& !(Rv0081)$
Rv0324	$(Rv0324 !(Rv3597c \& Rv1985c)) \& !(Rv0081)$
Rv1985c	$(Rv0081 Rv1985c Rv0324) \& !(Rv1994c Rv3597c Rv3133c)$
Rv2011c	$(Rv2011c Rv0324)$
Rv2034	$Rv0081 Rv2034 Rv0967$
Rv3681c	$Rv3597c (Rv0081 \& !Rv2011c)$
Rv0967	$Rv3133c Rv2034$
Rv0474	$Rv0081 Rv0474 \& !(Rv1985c)$
Rv0494	$(Rv0494) \& !(Rv0081)$
Rv0653c	$(Rv0653c) \& !(Rv3597c Rv0081)$
Rv1990c	$(Rv3597c Rv0081 Rv1990c) \& !(Rv3574)$
Rv2021c	$Rv0081 Rv3597c Rv3133c$
Rv2324	$Rv0081 Rv3597c$

Tabla 10 Funciones booleanas de la red reducida de 19 nodos. El símbolo “|” representa la función booleana “OR”, el símbolo “&” representa la función booleana “AND”, y el símbolo “!” representa la función booleana “NOT”.

7.4.2. Obtención de los atractores de la red de dormancia mínima

Aplicando el procedimiento descrito en la sección 7.4.3, se encontraron seis atractores para la RDM. En la figura 31 se pueden observar los estados de actividad

de los genes en cada uno de los seis atractores de la RDM. En el panel izquierdo de la figura 31 se muestran los 19 nodos de la RDM. Los rectángulos verdes y rojos representan el estado en el que se encuentran los 19 nodos para cada uno de los seis atractores, un rectángulo verde indica que el nodo está activo, mientras que un rectángulo rojo representa un estado inactivo del nodo en cuestión. Por ejemplo, El nodo Rv3133c se encuentra activo en los atractores uno y dos, mientras que se encuentra inactivo en los atractores tres, cuatro, cinco y seis. Por otra parte, el nodo Rv0324 presenta un patrón de activación opuesto, es decir, se encuentra inactivo en los atractores uno y dos, mientras que se encuentra activo en los atractores tres, cuatro, cinco y seis.

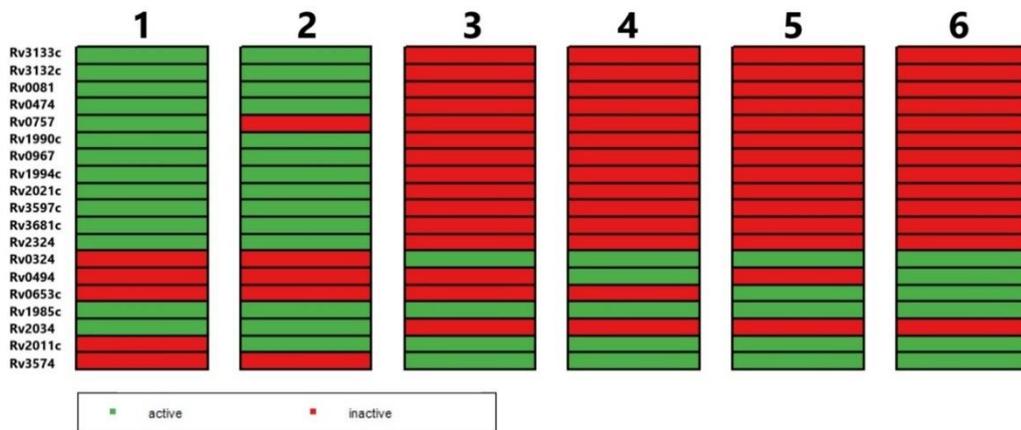


Fig. 31 Atractores de la RDM. En el lado izquierdo se observa a los genes de la RDM. En cada uno de los seis atractores los rectángulos verdes y rojos representan el estado de cada gen en el atractor. en color verde se representa un estado activo del gen, mientras que en color rojo se representa un estado inactivo.

8.4.2.1 Análisis de la reproducción de los estados experimentales por parte de los atractores de la RDM.

Para realizar dicho análisis se procede a comparar los estados de los nodos en los atractores de la RDM con sus estados experimentales descritos en la tabla 9.

Al realizar la comparación entre los estados generales y los atractores de la RDM, se comprueba que el modelo booleano puede reproducir tres de los estados generales. El atractor N°1 reproduce el estado general de Hipoxia Temprana, ver figura 31; el atractor N.º 2 reproduce el estado general de Reactivación y el atractor N°3 reproduce el estado general de Hipoxia Tardía.

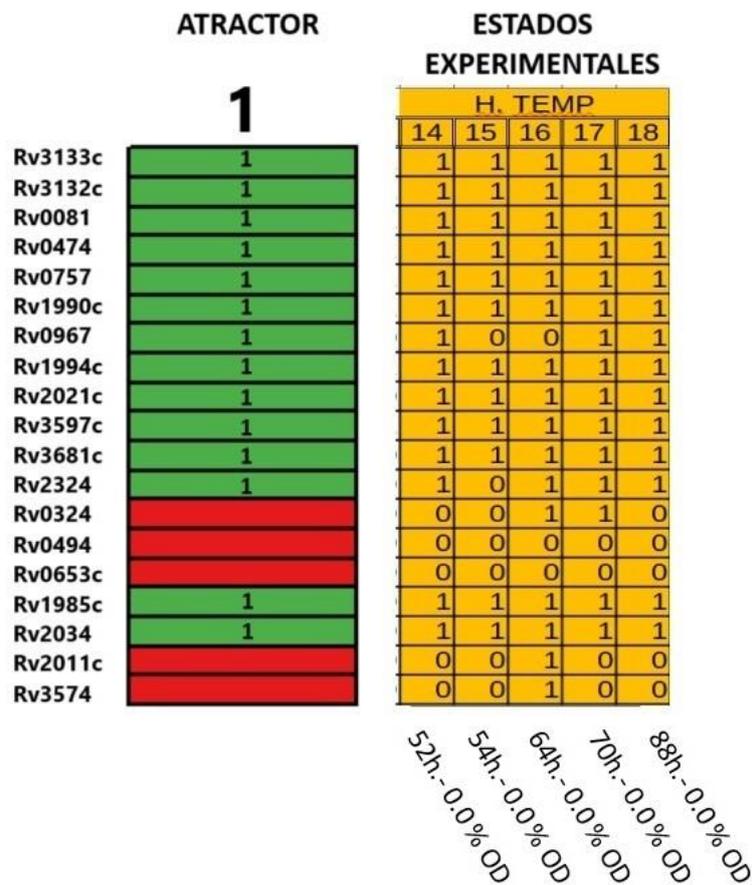


Fig. 32 Comparación entre el primer atractor de la RDM y el estado general de Hipoxia Temprana. En el panel izquierdo se muestra el atractor N°1 de la RDM. En el panel derecho se representan los resultados de binarización de la expresión genética diferencial para las 5 muestras (14 a 18) que conforman el estado general de Hipoxia Temprana. En esta representación, los rectángulos verdes y el valor “1” indican un estado activo, mientras que los rectángulos rojos y el valor “0” denotan un estado inactivo.

Como se muestra en la figura 32, el primer atractor puede representar la mayoría de los estados experimentales de Hipoxia Temprana. En este estado, hay catorce genes que mantienen el mismo patrón en las cinco muestras del estado de Hipoxia Temprana: activos para los genes Rv3133c, Rv3132c, Rv0081, Rv0474, Rv0757, Rv1990c, Rv1994c, Rv2021c, Rv3597c, Rv3681c, Rv1985c y Rv2034, e inactivos para los genes Rv0494 y Rv0653c. Además, se observa que cinco genes reproducen

principalmente el estado del primer atractor en la mayoría de las muestras: Rv0967 y Rv2324 se encuentran principalmente activos, mientras que Rv0324, Rv2011c y Rv3574 permanecen en su mayor parte inactivos. Por ejemplo, el gen Rv2011c está inactivo en las muestras 14,15,17 y 18, y solo está activo en la muestra 16, lo que es coherente con un estado inactivo en el primer atractor.

	ATRACTOR	ESTADOS EXPERIMENTALES
	2	R
		21
Rv3133c	1	1
Rv3132c	1	1
Rv0081	1	1
Rv0474	1	1
Rv0757	0	0
Rv1990c	1	1
Rv0967	1	1
Rv1994c	1	1
Rv2021c	1	1
Rv3597c	1	1
Rv3681c	1	1
Rv2324	1	1
Rv0324	0	1
Rv0494	0	0
Rv0653c	0	0
Rv1985c	1	1
Rv2034	1	1
Rv2011c	1	1
Rv3574	0	0

114h.- 0.45 % OD

Fig. 33 Comparación entre el segundo atractor de la RDM y el estado general de Reactivación. En el panel izquierdo se muestra el atractor N°2 de la RDM. En el panel derecho se presentan los resultados de binarización de la expresión genética diferencial para la muestra 21 (reoxigenación con 0.45% de oxígeno), que conforma el estado general de Reactivación. En esta representación, los rectángulos verdes y el valor “1” indican un estado activo, mientras que los rectángulos rojos y el valor “0” denotan un estado inactivo.

La figura 33 muestra que el segundo atractor puede representar de manera efectiva el estado general de Reactivación. En este estado, dieciocho genes mantienen un patrón similar al observado en la muestra 21 del estado de Hipoxia Temprana: activos para los genes Rv3133c, Rv3132c, Rv0081, Rv0474, Rv1990c, Rv0967, Rv1994c, Rv2021c, Rv3597c, Rv3681c y Rv2324, e inactivos para los genes Rv0757, Rv0494, Rv0653c y Rv3574. Sin embargo, existe una discrepancia en el estado de Rv0324 entre el resultado de la binarización y el atractor de Reactivación. Mientras el algoritmo de binarización indica que está inactivo, en el atractor de Reactivación se muestra activo. No obstante, al analizar la tabla S2 de expresión diferencial para dicho gen (ver anexo 2), se evidencia una disminución del nivel de expresión genética cercana al 50 % al pasar de la muestra 20 (Hipoxia Tardía) a la muestra 21 (Reactivación). Esta discrepancia podría atribuirse a un umbral (Threshold) muy bajo que no permite captar toda la dinámica de expresión de este gen.

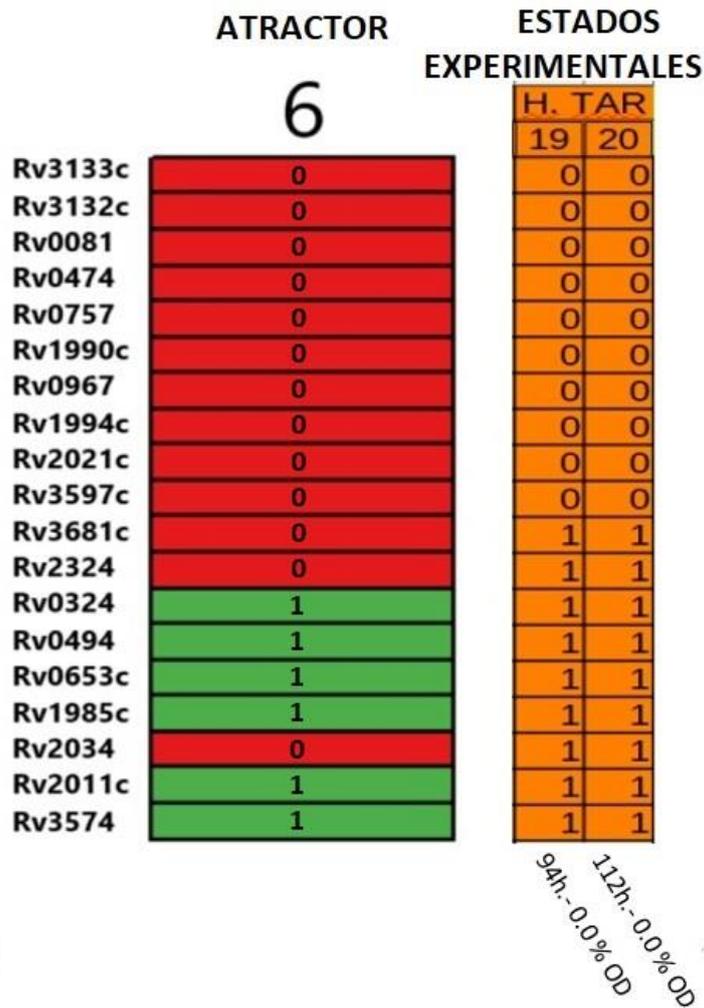


Fig. 34 Comparación entre el sexto atractor de la RDM y el estado general de Hipoxia Tardía. En el panel izquierdo se muestra el atractor N°6 de la RDM. En el panel derecho se presentan los resultados de binarización de la expresión genética diferencial para las muestras 19 y 20 (88 y 94 horas con 0,0% de oxígeno, respectivamente), que conforman el estado general de Hipoxia Tardía. En esta representación, los rectángulos verdes y el valor “1” indican un estado activo, mientras que los rectángulos rojos y el valor “0” denotan un estado inactivo.

La figura 34 muestra que el sexto atractor puede representar de manera efectiva el estado general de Hipoxia Tardía. En este estado, dieciséis genes mantienen un patrón similar al observado en las muestras 19 y 20 del estado general de Hipoxia Tardía: activos para los genes Rv0324, Rv0494, Rv0653c, Rv1985c, Rv2011c y Rv3574, e inactivos para los genes Rv3133c, Rv3132c, Rv0081, Rv0474, Rv0757, Rv1990c, Rv0967, Rv1994c, Rv2021c y Rv3597c. Sin embargo, existen discrepancias en los estados de los nodos Rv3681c, Rv2324 y Rv2034 entre el resultado de la binarización y el atractor de Hipoxia Tardía. Mientras que el algoritmo de binarización indica que están activos, en el atractor de Hipoxia Tardía se muestran inactivos. Sin embargo, al analizar la gráfica de expresión diferencial para el gen Rv3681c (ver tabla S2 del anexo 2), se evidencia una disminución del nivel de expresión genética del orden del 50 % al pasar de la muestra 18 (Hipoxia Temprana) a la muestra 19 (Hipoxia Tardía). En el caso de los genes Rv2324 y Rv2034, esta disminución es del orden del 10 y 20 % respectivamente. Por lo tanto, esta discrepancia, fundamentalmente para el nodo Rv3681c, podría estar relacionada con un umbral muy bajo que no permite captar la dinámica de expresión de estos genes. Con el objetivo de verificar si estos resultados son congruentes con otros modelos experimentales de hipoxia, se analizaron los estados de los nodos en el trabajo de Ignatov y colaboradores (81), donde se alcanza el estado de dormancia de MTB en un medio deficiente en potasio. En estas condiciones se observa que la expresión de los nodos Rv0081, Rv3133c, R0967, Rv3297c, Rv0757 y Rv3681c disminuye considerablemente en el estado general de hipoxia tardía, lo que está de acuerdo

con su estado inactivo en el atractor de hipoxia tardía. De forma similar para los nodos Rv2011c, Rv0324, Rv1985c, Rv0494 y Rv0653c se observa un aumento en sus niveles de expresión en estado general de hipoxia tardía, lo que está de acuerdo con su estado activo en el atractor de hipoxia tardía. Por lo tanto, de los tres nodos que muestran discrepancia, el nodo Rv3681c si sería congruente con los datos experimentales. Los atractores 3, 4 y 5 muestran un comportamiento similar al atractor de Hipoxia Tardía, con los genes Rv3133c, Rv3132c, Rv0081, Rv0474, Rv0757, Rv1990c, Rv1994c, Rv2021c, Rv3597c, Rv3681c, Rv1985c y Rv2034 en estado inactivo. Además, todos poseen cuencas de atracción de 512 estados. Las cuencas de atracción se definen como el número de estados del sistema que decaen hacia un determinado atractor. Por otro lado, el atractor de Reactivación posee la cuenca de atracción más grande con 407552 estados, mientras que la cuenca de atracción del atractor de Hipoxia Temprana tiene 114688 estados. La relación entre ambas cuencas de atracción es de 3.55.

7.4.3. Funciones reguladoras *in silico*

Con los algoritmos descritos en la sección 7.4.4 se procede a obtener las funciones reguladoras *in silico* a partir de los datos de expresión diferencial del trabajo de Peterson y colaboradores (65). Se muestra como ejemplo las funciones reguladoras para el nodo Rv1994c:

Alternative transition functions for gene Rv1994c:
--

$$\begin{aligned} Rv1994c = (& !Rv1994c \& !Rv3574 \& Rv0324 \& !Rv2034 \& Rv2011c) \\ & | (!Rv1994c \& Rv3574 \& Rv0324 \& !Rv2034 \& !Rv2011c) | (Rv3574 \\ & \& Rv0324 \& Rv2034 \& Rv2011c) | (Rv1994c \& !Rv0324 \& Rv2034 \\ & \& !Rv2011c) | (Rv1994c \& Rv3574 \& !Rv0324 \& !Rv2011c) \text{ (error:} \\ & 2) \end{aligned}$$

En el Anexo 4 se muestran el resto de funciones reguladoras *in silico*

Estas funciones *in silico* se utilizan para compararlas con las obtenidas experimentalmente en los trabajos de Turkarslan y colaboradores (73) y de Galagan y colaboradores (34), que como se describe en la sección 7.4.1 se usaron para obtener las funciones reguladoras de los nodos de la RDM y se encuentran en la tabla 10.

Al realizar la comparación se verificó que las funciones reguladoras *in silico* del nodo Rv3574 son las que más se ajustan a las funciones reguladoras obtenidas de los trabajos experimentales.

Función predicha *in silico*:

$$Rv3574 = (Rv0757 \& Rv0081) | (Rv3574 \& !Rv0081) \text{ (error: 1)}$$

Función reguladora experimental:

$$Rv3574 = (Rv3574 | Rv0324 \& !(Rv0081 | Rv3597c))$$

En ambos casos, el nodo Rv3574 actúa como un inductor, mientras que el nodo

Rv0081 se comporta como un inhibidor.

Como se puede observar en la tabla 11, la cantidad de funciones generadas *in silico* para cada uno de los 19 nodos varían dentro de un amplio rango, desde una sola función para el nodo Rv0324 hasta 69 para el nodo Rv2011c. Además, las funciones *in silico* también tienen una amplia variabilidad en la predicción de las funciones experimentales de la red. En el caso de los nodos Rv0081, Rv2011c, Rv3597c y Rv3132c las funciones *in silico* contienen todos los nodos presentes en sus correspondientes funciones experimentales. Mientras que para los nodos Rv0324, Rv2034, Rv3681c, Rv0653c, Rv0967, Rv2021c y Rv2034 sus funciones *in silico* no contienen ningún nodo presente en su función experimental. El resto de los 8 nodos presentan funciones *in silico* con al menos un nodo presente en sus funciones experimentales.

Nodo	Nodos de funciones <i>in silico</i> presentes en la correspondiente función experimental	Cantidad de funciones <i>in silico</i> generadas	Función experimental
Rv0081	Rv0081, Rv3574, Rv3133c	32	(!(Rv3574) Rv3133c)) & (!Rv0081)
Rv3133c	Rv0757	2	Rv3132c (Rv0757 & Rv2034) (Rv3133c & (Rv3132c (Rv0757 & Rv2034)))

Rv3132c	Rv2034, Rv3133c	21	Rv3133c Rv2034
Rv3574	RV3574, Rv0081	3	(Rv3574 Rv0324) & (!(Rv0081 Rv3597c))
Rv3597c	Rv3574, Rv2034, Rv1994c, Rv3597c, Rv0474.	50	(Rv3574 Rv2034 Rv1994c Rv3597c Rv0474) & !(Rv0324))
Rv0757	Rv3681c, Rv0757.	2	(Rv3681c Rv0081 Rv0757 Rv3597c) & !(Rv2011c)
Rv0324	0	1	(Rv0324 !(Rv3597c & Rv1985c)) & !(Rv0081)
Rv1985c	Rv1985c, Rv3133c.	2	Rv0324 Rv0081 (Rv1985c) & !(Rv1994c Rv3133c Rv3597c))
Rv2034	0	9	Rv0081 Rv2034 Rv0967
Rv3681c	0	4	Rv3597c (Rv0081& !Rv2011c)
Rv2011c	Rv0324, Rv2011c	69	(Rv2011c Rv0324)
Rv0474	Rv0474, Rv1985c.	29	Rv0081 Rv0474 & !(Rv1985c)
Rv0494	Rv0494,	1	(Rv0494) & !(Rv0081)
Rv0653c	0	2	Rv0653c & !(Rv3597c Rv0081)
Rv0967	0	5	Rv3133c Rv2034
Rv1990c	Rv1990c, Rv3574	20	(Rv3597c Rv0081 Rv1990c) & ! (Rv3574)
Rv1994c	Rv0081	8	(Rv0081 Rv1994c) & !(Rv0324)
Rv2021c	0	4	Rv0081 Rv3597c Rv3133c

Rv2324	0	10	Rv3597c Rv0081
--------	---	----	------------------

Tabla 11 Nodos de funciones *in silico* presentes en las funciones experimentales de la red reducida de 19 nodos. En la columna de la izquierda se encuentran los 19 nodos de la red. Se observa una gran variabilidad en la cantidad de funciones *in silico* generadas para cada nodo. En la primera fila de la tabla se puede apreciar que los nodos de las funciones *in silico* del nodo Rv0081: Rv0081, Rv3574, Rv3133c se encuentran presentes en su función experimental: $(!(Rv3574) | Rv3133c) \& (!(Rv0081))$.

7.5. ANÁLISIS DE MUTACIONES EN LA RED

7.5.1. Verificación de importancia de los nodos en la estructura general

Siguiendo el procedimiento descrito en la sección 7.5.1, se eliminó un nodo de la red y se procedió a obtener los atractores correspondientes sin el nodo en cuestión. Este proceso se repitió para los dieciocho nodos restantes. Los resultados obtenidos se resumen en la tabla 12.

GEN ELIMINADO	VARIACIÓN EN LOS ATRACTORES			Cantidad de Atractores
	Hipoxia Temprana	Hipoxia Tardía	Reactivación	
Rv1994c	0	0	0	6
Rv3133c	3	0	3	11 *
Rv3132c	0	0	1	10

Rv3574	0	2	0	50
Rv3597c	0	2	1	54
Rv0757	0	0	0	6
Rv0081	2	4	3	56
Rv0324	0	AUSENTE	0	2
Rv1985c	0	0	0	10
Rv2011c	0	1	AUSENTE	9
Rv2034	0	0	0	8**
Rv3681c	0	0	0	6
Rv0967	0	0	0	6
Rv0474	0	0	0	6
Rv0494	0	0	0	4
Rv0653c	0	0	0	4
Rv1990c	<u>0</u>	<u>0</u>	<u>0</u>	<u>6</u>
Rv2021c	0	0	0	6
Rv2324	0	0	0	6

* 8 atractores de un estado y 3 atractores de dos estados.

** 7 atractores de un estado y 1 atractor de dos estados.

Tabla 12 Variación de los atractores al eliminar un nodo. La tabla muestra la cantidad de nodos que cambian su estado en cada atractor particular al eliminar un nodo en comparación con los atractores de la RDM original. En el caso de los genes Rv0324 y Rv2011c, el término “AUSENTE” representa la ausencia del atractor de Hipoxia Tardía y de Reactivación, respectivamente, cuando se elimina el nodo en cuestión. Los valores numéricos representan la cantidad de nodos que cambian de estado respecto al atractor de la RDM original, por ejemplo, para el caso de la eliminación del nodo Rv2324, todos los atractores del sistema se mantienen sin variaciones.

El análisis de la RDM luego de la eliminación de nodos específicos sugiere que los nodos Rv0081, Rv0324, Rv3574, Rv2011c, Rv3597c, Rv3132c y Rv3133c son los más críticos para mantener los atractores de la RDM. Se observa que Rv0324 y Rv2011c son fundamentales para alcanzar los atractores de Hipoxia Tardía y de Reactivación, respectivamente. Además, se evidencia que Rv0081 y Rv3133c son nodos críticos para alcanzar el estado de Hipoxia Temprana, lo que concuerda con la evidencia experimental (5), (6). Los nodos Rv3574, Rv3597c también son importantes para alcanzar el estado de Hipoxia Tardía, que es el de mayor resistencia de MTB.

7.5.2. Análisis de la trayectoria hacia el atractor

Siguiendo el procedimiento descrito en la sección 7.5.2, se encienden genes particulares de la RDM, lo que constituye un estado inicial y se analiza la trayectoria desde ese estado inicial hasta el atractor. Este análisis se realizó para los nodos críticos encontrados en la sección anterior. A continuación, se muestran los resultados de estas simulaciones.

1) Partiendo de los nodos Rv3133c y Rv3132c activados y con todos los demás nodos apagados se llega al atractor de Hipoxia Temprana en tres pasos como se muestra en la figura 35. En el panel izquierdo se muestran los genes de la RDM. En la parte superior se indican los pasos de tiempo. Como se observa en la figura, en el estado de tiempo “0”, la red tiene activos a los genes Rv3133c y Rv3132c, mientras que el resto de diecisiete genes se encuentran inactivos. El siguiente estado se obtiene evaluando las funciones booleanas para cada gen que se muestran en la tabla 10, con lo que obtienen los estados de cada gen al tiempo “1”. En este siguiente estado se observa que se encuentran activos los genes Rv3133c, Rv3132c, Rv0081, Rv0967c y Rv2021c, mientras que el resto de catorce genes se encuentran inactivos. Este proceso se repite dos veces más hasta que se alcanza el estado de tiempo “3” donde se alcanza al atractor de Hipoxia Temprana y los estados de los genes ya no cambian al evaluar sus funciones reguladoras.

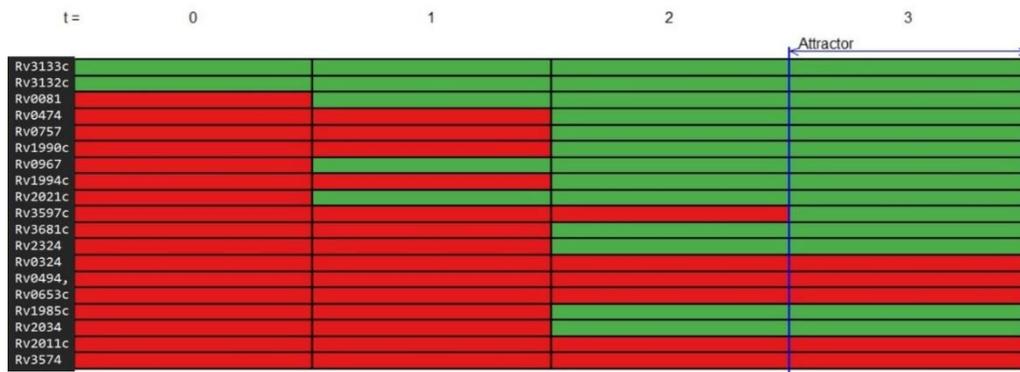


Fig. 35 Estados hasta el atractor partiendo de los nodos Rv3133c y Rv3132c activados. En el panel izquierdo se muestran los genes de la RDM. En la parte superior se indican los pasos de tiempo. El sistema parte de los genes Rv3133c y Rv3132c activos y después de tres pasos de tiempo el sistema alcanza el atractor de Hipoxia Temprana.

2) Partiendo del nodo Rv0081 activo, se llega al atractor de Hipoxia Temprana en dos pasos de tiempo como se muestra en la figura 36.

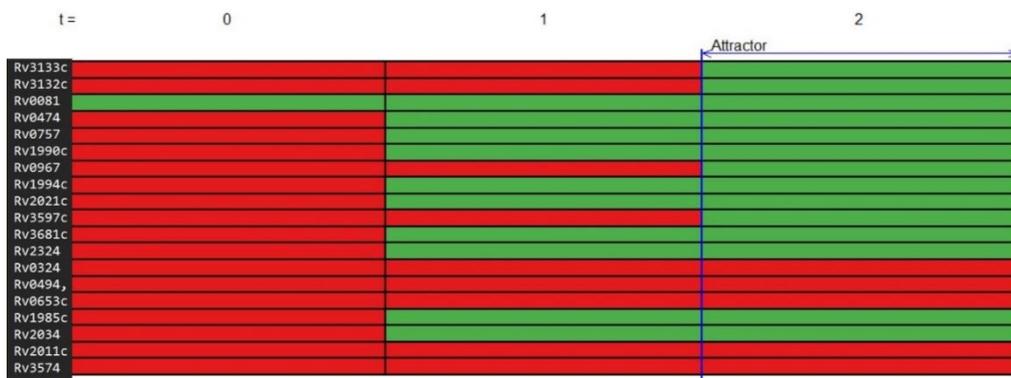


Fig. 36 Estados hasta el atractor partiendo del nodo Rv0081 activo.

3) Partiendo del nodo Rv3133c activo, se puede llegar al atractor de Hipoxia Temprana en tres pasos de tiempo como se muestra en la figura 37.

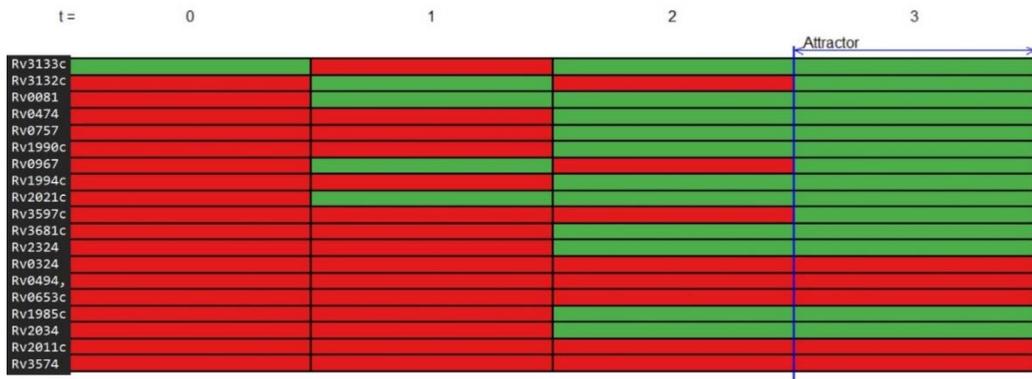


Fig. 37 Estados hasta el atractor partiendo del nodo Rv3133c activo.

4) Partiendo del nodo Rv3132c activo, se puede llegar al atractor de Hipoxia Temprana en tres pasos de tiempo como se muestra en la figura 38.

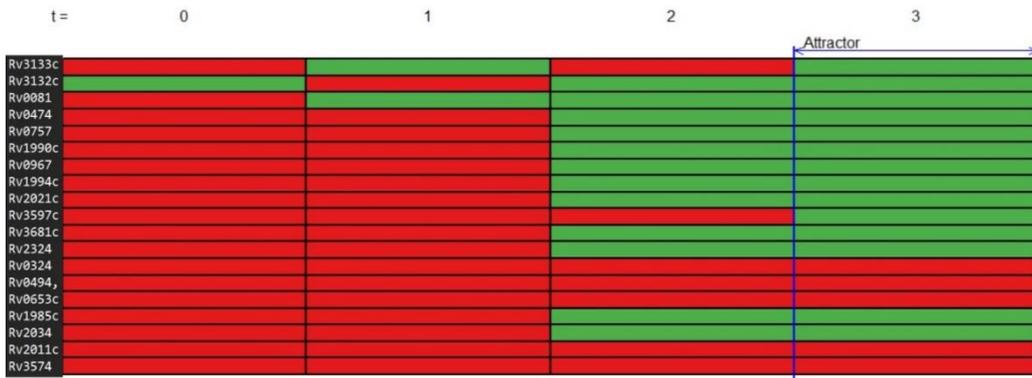


Fig. 38 Estados hasta el atractor partiendo del nodo Rv3132c activo.

5) Partiendo de Rv3133c, Rv3132c, y Rv0081 activos, se puede llegar al atractor de Hipoxia Temprana en dos pasos de tiempo como se muestra en la figura 39.

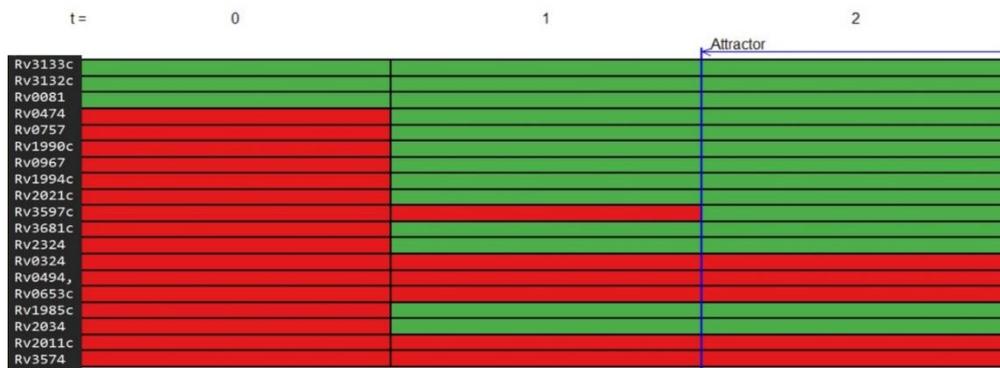


Fig. 39 Estados hasta el atractor partiendo de los nodos Rv3132c, Rv3133c y Rv0081 activos.

La evidencia experimental ha demostrado que el sistema de dos componentes DevR/DevS (Rv3132c/Rv3133c) es el primero en activarse. Luego se activan los genes de respuesta temprana a la hipoxia, como el gen Rv0081, que le permiten a MTB alcanzar el estado de Hipoxia Temprana (5, 6).

Los análisis de trayectorias hacia el atractor demuestran que, partiendo de un estado inicial con alguno de los nodos Rv0081, Rv3133c y Rv3132c activos, se alcanza el atractor de Hipoxia Temprana. Además, cuando se parte de un estado inicial donde solo el nodo Rv0081 se encuentra activo (figura 36), el tiempo para alcanzar el atractor de Hipoxia Temprana es menor respecto del estado inicial en el que se encuentran activos los nodos Rv3132c y Rv3133c (figuras 37 y 38). Por lo tanto, el

modelo reproduce la trayectoria de activación hacia el estado de Hipoxia Temprana, que inicia con la activación del sistema de dos componentes DevR/DevS (Rv3132c/Rv3133c), seguido de la activación del nodo Rv0081 para que se puedan activar los genes de la Hipoxia Temprana.

6) Partiendo del nodo Rv0324 activo, se puede llegar al atractor de Reactivación en seis pasos de tiempo como se observa en la figura 40.

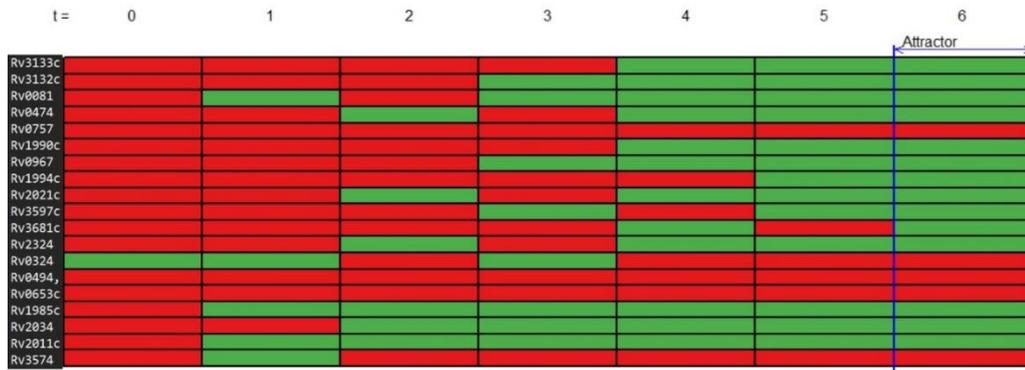


Fig. 40 Estados hasta el atractor partiendo del nodo Rv0324 activo.

7) Partiendo del nodo Rv3574 activo, se puede llegar al atractor de Hipoxia Temprana en cinco pasos de tiempo como se observa en la figura 41.

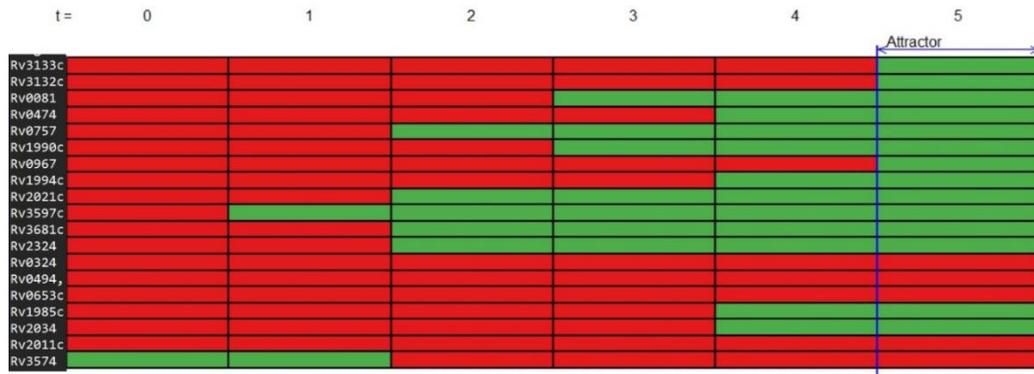


Fig. 41 Estados hasta el atractor partiendo del nodo Rv3574 activo.

8) Partiendo del nodo Rv3597c activo, se puede llegar al atractor de Hipoxia Temprana en tres pasos de tiempo como se observa en la figura 42.



Fig. 42 Estados hasta el atractor partiendo del nodo Rv3597c activo.

9) Partiendo del nodo Rv2011c activo, se puede llegar al atractor de Reactivación en cuatro pasos de tiempo como se observa en la figura 43.



Fig. 43 Estados hasta el atractor partiendo del nodo Rv2011c activo.

10) Solo partiendo de los nodos Rv0324 y Rv3574 activos, se puede llegar a un atractor similar al atractor de Hipoxia Tardía en solo un paso de tiempo como se observa en la figura 44.

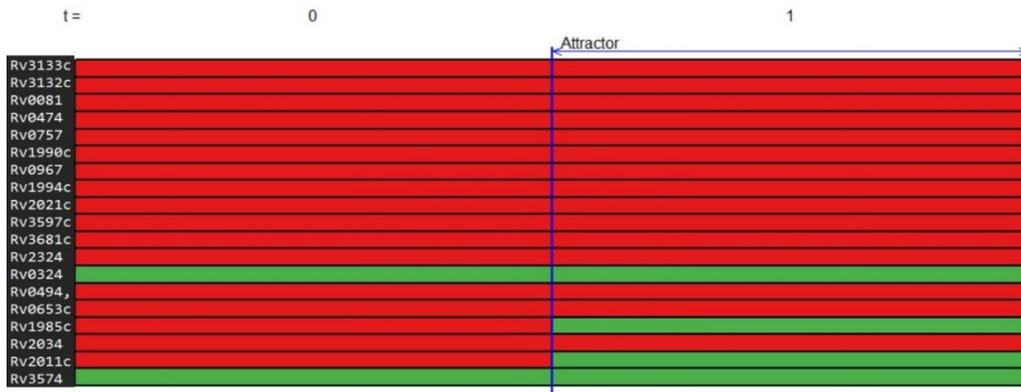


Fig. 44 Estados hasta el atractor partiendo de los nodos Rv0324 y Rv3574 activos.

En la tabla 13 se muestran los resultados del análisis de trayectorias para todos los pares genes posibles.

Nodos	Rv0081	Rv3133c	Rv3132c	Rv3574	Rv3597c	Rv0757
Rv2324	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv2021c	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv0967c	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv1990c	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv0967	H(2)	H(2)	H(3)	H(4)	H(2)	H(2)
Rv0653c	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv0494	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv0757	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv2011c	R(3)	R(4)	R(4)	R(6)	R(4)	R(4)
Rv0757	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv2021c	H(2)	H(2)	H(2)	H(4)	H(2)	H(2)
Rv1985c	H(2)	R(4)	R(4)	R(6)	H(3)	R(4)
Rv0324	R(3)	R(4)	R(4)	TA(0)	R(4)	R(6)
Rv0757	H(2)	H(3)	H(3)	H(5)	H(3)	H(3)
Rv3597c	H(2)	H(3)	H(3)	H(4)	H(3)	H(3)
Rv3574	H(2)	H(4)	H(4)	H(5)		
Rv3132	H(2)	H(3)	H(3)			
Rv3133	H(2)	H(3)				
Rv0081	H(2)					

11) Partiendo de los nodos Rv0324, Rv3574, Rv0494 y Rv0653c activos, se puede llegar al atractor de Hipoxia Tardía en solo un paso de tiempo como se observa en la figura 45.

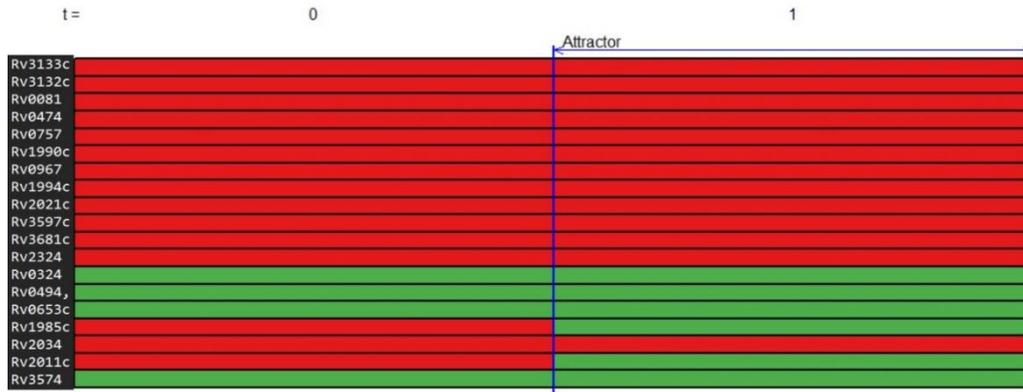


Fig. 45 Estados hasta el atractor partiendo de los nodos Rv0324, Rv3574, Rv0494 y Rv0653c activos.

Estos resultados sugieren que tanto Rv3574 y Rv0324 son críticos para alcanzar el estado de hipoxia tardía y que el nodo Rv2011c es esencial para alcanzar el atractor de Reactivación, como se mostró en la sección anterior.

Estos resultados están de acuerdo con la evidencia experimental que demuestra que el nodo Rv0324 es fundamental para generar la respuesta tardía a la hipoxia (34). También, el nodo Rv3574, que está implicado en la adaptación al metabolismo lipídico, es fundamental para que MTB se adapte energéticamente al estado de hipoxia prolongado (82).

7.5.3. Mutaciones de activación

Siguiendo el procedimiento descrito en la sección 7.5.3 se realizaron las mutaciones de activación. Los resultados se muestran en la tabla 14.

GEN FIJADO EN ESTADO ACTIVO	NODOS DE VARIACIÓN EN LOS ATRACTORES			Cantidad de atractores
	Hipoxia Temprana	Hipoxia Tardía	Reactivación	
Rv3133c	0	AUSENTE	0	2
Rv3132c	0	AUSENTE	0	2
Rv0081	0	AUSENTE	0	2
Rv0474	0	1	0	6
Rv0757	0	1	0	6
Rv1990c	0	1	0	6
Rv0967	0	AUSENTE	0	2
Rv1994c	0	1	0	6
Rv1985c	0	0	0	6
Rv2021c	0	1	0	6
Rv3597c	0	AUSENTE	0	2

Rv3681c	0	1	0	6
Rv2324	0	1	0	6
Rv0324	AUSENTE	0	4	5
Rv0494	1	0	1	5
Rv0653c	1	0	1	5
Rv2034	<u>0</u>	<u>AUSENTE</u>	<u>0</u>	<u>2</u>
Rv2011c	AUSENTE	0	0	5
Rv3574	2	0	2	22

Tabla 14 Variación de los atractores al fijar en estado activo un nodo. La tabla muestra la cantidad de nodos que cambian su estado en cada atractor particular al ejecutar la mutación de activación respecto de los atractores de la RDM sin mutaciones. El término “AUSENTE” representa la ausencia del atractor. El valor numérico hace referencia a la cantidad de nodos que varían en el atractor, por ejemplo, para la mutación de activación del gen Rv3574, se verifica que hay dos nodos que cambian de estado tanto para el atractor de Hipoxia Temprana como para el atractor de Reactivación, mientras que el atractor de Hipoxia Tardía se mantiene sin variaciones.

1) La mutación de activación de los nodos Rv0081, Rv3133c y Rv3132c, la red mantiene el atractor de Hipoxia Temprana como era de esperarse por el hecho de que estos nodos son los responsables de la respuesta temprana a la hipoxia. También se observa que se mantiene el atractor de Reactivación y desaparece el atractor de Hipoxia Tardía. Ver figura 46. Por lo tanto, esta mutación de activación no permite alcanzar el atractor de Hipoxia Tardía al sistema.

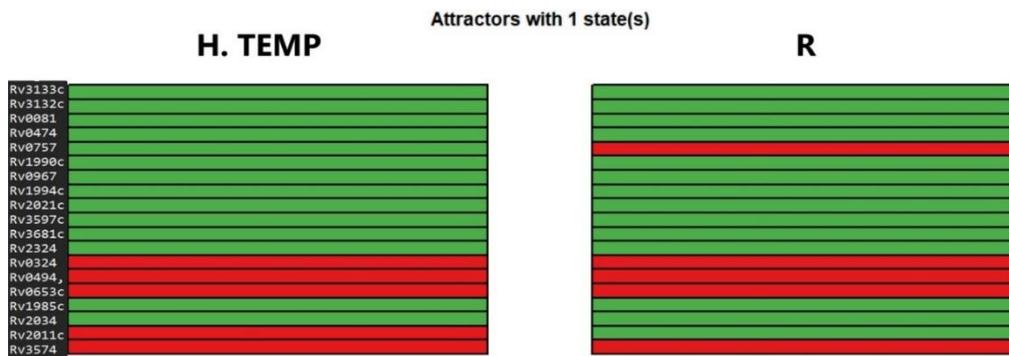


Fig. 46 Atractores del sistema para las mutaciones de activación de los nodos Rv0081, Rv3133c y Rv3132c.

2) La mutación de activación de los genes Rv0967, Rv3597c y Rv2034 presentan el mismo comportamiento anterior, como se observa en la tabla 14.

3) La mutación de activación del gen Rv0324 mantiene el atractor de Hipoxia Tardía. Además, se observa un atractor similar al de Reactivación y desaparece el atractor de Hipoxia Temprana, como se observa en la figura 47.

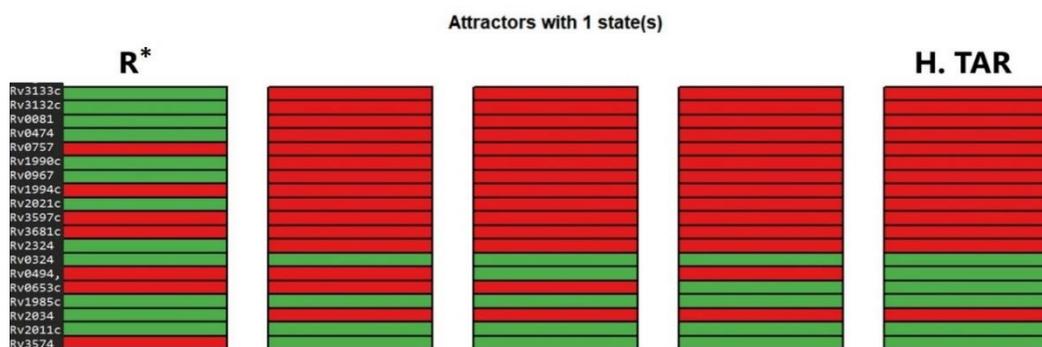


Fig. 47 Atractores del sistema con la mutación de activación del nodo Rv0324. Se observa la conservación del atractor de Hipoxia Tardía y un atractor similar al de Reactivación, con la diferencia en los nodos: Rv1994c, Rv3597c y Rv3681c, que se encuentran activos.

4) La mutación de activación del gen Rv3574 mantiene el atractor de Hipoxia Tardía y se observa una distorsión de los atractores de Reactivación y de Hipoxia Temprana (el nodo Rv1990c se encuentra inactivo en ambos atractores). También se evidencia una disminución en sus cuencas de atracción, de 407552 a 198144 para el atractor de Reactivación y de 114168 a 55808 para el atractor de Hipoxia Tardía, mientras que las cuencas de atracción de dos atractores similares al de Reactivación aumentan de 512 a 2048. Además, la cantidad de atractores del sistema aumenta hasta 22, como se observa en la figura 48.

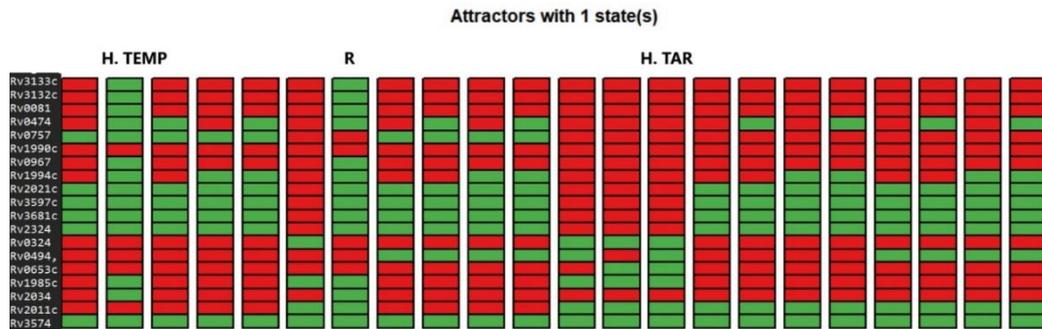


Fig. 48 Atractores del sistema con la mutación de activación del nodo Rv3574.

5) La mutación de activación del gen Rv2011c mantiene inalterados los atractores de Hipoxia Tardía y de Reactivación, mientras que desaparece el atractor de Hipoxia Temprana, como se observa en la figura 49.

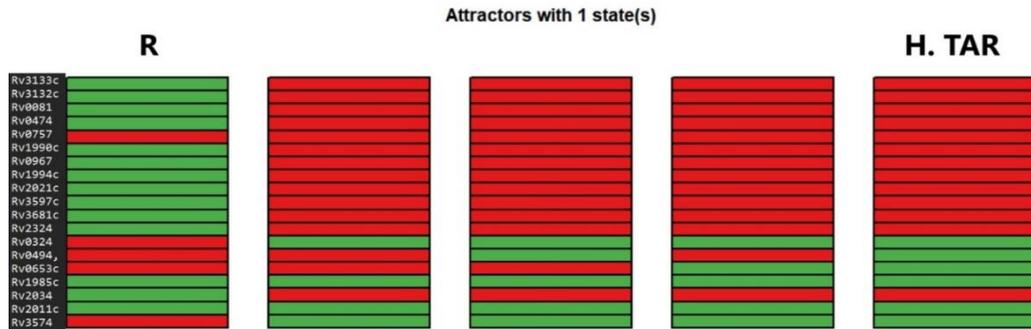


Fig. 49 Atractores del sistema con la mutación de activación del nodo Rv2011c.

Como se observa en la tabla 14, la evaluación de mutación de activación de los nodos Rv0474, Rv0757, Rv1990c, Rv1994c, Rv1985c, Rv2021c, Rv3681c, Rv2324, Rv0494 y Rv0653c no genera cambios significativos en los atractores del sistema. Esto está en concordancia con el hecho de que tienen un menor grado de conectividad de salida, por lo que regulan a pocos nodos, como se observa en la tabla 15.

GENE	Kout	Función Booleana
Rv3133c	6	Rv3132c (Rv0757 & Rv2034) (Rv3133c & (Rv3132c (Rv0757 & Rv2034)))
Rv3132c	1	Rv3133c Rv2034
Rv3574	4	(Rv3574 Rv0324) & (!(Rv0081 Rv3597c))
Rv3597c	10	(Rv1994c Rv3574 Rv2034 Rv0474 Rv3597c) & (!(Rv0324))
Rv0757	2	(Rv3681c Rv0081 Rv0757 Rv3597c) & (! Rv2011c)
Rv0081	14	(!(Rv3574) Rv3133c) & (!Rv0081)
Rv0324	6	(Rv0324 (!(Rv3597c & Rv1985c)) & (! Rv0081))
Rv2034	5	Rv0081 Rv2034 Rv0967
Rv0474	2	Rv0081 Rv0474 & (!Rv1985c)
Rv1990c	1	(Rv3597c Rv0081 Rv1990c) & ! (Rv3574)
Rv0967	1	Rv3133c Rv2034
Rv1994c	3	(Rv0081 Rv1994c) & (! Rv0324)
Rv2021c	0	Rv0081 Rv3597c Rv3133c
Rv3681c	1	Rv3597c (Rv0081 & !Rv2011c)
Rv2324	0	Rv0081 Rv3597c
Rv0494	1	(Rv0494) & (! Rv0081)
Rv0653c	1	(Rv0653c) & !(Rv3597c Rv0081)

Rv1985c	3	(Rv0081 Rv1985c Rv0324) & (! (Rv1994c Rv3597c Rv3133c))
Rv2011c	3	(Rv2011c Rv0324)

Tabla 15 Valores de conectividad de salida (Kout) para los 19 nodos de la red. La tabla muestra la cantidad la cantidad de regulaciones para cada nodo. Por ejemplo, el nodo Rv3681c solo regula al nodo Rv0757, por lo tanto, tiene un Kout de 1. En verde se muestran los nodos que no presentan cambios significativos en sus atractores luego de una mutación de activación. En la columna de la izquierda se visualizan las funciones reguladoras para cada nodo.

7.5.4. Verificación de mutaciones de inactivación

Siguiendo el procedimiento descrito en la sección 7.5.4 se realizaron las mutaciones de inactivación, que a diferencia del análisis de la sección 8.5.1, donde se eliminaba un nodo de la red, en este caso se lo mantiene, pero en estado inactivo.

1) La mutación de inactivación del nodo Rv0081 distorsiona la red y no se pueden alcanzar los atractores de Hipoxia Temprana y de Reactivación, pero se mantiene el atractor de Hipoxia Tardía. Además, se observa un aumento considerable en el número de atractores que llegan a ser 64. Ver figura 50.

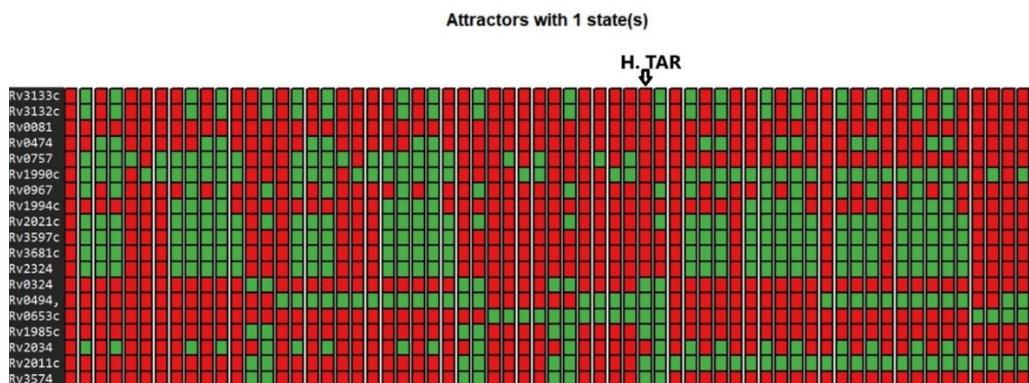


Fig. 50 Atractores de la red con mutación de inactivación del nodo Rv0081.

2) La mutación de inactivación de los nodos Rv3133c y Rv3132c mantiene el atractor de Hipoxia Tardía sin cambios y los atractores de Hipoxia Temprana y Reactivación con solo la variación del nodo mutado. Además, se observa un aumento del número de atractores que llegan a ser 10. Ver figura 51.

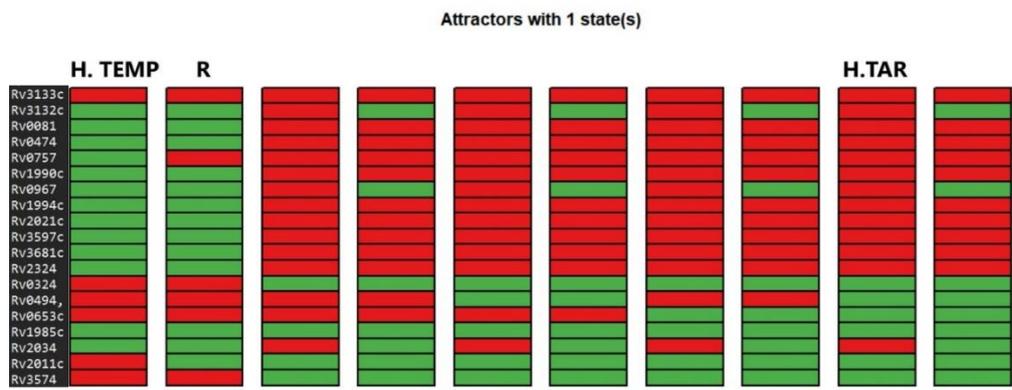


Fig. 51 Atractores de la red con mutación de inactivación del nodo Rv3133c.

3) La mutación de inactivación de los nodos Rv0474, Rv0757, Rv1990c, Rv0967, Rv1994c, Rv1985c, Rv2021c, Rv3597c, Rv3681c, Rv2324, Rv0494, Rv0653c y Rv2034 no afecta significativamente a los atractores del sistema como se evidencia en la tabla 13. Una particularidad de la mutación de los nodos Rv3597c y Rv2034 es que se elevan el número de atractores, a 54 y 10 respectivamente, ambos tienen un alto grado de conectividad de salida igual a cinco en la RDM.

4) La mutación de inactivación de los genes Rv0324 y Rv3574 conserva intactos la estructura de los atractores de Hipoxia Temprana y Reactivación, pero se pierde el atractor de Hipoxia Tardía, lo que refleja el hecho de que estos nodos son esenciales para alcanzar el estado general de Hipoxia Tardía. Ver figura 52.

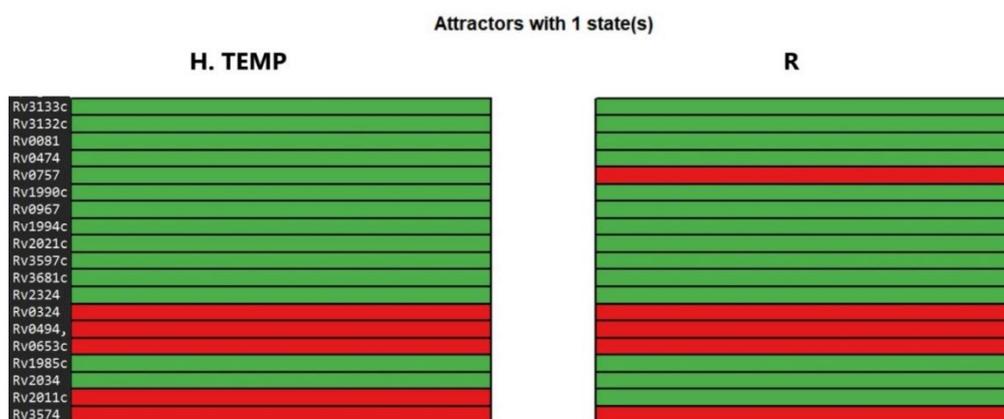


Fig. 52 Atractores de la red con mutación de inactivación de los nodos Rv0324 y Rv3574.

5) La mutación de inactivación del gen Rv2011c conserva en general la estructura de los atractores de Hipoxia Tardía y de Hipoxia Temprana, pero se pierde el atractor de Reactivación. Lo que sugiere que este nodo es esencial para alcanzar el estado general de Reactivación. Ver figura 53.

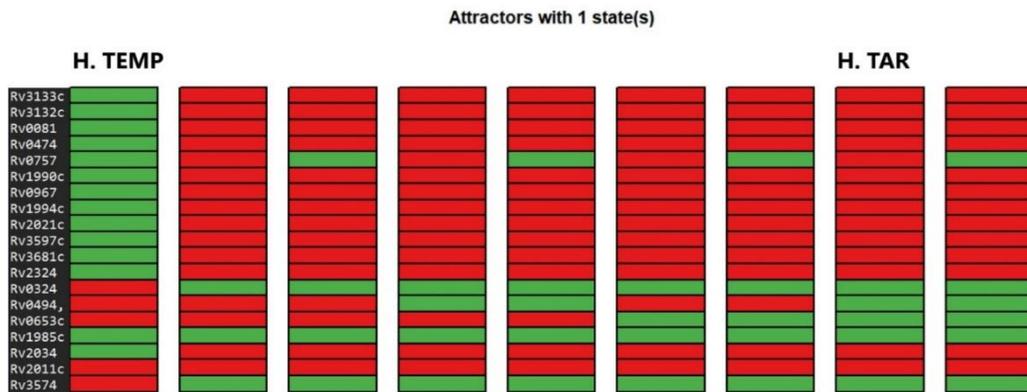


Fig. 53 Atractores de la red con mutación de inactivación del nodo Rv2011c.

Todos los resultados de las mutaciones de inactivación se resumen en la tabla

16.

GEN FIJADO EN ESTADO INACTIVO	NODOS DE VARIACIÓN EN LOS ATRACTORES			Cantidad de Atractores
	Hipoxia Temprana	Hipoxia Tardía	Reactivación	
Rv3133c	1	0	1	10
Rv3132c	1	0	2	10
Rv0081	AUSENTE	0	AUSENTE	64
Rv0474	1	0	1	6
Rv0757	1	0	0	6

Rv1990c	1	0	1	6
Rv0967	1	0	1	6
Rv1994c	1	0	1	6
Rv1985c	1	1	1	10
Rv2021c	1	0	1	6
Rv3597c	1	0	2	54
Rv3681c	1	0	1	6
Rv2324	1	0	1	6
Rv0324	0	AUSENTE	0	2
Rv0494	0	1	0	4
Rv0653c	0	1	0	4
Rv2034	<u>1</u>	<u>0</u>	<u>1</u>	<u>10*</u>
Rv2011c	0	1	AUSENTE	5
Rv3574	0	AUSENTE	0	22

*** 8 atractores de un estado y 2 atractores de dos estados.**

Tabla 16 Variación de los atractores al fijar en estado inactivo un nodo.

La tabla muestra la cantidad de nodos que cambian su estado en cada atractor particular al ejecutar la mutación de inactivación respecto de los atractores de la RDM sin mutaciones.

7.6.INTRODUCCIÓN DE UN NODO RELACIONADO CON LAS CITOQUINAS

7.6.1. Nodo relacionado a las citoquinas

Se realizó una búsqueda sistemática de los genes de MTB relacionados a las citoquinas del huésped y se encontró que la expresión del gen HspX (Rv2031c) está relacionada con la concentración relativa de IFN- γ y la IL 10, que son citoquinas con funciones inflamatoria y antiinflamatoria, respectivamente (68).

7.6.2. Análisis de la integración del nodo Rv2031c con la red dormancia

De la búsqueda bibliográfica se verifica que el gen Rv3133c se comporta como un activador del gen Rv2031c, por lo que constituye una función booleana activadora. Además, según la base de datos de Turkarslan y colaboradores (73), que contiene información de los genes que son activados o inhibidos por los factores de transcripción de MTB, con sus correspondientes valores de *p-value*, esta función activadora tiene un alto valor de inducción, mayor a cinco veces y un *p-value* de 2.93×10^{-26} , por lo tanto, es estadísticamente representativa.

7.6.3. Análisis bioinformático de la relación de Rv2031c con la red de dormancia.

Análisis de Componentes Principales:

Mediante este algoritmo de agrupamiento en Componentes Principales (CP), se puede determinar qué genes están más relacionados entre sí según sus perfiles de expresión genética. Para este análisis se usaron los datos experimentales descritos en la sección 8.1, tomados del trabajo de Peterson y Colaboradores (65). La cercanía en una posición particular del plano de componentes determina el grado de similitud entre los genes para las componentes principales analizadas. Como se observa en la figura 54, los genes Rv3133c y Rv2031c se encuentran muy cercanos entre sí, tanto en las componentes CP1 y CP2 como en las componentes CP1 y CP3. Esta proximidad sugiere una alta correlación entre estos genes, respaldando así la elección de Rv3133c como nodo regulador de Rv2031c.

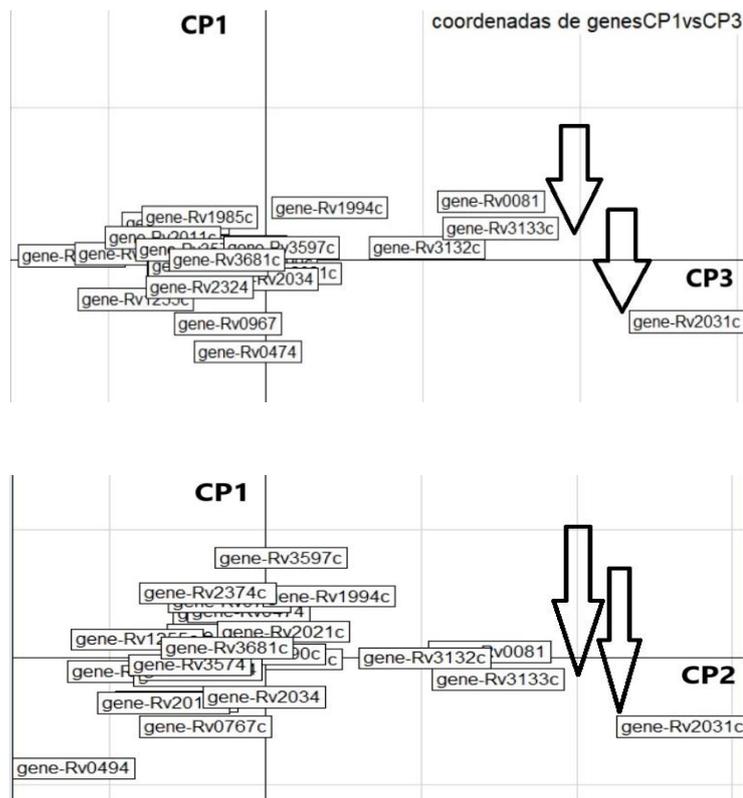


Fig. 54 Análisis de agrupamiento de los genes en componentes principales. Se observa la proximidad de los genes Rv2031c y Rv3133c, con valores cercanos a cero en CP1 y valores positivos en CP3 en el panel superior, así como valores positivos en CP2 en el panel inferior. La flecha indica la posición de las coordenadas de los genes Rv2031c y Rv3133c en el plano definido por las componentes principales CP1 y CP3 (panel superior) y CP1 y CP2 (panel inferior).

Análisis de funciones *in silico* del nodo Rv2031c:

Se cargaron los perfiles de expresión de Rv2031c en la red mínima de 19 nodos. El algoritmo “bestfit” se usó con K=2 (funciones con un máximo de 2 genes reguladores). Encontrándose al gen Rv3133c como su regulador en la segunda función reguladora generada por el algoritmo.

Transition functions:
Alternative transition functions for gene Rv2031c:
 $Rv2031c = (Rv3574) \mid (Rv2031c)$ (error: 1)
 $Rv2031c = (Rv3574) \mid (Rv3133c)$ (error: 1)

En base a estos resultados, la función de regulación propuesta para el gen Rv2031c es:

Rv2031c, Rv3133c

Tanto el análisis de componentes principales, como el análisis de funciones *in silico*, corroboran el dato experimental encontrado en la base de datos de Turkarslan y colaboradores (73), que asigna a Rv3133c como un regulador de Rv3031c.

Por lo tanto, la red de dormancia mínima, ampliada con este nodo, sería la siguiente:

GEN	FUNCIONES BOOLEANAS
Rv1994c	(Rv0081 Rv1994c) & (! Rv0324)
Rv3133c	Rv3132c (Rv0757 & Rv2034) (Rv3133c & (Rv3132c (Rv0757 & Rv2034)))
Rv3132c	Rv3133c Rv2034
Rv3574	(Rv3574 Rv0324 & !(Rv0081 Rv3597c))
Rv3597c	(Rv1994c Rv3574 Rv2034 Rv0474 Rv3597c) & !(Rv0324)
Rv0757	(Rv3681c Rv0081 Rv0757 Rv3597c) & (! Rv2011c)
Rv0081	!(Rv3574) Rv3133c)) & !(Rv0081 Rv1033c))
Rv0324	(Rv0324 !(Rv3597c & Rv1985c)) & (! Rv0081)
Rv1985c	(Rv0081 Rv1985c Rv0324) & (! (Rv1994c Rv3597c Rv3133c))
Rv2011c	(Rv2011c Rv0324)
Rv2034	Rv0081 Rv2034 Rv0967
Rv3681c	Rv3597c (Rv0081& !Rv2011c)
Rv0967	Rv3133c Rv2034
Rv0474	Rv0081 Rv0474 & !(Rv1985c)

Rv0494	(Rv0494) & (! Rv0081)
Rv0653c	(Rv0653c Rv1033c) & !(Rv3597c Rv0081)
Rv1990c	(Rv3597c Rv0081 Rv1990c) & ! (Rv3574)
Rv2021c	Rv0081 Rv3597c Rv3133c
Rv2324	Rv3597c Rv0081
Rv2031c	Rv3133c

Tabla 17 Funciones booleanas de la RMD ampliada con el nodo Rv2031c.

7.6.4. Análisis de los atractores de la red de dormancia ampliada con el nodo Rv2031c

1) Binarización de los estados de expresión diferencial del nodo Rv2031c.

Para poder obtener el estado experimental del gen Rv2031c en los estados generales de Hipoxia Temprana, Hipoxia Tardía y Reactivación, se binarizaron sus estados de expresión diferencial que se muestran en la figura 55.

Gen	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	threshold
Rv2031c	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	3.186

Tabla 18 Binarización de la expresión diferencial del gen Rv2031c.

2) Análisis de la Reproducción de los estados experimentales por parte de los atractores del modelo booleano de dormancia ampliado con el nodo Rv2031c.

Al obtenerse los atractores se la red de dormancia ampliada se generan los mismos seis atractores que los de la RMD. Al analizar los estados del nodo Rv2031c en la figura 56, se verifica que se encuentra activo en los estados generales de Hipoxia Temprana y Reactivación e inactivo en el estado de Hipoxia Tardía, reproduciendo sus estados experimentales descritos en la tabla 18 y en la figura 55. Además, se mantienen las mismas proporciones entre las cuencas de atracción de los atractores de Reactivación (815104 estados), Hipoxia Temprana (229376 estados) e Hipoxia Tardía (1024 estados), que las obtenidas en la red antes de la adición del nodo Rv2031c.

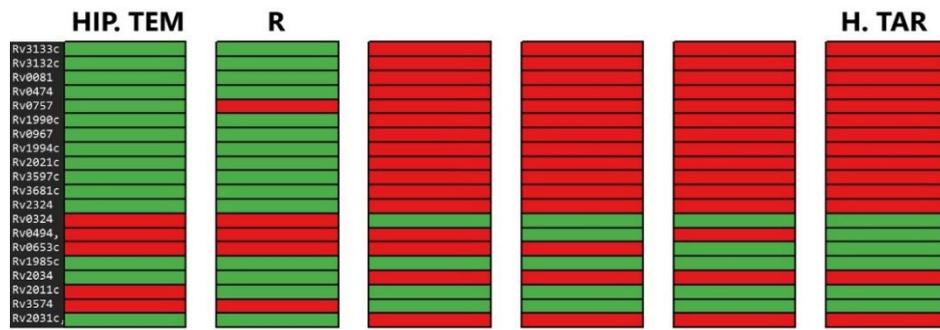


Fig. 56 Atractores de la red mínima de dormancia ampliada con el nodo Rv2031c.

La estructura de la RMD ampliada con el nodo Rv2031c se muestra en la figura 57.

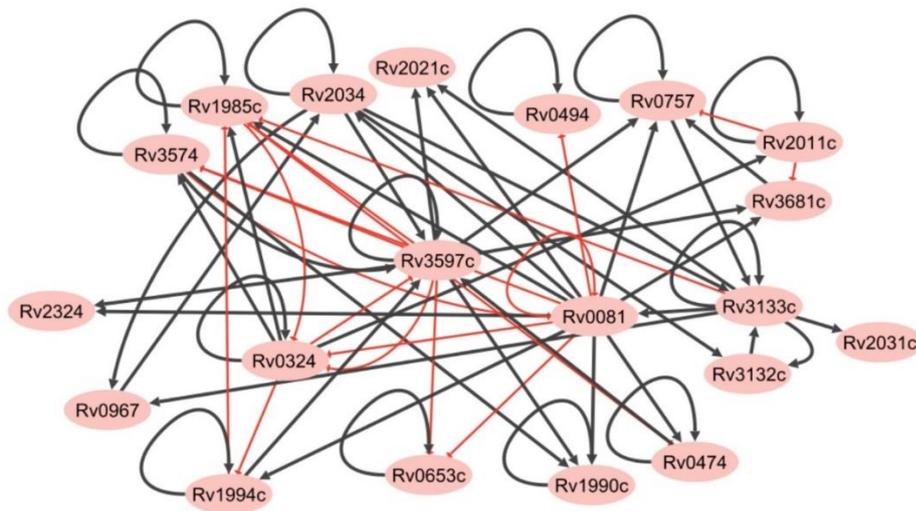


Fig. 57 red mínima de dormancia ampliada con el nodo Rv2031c.

7.7.COMPARACIÓN DE LA RED MÍNIMA DE DORMANCIA CON REDES ALEATORIAS DEL MISMO GRADO DE ENTRADA PARA LOS NODOS.

Con el algoritmo descrito en la sección 7.7 se procede a generar dos redes aleatorias de 20 nodos que tengan la misma conectividad de entrada que la RDM y se obtienen sus atractores.

Atractores de la Red Aleatoria 1:

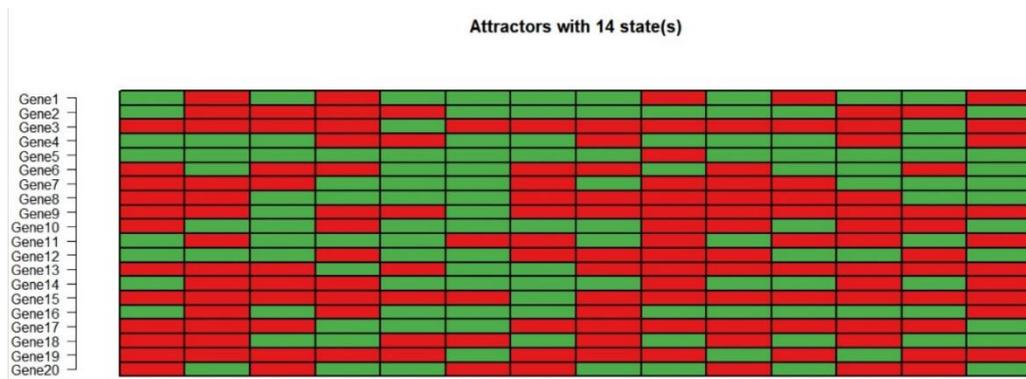


Fig. 58 Atractor de 14 estados de la red aleatoria 1.



Fig. 59 Atractor de 4 estados de la red aleatoria 1:



Fig. 60 Atractor de 7 estados de la red aleatoria 1

Atractores de la Red Aleatoria 2:

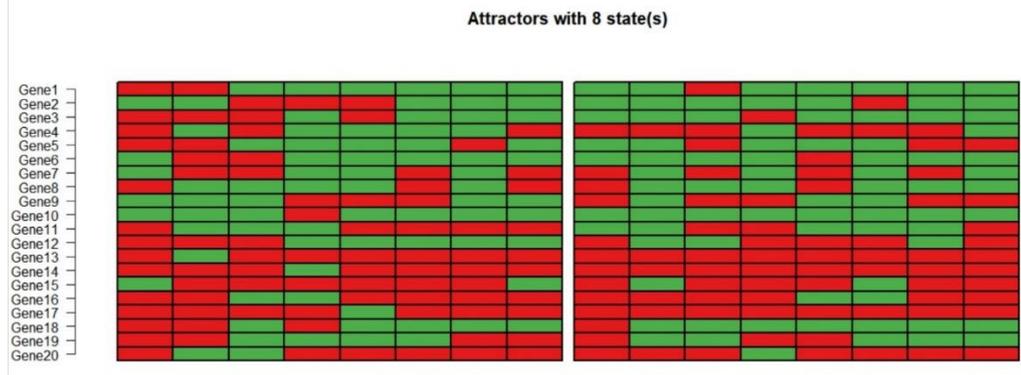


Fig. 61 Atractores de 8 estados de la red aleatoria 2

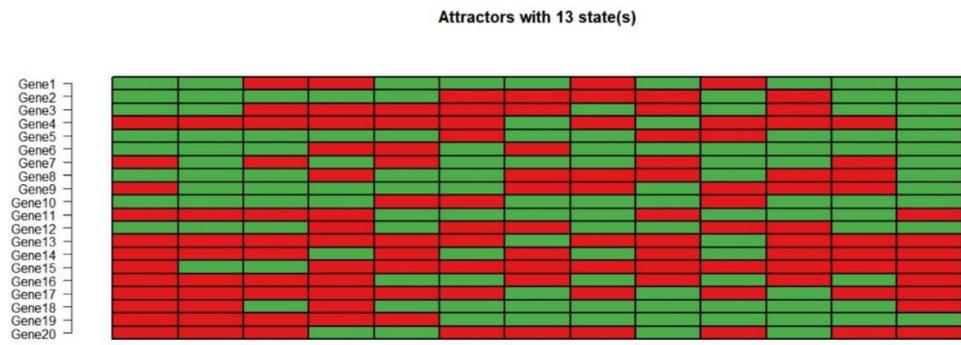


Fig. 62 Atractor de 13 estados de la red aleatoria 2

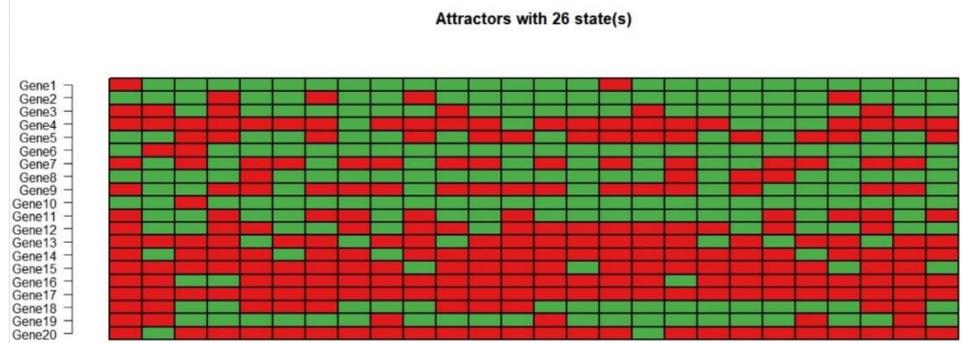


Fig. 63 Atractor de 26 estados de la red aleatoria 2

Al comparar los atractores de las dos redes aleatorias con los atractores del modelo de dormancia, en ningún caso obtenemos alguno de los atractores de Hipoxia Temprana, Reactivación o de Hipoxia Tardía de la red de dormancia mínima. Además, los atractores de estas redes aleatorias tienen estructuras muy diversas, algunas de los cuales son muy complejas a pesar de contar con el mismo número de nodos y de reguladores. En la tabla 19 se muestran la cantidad de atractores obtenidos para 15 redes aleatorias. Se observan atractores con una gran variedad de estados: 1, 2, 3, 4, 5, 6, 7, 8, 11, 13, 14, 18, 21, 22, 23, 26, 30, 46, 48, 55, 62, 93, 105, 131 y 139 estados, que no tienen ningún tipo de similitud con algunos de los atractores de la red mínima de dormancia, que consta de 6 atractores de un solo estado

Nº DE RED ALEATORIA	CANTIDAD DE ATRACTORES	ESTADOS DEL ATRACTOR
1	1	14
	1	4
	1	7
2	2	8
	1	13
	1	26
3	3	4
	2	6
	1	8

4	1	1
	1	2
	1	46
5	1	2
	1	11
	1	23
6	1	2
	1	8
	1	7
	1	30
7	1	1
	1	8
	1	21
8	2	6
	1	12
9	2	3
	1	8
	1	13
	1	17
	1	26
	1	131
	1	139
10	2	4
	1	55
11	3	2
	2	3
	1	8
	1	18
	1	30
	1	93
12	7	1
	2	2
	2	4
	1	5

	1	7
13	1	4
	1	7
14	3	2
	1	4
	1	48
	1	55
	1	62
	1	62
15	1	4
	1	22
	1	105

Tabla 19 Atractores de 15 redes aleatorias. En la primera columna se muestra el número de la red aleatoria. En la columna central figuran la cantidad de atractores con una cantidad particular de estados. Por ejemplo, la red aleatoria N°13 posee un atractor de 4 estados y un atractor de 7 estados.

7.8.INTRODUCCIÓN DE NODOS RELACIONADOS AL ESTADO DE REACTIVACIÓN.

7.8.1. Nodos relacionados al estado de reactivación

Luego de una búsqueda bibliográfica se determinó que los genes RpfA (Rv0867c), RpfB (Rv1009), RpfC (Rv1884c), RpfD (Rv2389c), RpfE (Rv2450c), FtsZ (Rv2150c) y ripA (Rv1477) están relacionados al proceso de reactivación (77, 78, 79, 80).

7.8.2. Análisis de la interacción de los nodos relacionados a la reactivación con la red de dormancia de MTB

Se realizó la búsqueda de las funciones reguladoras de los nodos RpfA (Rv0867c), RpfB (Rv1009), RpfC (Rv1884c), RpfD (Rv2389c), RpfE (Rv2450c), FtsZ (Rv2150c) y ripA (Rv1477) en las bases de datos de Turkarslan y colaboradores (73), disponibles en la plataforma “MTB Network Portal v2” y de Galagan y colaboradores (34) y solo se obtuvieron interacciones significativas para los nodos

Rv1884c, Rv2389c, Rv2450c y Rv2150c, sin embargo, fue necesario introducir a los nodos Rv1033c y Rv1353c dado que estos nodos regulan a los nodos relacionados a la reactivación y además son nodos que regulan a muchos otros factores de transcripción 607 y 1284 respectivamente según el trabajo de Galagan y colaboradores (34).

De igual forma que en la sección 7.4.1, sus perfiles de expresión diferencial se usaron para verificar sus funciones reguladoras.

En el anexo 5 se muestran sus perfiles de expresión diferencial.

Del análisis anterior se encontraron las funciones reguladoras para los nodos relacionados a la reactivación que se muestran en la tabla 20.

GEN	FUNCIONES BOOLEANAS
Rv1994c	(Rv0081 Rv1994c) & (! Rv0324)
Rv3133c	Rv3132c (Rv0757 & Rv2034) (Rv3133c & (Rv3132c (Rv0757 & Rv2034)))
Rv3132c	Rv3133c Rv2034
Rv3574	(Rv3574 Rv0324 & !(Rv0081 Rv3597c))
Rv3597c	(Rv1994c Rv3574 Rv2034 Rv0474 Rv3597c) & !(Rv0324)
Rv0757	(Rv3681c Rv0081 Rv0757 Rv3597c) & (! Rv2011c)
Rv0081	(!(Rv3574) Rv3133c)) & !(Rv0081 Rv1033c))
Rv0324	(Rv0324 !(Rv3597c & Rv1985c)) & (! Rv0081)
Rv1985c	(Rv0081 Rv1985c Rv0324) & (! (Rv1994c Rv3597c Rv3133c))
Rv2011c	(Rv2011c Rv0324)
Rv2034	Rv0081 Rv2034 Rv0967
Rv3681c	Rv3597c (Rv0081& !Rv2011c)
Rv0967	Rv3133c Rv2034
Rv0474	Rv0081 Rv0474 & !(Rv1985c)
Rv0494	(Rv0494) & (! Rv0081)
Rv0653c	(Rv0653c Rv1033c) & !(Rv3597c Rv0081)
Rv1990c	(Rv3597c Rv0081 Rv1990c) & ! (Rv3574)

Rv2021c	Rv0081 Rv3597c Rv3133c
Rv2324	Rv3597c Rv0081
Rv2031c	Rv3133c
Rv1033c	Rv1033c (Rv1353c & !(Rv0081 Rv1994c Rv0324 Rv3597c))
Rv2389c	(Rv0757 Rv0324 Rv1033c) & (!Rv0081)
Rv2450c	Rv0081 Rv0967 Rv2021c
Rv1884c	(Rv1353c Rv0967 Rv3597c) & !(Rv2011c)
Rv1353c	Rv1353c ((Rv0324 Rv0757) & ! Rv3574))
Rv2150c	Rv0967 (Rv1033c & !(Rv0081 Rv2011c))

Tabla 20 Funciones booleanas de la RMD ampliada con los nodos relacionados a la reactivación. Esta red de 26 nodos contiene el nodo relacionado a las citoquinas Rv2031c y los 6 nodos adicionales relacionados a la reactivación.

7.8.3. Análisis de los atractores de la red ampliada con los nodos relacionados a la reactivación

1) Binarización de los estados de expresión diferencial de los nodos relacionados a la Reactivación.

Usando los algoritmos descritos en la sección 7.2 se realizó la binarización de la expresión diferencial para los genes Rv1884c, Rv2389c, Rv2450c, Rv2150c, Rv1033c y Rv1353c como se muestra en la tabla 21. Estos estados experimentales deberían poder ser reproducidos por el modelo booleano de dormancia.

GENES	H. TEMP						H. TAR		R						thresholds
	13	14	15	16	17	18	19	20	21	22	23	24	25		
Rv1033c	0	0	0	0	0	0	1	1	1	0	0	0	0	0.16275	
Rv2389c	0	0	0	0	0	0	1	1	0	0	0	0	0	0.238557692307692	
Rv2450c	0	0	0	0	0	1	0	0	1	1	1	1	1	0.118298611111111	
Rv1884c	1	1	1	0	0	1	0	0	0	0	0	0	0	0.142875	
Rv1353c	0	0	0	1	1	0	1	1	1	0	0	0	0	0.338933823529412	
Rv2150c	1	1	1	1	1	0	0	0	1	1	1	1	0	0.431349206349206	

Tabla 21 Binarización de la expresión diferencial de los genes relacionados al estado general de Reactivación. Se muestran los estados binarizados de los nodos relacionados a la reactivación para las muestras que comprenden los estados generales de Hipoxia Temprana, Hipoxia Tardía y Reactivación.

2) Análisis de la Reproducción de los estados experimentales binarizados por parte de los atractores del modelo booleano de dormancia ampliado con los nodos relacionados a la Reactivación.

Una vez cargadas los nodos adicionales se procede a obtener los atractores de la red ampliada. Se obtuvieron 18 atractores, tres de los cuales representan a los estados generales de Hipoxia Temprana, Reactivación y de Hipoxia Tardía. Estos atractores son a su vez los que tienen las cuencas de atracción más grandes (cantidad de estados del sistema que se dirigen hacia el atractor). El atractor de Hipoxia Temprana tiene una cuenca de atracción que consta de 3538944 estados que representan al 10,55% del total de estados, el atractor de Reactivación posee una cuenca de atracción que consta de 12730368 estados que representan al 37,97% del total de estados y el atractor de Hipoxia Tardía tiene una cuenca de atracción que consta de 16384 estados que representan al 0.05 % de estados del sistema. En la figura 64 se muestran estos tres atractores.

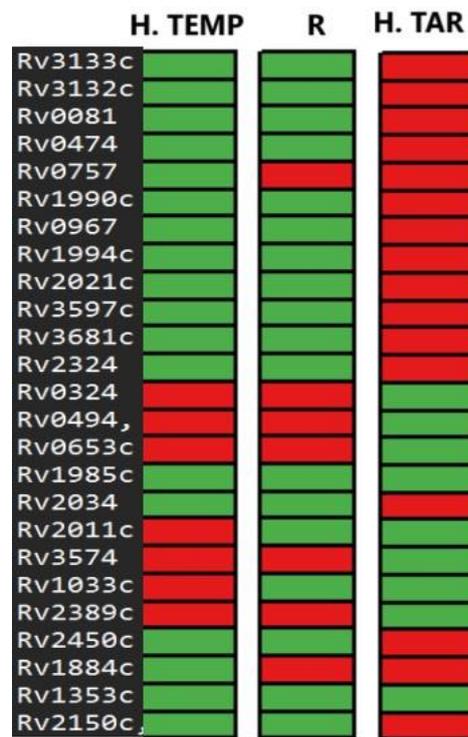


Fig. 64 Atractores de la red de dormancia ampliada con los nodos relacionados a la Reactivación. En el panel izquierdo se muestran todos los nodos

de la red. Se muestra a los tres atractores identificados con los estados generales de Hipoxia Temprana (H. TEMP), Reactivación (R) e Hipoxia Tardía (H. TAR).

Al analizar sus estados en los atractores y compararlos con sus estados experimentales binarizados de la tabla 17 se verifica lo siguiente:

El atractor de Reactivación reproduce los estados experimentales binarizados de los nodos Rv1033c, Rv2389c, Rv2450c, Rv1884c, Rv1353c y Rv2150c.

El atractor de Hipoxia Temprana reproduce los estados experimentales binarizados de los nodos Rv1033c, Rv2389c y Rv2150c.

El atractor de Hipoxia Tardía reproduce los estados experimentales binarizados de los nodos Rv1033c, Rv2389c, Rv2450c, Rv1884c, Rv1353c y Rv2150c.

Por lo tanto, el modelo booleano de dormancia puede reproducir el 100 % de los estados de experimentales binarizados de los seis nodos adicionales con los atractores de Reactivación e Hipoxia Tardía y el 50% de los estados experimentales binarizados de dichos nodos con el atractor de Hipoxia Temprana.

7.8.4. Análisis de trayectorias hacia el atractor de reactivación

Con el objetivo de determinar los principales genes asociados al proceso de Reactivación, se realizó un análisis de las trayectorias hacia dicho atractor. Se evaluaron los nodos Rv2011c, Rv0324 y Rv1985c que en la sección 8.5 mostraron

ser importantes para el estado de Reactivación, junto con los nodos adicionales: Rv1884c, Rv2389c, Rv2450c, Rv2150c, Rv1033c y Rv1353c, encontrándose que los nodos Rv2011c, Rv1033c, Rv1353c y Rv2389c son suficientes para alcanzar el atractor de Reactivación en solo cuatro pasos de tiempo como se muestra en las figuras 65 y 65.

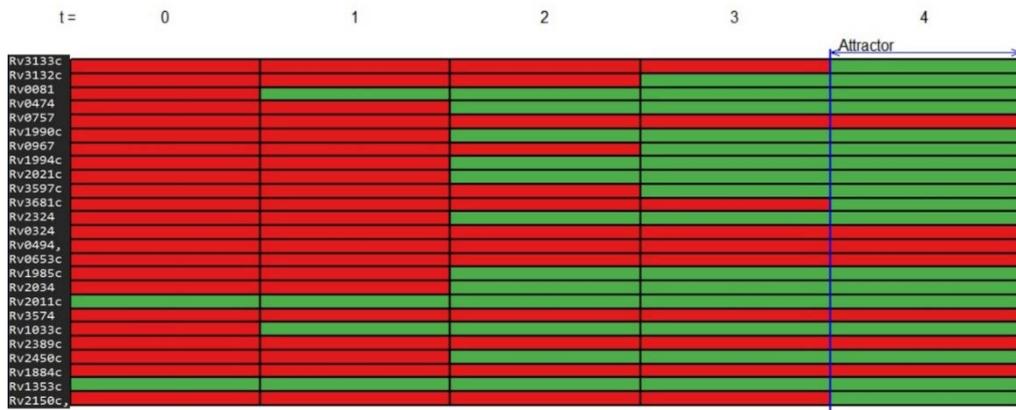


Fig. 65 Estados hasta el atractor partiendo de los nodos Rv2011c y Rv1353c activos.



Fig. 66 Estados hasta el atractor partiendo de los nodos Rv2011c, Rv1033c y Rv2389c activos

En la figura 67 se muestra la red de dormancia ampliada con la introducción de los nodos relacionados a la Reactivación, ubicados en el lado derecho de la red, mientras que los nodos con mayor conectividad como Rv0081, Rv3597c y Rv0324 se ubican en el centro de la gráfica.

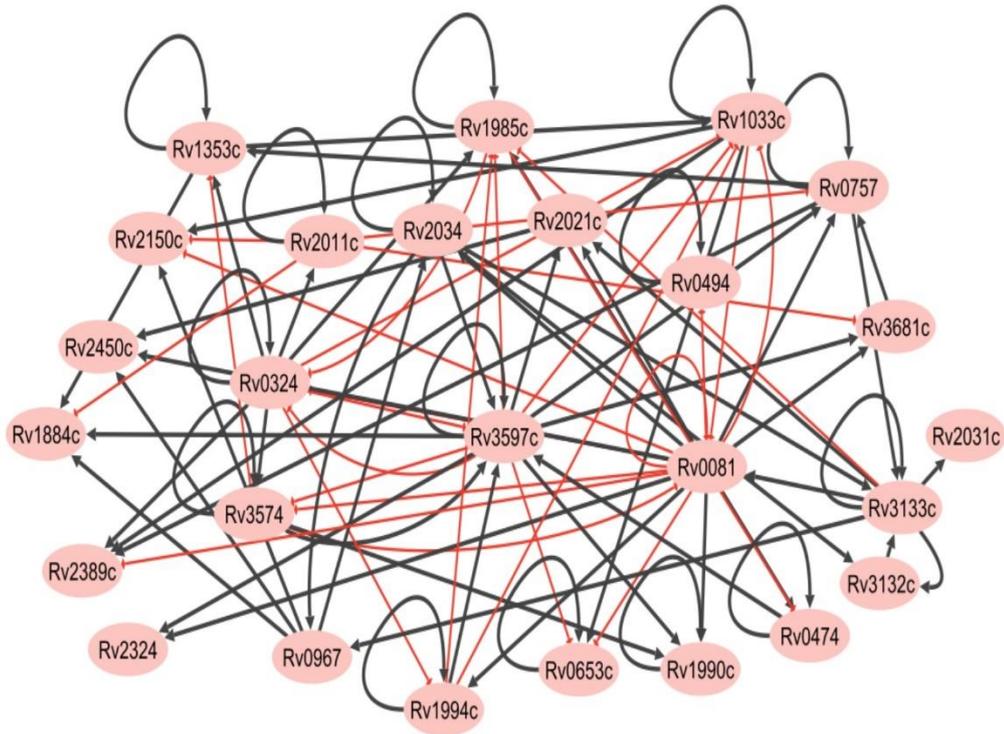


Fig. 67 Red de dormancia ampliada con los nodos relacionados a la Reactivación y el nodo relacionado a las citoquinas Rv2031c.

Estos resultados muestran que la red RDM de 19 nodos posee tres atractores relacionados a los estados experimentales binarizados de Hipoxia Temprana, Hipoxia Tardía y Reactivación. Al realizarse simulaciones de mutaciones y de trayectorias hacia el atractor, se ha podido determinar la importancia de nodos particulares para cada atractor, por ejemplo, el nodo Rv2011c es importante para alcanzar el atractor de Reactivación. Luego de introducirse el nodo Rv2031c, relacionado a las citoquinas, se observa que los atractores de la red RDM se mantienen. Finalmente, luego de introducirse los nodos relacionados a la Reactivación, la red se expande hasta

los 26 nodos. En esta red ampliada también se conservan dichos atractores y se verifica que los atractores de Hipoxia Tardía y de Reactivación son los que más se ajustan a los datos experimentales binarizados. Por lo tanto, esta red de 26 nodos, al contener los nodos implicados al proceso de reactivación, es la más adecuada para simular este proceso complejo, que le permite al patógeno volver a producir una enfermedad activa.

VIII. DISCUSIÓN

Partiendo de una red booleana previa, del trabajo de Bose y colaboradores (64), se modificaron las funciones reguladoras de los nodos con el fin de adecuarlos más cercanamente a los datos experimentales de estudios más recientes de expresión genética. Para hacer un análisis exhaustivo del comportamiento de esta red de dormancia mediante mutaciones de sus nodos y evaluación de las trayectorias desde un estado inicial, se procedió a reducir la red, retirando los nodos de salida y manteniendo los nodos hub; además, se realizó una corrección de las funciones reguladoras en base a los perfiles de expresión genética de los genes reguladores. Con estas correcciones y la reducción de nodos se obtuvo una red de dormancia mínima de MTB con seis atractores, tres de los cuales reflejan el comportamiento experimental de Hipoxia Temprana, Hipoxia Tardía y Reactivación.

Se utilizaron datos experimentales de expresión diferencial del trabajo de Peterson y colaboradores (65), para obtener *de novo* funciones reguladoras *in silico*. Estas funciones reguladoras *in silico* se compararon con las funciones reguladoras de los nodos que se derivaron de los trabajos de Galagan y colaboradores (34) y Turkarslan y colaboradores (73). Se observó que las funciones reguladoras *in silico* del gen Rv3574 son las que más se ajustan a las funciones derivadas de los datos experimentales. Por otro lado, mediante un análisis de componentes principales sobre los datos experimentales de expresión genética (65), se pudo evidenciar la

existencia de al menos tres estados generales que se corresponden con un estado general de Hipoxia temprana, otro de Hipoxia Tardía y uno de Reactivación. Estos datos luego fueron usados para comprobar la eficacia de la red de dormancia para reproducir estos estados experimentales, mediante el análisis de sus atractores.

Al realizar la comparación entre los estados de los 19 nodos de la RDM en los atractores con sus estados de expresión binarizados, se observó una alta concordancia. El atractor de Hipoxia temprana reproduce el estado experimental binarizado de 18 de los 19 nodos. El atractor de Hipoxia Tardía reproduce el estado binarizado de 16 de los 19 nodos, solo los genes Rv3681c, Rv2324 y Rv2034 tienen un estado opuesto al predicho por el atractor. Mientras que el atractor de Reactivación reproduce el estado experimental binarizado para 18 de los 19 nodos de la RDM, solo el gen Rv0324 muestra un estado opuesto al predicho por el atractor. Al analizar el perfil de expresión de este último, se observa una disminución del 50 % en su actividad lo que estaría de acuerdo con el estado inactivo que predice el atractor de Reactivación, por lo tanto, la discrepancia se podría atribuir a un valor alto del umbral asignado por el algoritmo de binarización.

Los análisis de eliminación de los nodos Rv0081 y Rv3133c mostraron una alteración en los atractores de Hipoxia Temprana y de Reactivación, pero solo en el caso de la eliminación del gen Rv0081 se evidenció una gran alteración del atractor de Hipoxia Tardía, esto está de acuerdo a la evidencia experimental que

sugiere que el gen Rv0081 es esencial para la transición entre el estado general de Hipoxia Temprana al de Hipoxia Tardía (35).

Los análisis de mutaciones de activación para los genes Rv0081, Rv3133c, Rv3132c y Rv0967 mostraron la conservación de los atractores de Hipoxia Temprana y de Reactivación, pero, a la vez, mostraron la pérdida del atractor de Hipoxia Tardía. Esto último está de acuerdo con la evidencia experimental que sugiere que estos nodos son importantes solo para la respuesta temprana a la hipoxia (5), (6). Las mutaciones de activación de los genes Rv0324 y Rv2011c mostraron la conservación del atractor de Hipoxia Tardía y la pérdida del atractor de Hipoxia Temprana, mientras que solo para el gen Rv2011c se mostró la conservación del atractor de Reactivación, estos resultados están de acuerdo con la evidencia experimental que sugiere que estos genes son fundamentales para la respuesta tardía a la hipoxia (34). Además, sugieren un papel preponderante del gen Rv2011c para la reactivación.

Los análisis de mutaciones de inactivación revelaron la importancia del nodo Rv0081 para obtener los atractores de Hipoxia Temprana, que se pierde al inactivar dicho nodo. Además, Rv0324 resultó ser fundamental para obtener el atractor de Hipoxia Tardía, que es el estado de mayor resistencia de MTB, mientras que el gen Rv2011c resultó esencial para obtener el atractor de Reactivación que se pierde al inactivar dicho gen.

El análisis de trayectorias desde un estado inicial se alinea con la evidencia experimental, que demuestra que los nodos Rv3133c y Rv0081 deben ser los primeros en activarse para poder alcanzar el atractor de Hipoxia Temprana (5, 6, 35). Este análisis también reveló la importancia de los genes Rv0324 y Rv3574 para poder alcanzar el atractor de Hipoxia Tardía, esto está de acuerdo con la evidencia experimental que asigna un rol central al gen Rv0324 en la respuesta tardía a la hipoxia (34), y al hecho de que el gen Rv3574 es fundamental para la adaptación del metabolismo lipídico, el cual es necesario para que MTB se adapte energéticamente al estado de hipoxia prolongado (82).

Los resultados del análisis de dos redes aleatorias de igual número de nodos y de reguladores que la RMD mostraron que en ningún caso se obtuvo alguno de los atractores característicos de Hipoxia Temprana, Hipoxia Tardía o Reactivación. Esto demuestra que los atractores característicos del modelo de dormancia son un producto específico de las regulaciones identificadas en el organismo MTB.

Uno de los nodos de la RDM es el factor HspX. Dicho factor está sobreexpresado durante la dormancia. En estudios previos, se ha demostrado que su expresión está regulada por la relación de concentración IL10/ IFN- γ (68). Del análisis de este nodo en la RDM se ha encontrado que este nodo crítico está directamente relacionado al sistema de dos componentes DosR/DevS (que es fundamental para el ingreso a la dormancia) mediante el nodo Rv3133c. Por lo tanto, estos resultados sugieren que HspX forma parte de un sistema de interacción entre la RDM y la red

de factores del sistema inmunológico, que en última instancia media la resucitación de MTB, conduciendo a la reactivación de la tuberculosis.

La ampliación de la RDM con los genes relacionados a la reactivación: RpfD (Rv2389c), RpfC (Rv1884c), RpfE (Rv2450c) y FtsZ (Rv2150c), se logró mediante la introducción de dos nodos adicionales: Rv1033c y Rv1353c. El análisis de los atractores de esta red de dormancia ampliada mostró una reproducción del 100 % de los estados experimentales binarizados de estos seis nodos para los atractores de Hipoxia Tardía y Reactivación. Sin embargo, el atractor de Hipoxia Temprana solo pudo reproducir los estados experimentales binarizados de los nodos Rv1033c, Rv1353c y Rv2150c.

Al analizarse las trayectorias hacia el atractor de Reactivación de la RDM ampliada, se encontró que los nodos Rv1033c, Rv1353c, Rv2011c y Rv2389c son esenciales para alcanzar dicho atractor.

Los resultados sugieren que la inactivación de Rv0324 le impediría al patógeno alcanzar el estado de Hipoxia Tardía. Además, trabajos previos han demostrado que este gen es fundamental en la resistencia a antibióticos (83). Por lo tanto, surge Rv0324 como un buen candidato de blanco terapéutico para evitar el estado de dormancia. El papel fundamental de los genes Rv1033c, Rv1353c, Rv2011c y Rv2389c en el establecimiento del atractor de Reactivación sugiere que estos nodos

podrían ser buenos candidatos como blancos terapéuticos durante la tuberculosis en estado de latencia, con el fin de evitar la reactivación.

El gen Rv1033c es parte de un sistema de transducción de señales de dos componentes similar al sistema DosR/DevS. Debido a su importancia en el establecimiento del atractor de Reactivación, podría estar implicado en las vías de señalización que permiten a MTB alcanzar este estado, constituyéndose así en otro candidato importante para ser un blanco terapéutico en la fase de latencia de la tuberculosis.

El hecho de que el gen RpfD (Rv2389c) sea el único factor promotor de resucitación activo durante el estado general de Hipoxia Tardía sugiere un mecanismo de reactivación que involucra su activación por parte de Rv1033c y que finalmente conduzca a la activación del gen RpfE (Rv2450c) y FtsZ (Rv2150c), este último es un factor implicado en la duplicación celular (78).

IX. CONCLUSIONES

- La red de dormancia mínima de MTB de 19 nodos posee tres atractores que reproducirían el comportamiento experimental de los estados binarizados de Hipoxia Temprana, Hipoxia Tardía y Reactivación.
- Del análisis de las trayectorias hacia un atractor, se sugiere la importancia de los nodos Rv0081 y DevR/S como iniciadores del estado general de Hipoxia Temprana.
- Del análisis de las mutaciones de los nodos se ha podido evidenciar la importancia de los genes Rv0324 y Rv3574 para alcanzar el estado binarizado de Hipoxia Tardía, por lo que serían posibles candidatos de blancos terapéuticos. Además, sugieren un papel preponderante del gen Rv2011c para alcanzar el estado general de Reactivación.
- Al introducir el nodo HspX, que está modulado por la relación de las concentraciones de interleuquinas provenientes del huésped, IL10/IFN- γ , se comprobó que la red reflejaría el comportamiento experimental para los estados de Hipoxia Temprana, Hipoxia Tardía y de Reactivación de este nodo.
- La incorporación de los nodos relacionados con la resucitación (RpfD, RpfC y RpfE), así como de los genes FtsZ, Rv1033c y Rv1353c, a la red de

dormancia ha demostrado que los atractores correspondientes a la Hipoxia Tardía y la Reactivación reproducirían el comportamiento experimental observado para estos seis nodos.

- A partir del análisis de las trayectorias de la red de dormancia ampliada con los genes asociados a la reactivación, se verifica la importancia de los genes Rv1033c y Rv2389c para alcanzar el estado general de Reactivación, lo que los convierte en posibles blancos terapéuticos para el tratamiento de la tuberculosis en estado de latencia, con el potencial de prevenir su reactivación.

X. DIRECCIONES A FUTURO

- Para mejorar el modelo se puede introducir niveles de expresión genética para los nodos de la red, en lugar de estados binarios.
- Se puede evaluar si la red de dormancia de MTB se encuentra operando en punto crítico. Para esto, se generarían mutaciones en la red con el objetivo de determinar los nodos esenciales para mantener el estado crítico de la red.
- Se podría ampliar la red, y así generar un modelo booleano multinivel que incluya nodos de las células inmunitarias del huésped, nodos mediadores como las citoquinas y nodos de la red de dormancia de MTB.

XI. LIMITACIONES

- La binarización de los estados de expresión genética no tiene en cuenta el espectro de modulaciones en los niveles de expresión genética.
- El modelado sincrónico no toma en cuenta algún tipo de retardo en la interacción entre los nodos de la red. Además, es determinista y no contempla la variabilidad natural en las poblaciones bacterianas.
- El modelo se basa en datos de expresión genética in vitro, sin tomar en cuenta algún factor del huésped que pueda alterar la expresión genética de MTB.
- El modelo de dormancia se basa en resultados experimentales de expresión genética en estado de dormancia inducido por hipoxia y solo en el estado de Hipoxia Tardía se considera adicionalmente los resultados de un modelo de dormancia por falta de potasio.
- El modelo no incluye interacciones de modificación post traduccional como fosforilación, metilación, entre los nodos.

XII. BIBLIOGRAFÍA

1. OMS. Global tuberculosis report 2022. Geneva; 2022.
2. Houben RMGJ, Dodd PJ. The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med.* octubre de 2016;13(10):e1002152.
3. Gong W, Wu X. Differential Diagnosis of Latent Tuberculosis Infection and Active Tuberculosis: A Key to a Successful Tuberculosis Control Strategy. *Front Microbiol.* 2021;12:745592.
4. Gupta A, Kaul A, Tsolaki AG, Kishore U, Bhakta S. Mycobacterium tuberculosis: immune evasion, latency and reactivation. *Immunobiology.* marzo de 2012;217(3):363-74.
5. Rustad TR, Harrell MI, Liao R, Sherman DR. The enduring hypoxic response of Mycobacterium tuberculosis. *PloS One.* 30 de enero de 2008;3(1):e1502.
6. Dasgupta N, Kapur V, Singh KK, Das TK, Sachdeva S, Jyothisri K, et al. Characterization of a two-component system, devR-devS, of Mycobacterium tuberculosis. *Tuber Lung Dis Off J Int Union Tuberc Lung Dis.* 2000;80(3):141-59.
7. Borah K, Xu Y, McFadden J. Dissecting Host-Pathogen Interactions in TB Using Systems-Based Omic Approaches. *Front Immunol.* 2021;12:762315.
8. Verma A, Kaur M, Singh LV, Aggarwal D, Verma I, Radotra BD, et al. Reactivation of latent tuberculosis through modulation of resuscitation promoting factors by diabetes. *Sci Rep.* 5 de octubre de 2021;11(1):19700.
9. Albert I, Thakar J, Li S, Zhang R, Albert R. Boolean network simulations for life scientists. *Source Code Biol Med.* 14 de noviembre de 2008;3:16.
10. Albert R, Thakar J. Boolean modeling: a logic-based dynamic approach for understanding signaling and regulatory networks and for making useful predictions. *Wiley Interdiscip Rev Syst Biol Med.* octubre de 2014;6(5):353-69.
11. Alvarez-Buylla ER, Chaos A, Aldana M, Benítez M, Cortes-Poza Y, Espinosa-Soto C, et al. Floral morphogenesis: stochastic explorations of a gene network epigenetic landscape. *PloS One.* 2008;3(11):e3626.

12. Albert R, Othmer HG. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *J Theor Biol.* 7 de julio de 2003;223(1):1-18.
13. Kim S, Kim J, Cho KH. Inferring gene regulatory networks from temporal expression profiles under time-delay and noise. *Comput Biol Chem.* agosto de 2007;31(4):239-45.
14. Akutsu T, Miyano S, Kuhara S. Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. *Pac Symp Biocomput Pac Symp Biocomput.* 1999;17-28.
15. Serra R, Villani M, Graudenzi A, Kauffman SA. Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data. *J Theor Biol.* 7 de junio de 2007;246(3):449-60.
16. Magombedze G, Mulder N. A mathematical representation of the development of *Mycobacterium tuberculosis* active, latent and dormant stages. *J Theor Biol.* 7 de enero de 2012;292:44-59.
17. Kolloli A, Subbian S. Host-Directed Therapeutic Strategies for Tuberculosis. *Front Med.* 2017;4:171.
18. Krug S, Parveen S, Bishai WR. Host-Directed Therapies: Modulating Inflammation to Treat Tuberculosis. *Front Immunol.* 2021;12:660916.
19. Chisholm RH, Tanaka MM. The emergence of latent infection in the early evolution of *Mycobacterium tuberculosis*. *Proc Biol Sci.* 25 de mayo de 2016;283(1831).
20. Hett EC, Rubin EJ. Bacterial growth and cell division: a mycobacterial perspective. *Microbiol Mol Biol Rev MMBR.* marzo de 2008;72(1):126-56, table of contents.
21. Ernst JD. Macrophage receptors for *Mycobacterium tuberculosis*. *Infect Immun.* abril de 1998;66(4):1277-81.
22. Bhat KH, Yaseen I. *Mycobacterium tuberculosis*: Macrophage Takeover and Modulation of Innate Effector Responses. En: Ribón W, editor. *Mycobacterium* [Internet]. Rijeka: IntechOpen; 2018. Disponible en: <https://doi.org/10.5772/intechopen.75003>

23. Nikitushkin VD, Demina GR, Kaprelyants AS. Rpf Proteins Are the Factors of Reactivation of the Dormant Forms of Actinobacteria. *Biochem Biokhimiia*. diciembre de 2016;81(13):1719-34.
24. Lähdesmäki H, Shmulevich I, Yli-Harja O. On Learning Gene Regulatory Networks Under the Boolean Network Model. *Mach Learn*. 1 de julio de 2003;52(1):147-67.
25. Galagan J, Lyubetskaya A, Gomes A. ChIP-Seq and the complexity of bacterial transcriptional regulation. *Curr Top Microbiol Immunol*. 2013;363:43-68.
26. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature*. 11 de junio de 1998;393(6685):537-44.
27. Jacques PE, Gervais AL, Cantin M, Lucier JF, Dallaire G, Drouin G, et al. MtbRegList, a database dedicated to the analysis of transcriptional regulation in *Mycobacterium tuberculosis*. *Bioinforma Oxf Engl*. 15 de mayo de 2005;21(10):2563-5.
28. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 20 de octubre de 1995;270(5235):467-70.
29. Balázsi G, Heath AP, Shi L, Gennaro ML. The temporal response of the *Mycobacterium tuberculosis* gene regulatory network during growth arrest. *Mol Syst Biol*. 2008;4:225.
30. Sanz J, Navarro J, Arbués A, Martín C, Marijuán PC, Moreno Y. The transcriptional regulatory network of *Mycobacterium tuberculosis*. *PloS One*. 2011;6(7):e22178.
31. Eoh H. Metabolomics: a window into the adaptive physiology of *Mycobacterium tuberculosis*. *Tuberc Edinb Scotl*. diciembre de 2014;94(6):538-43.
32. Wayne LG, Hayes LG. An in vitro model for sequential study of shutdown of *Mycobacterium tuberculosis* through two stages of nonreplicating persistence. *Infect Immun*. junio de 1996;64(6):2062-9.
33. Eoh H, Rhee KY. Multifunctional essentiality of succinate metabolism in adaptation to hypoxia in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A*. 16 de abril de 2013;110(16):6554-9.

34. Galagan JE, Minch K, Peterson M, Lyubetskaya A, Azizi E, Sweet L, et al. The Mycobacterium tuberculosis regulatory network and hypoxia. *Nature*. 11 de julio de 2013;499(7457):178-83.
35. Sun X, Zhang L, Jiang J, Ng M, Cui Z, Mai J, et al. Transcription factors Rv0081 and Rv3334 connect the early and the enduring hypoxic response of Mycobacterium tuberculosis. *Virulence*. 2018;9(1):1468-82.
36. Voskuil MI, Schnappinger D, Rutherford R, Liu Y, Schoolnik GK. Regulation of the Mycobacterium tuberculosis PE/PPE genes. *Tuberc Edinb Scotl*. 2004;84(3-4):256-62.
37. Grundner C, Ng HL, Alber T. Mycobacterium tuberculosis protein tyrosine phosphatase PtpB structure reveals a diverged fold and a buried active site. *Struct Lond Engl* 1993. noviembre de 2005;13(11):1625-34.
38. Yaseen I, Kaur P, Nandicoori VK, Khosla S. Mycobacteria modulate host epigenetic machinery by Rv1988 methylation of a non-tail arginine of histone H3. *Nat Commun*. 16 de noviembre de 2015;6:8922.
39. Li J, Chai QY, Zhang Y, Li BX, Wang J, Qiu XB, et al. Mycobacterium tuberculosis Mce3E suppresses host innate immune responses by targeting ERK1/2 signaling. *J Immunol Baltim Md* 1950. 15 de abril de 2015;194(8):3756-67.
40. Chai Q, Wang L, Liu CH, Ge B. New insights into the evasion of host innate immunity by Mycobacterium tuberculosis. *Cell Mol Immunol*. septiembre de 2020;17(9):901-13.
41. Houben D, Demangel C, van Ingen J, Perez J, Baldeón L, Abdallah AM, et al. ESX-1-mediated translocation to the cytosol controls virulence of mycobacteria. *Cell Microbiol*. agosto de 2012;14(8):1287-98.
42. Kim JK, Yuk JM, Kim SY, Kim TS, Jin HS, Yang CS, et al. MicroRNA-125a Inhibits Autophagy Activation and Antimicrobial Responses during Mycobacterial Infection. *J Immunol Baltim Md* 1950. 1 de junio de 2015;194(11):5355-65.
43. Wang J, Li BX, Ge PP, Li J, Wang Q, Gao GF, et al. Mycobacterium tuberculosis suppresses innate immunity by coopting the host ubiquitin system. *Nat Immunol*. marzo de 2015;16(3):237-45.
44. Peddireddy V, Doddam SN, Qureshi IA, Yerra P, Ahmed N. A putative nitroreductase from the DosR regulon of Mycobacterium tuberculosis induces pro-

- inflammatory cytokine expression via TLR2 signaling pathway. *Sci Rep.* 20 de abril de 2016;6:24535.
45. Gautam US, Mehra S, Ahsan MH, Alvarez X, Niu T, Kaushal D. Role of TNF in the altered interaction of dormant *Mycobacterium tuberculosis* with host macrophages. *PloS One.* 2014;9(4):e95220.
 46. Ritter K, Rousseau J, Hölscher C. The Role of gp130 Cytokines in Tuberculosis. *Cells.* 15 de diciembre de 2020;9(12).
 47. Esin S, Batoni G. Natural killer cells: a coherent model for their functional role in *Mycobacterium tuberculosis* infection. *J Innate Immun.* 2015;7(1):11-24.
 48. Essone PN, Leboueny M, Maloupazoa Siawaya AC, Alame-Emane AK, Aboumegone Biyogo OC, Dapnet Tadatsin PH, et al. *M. tuberculosis* infection and antigen specific cytokine response in healthcare workers frequently exposed to tuberculosis. *Sci Rep.* 3 de junio de 2019;9(1):8201.
 49. Luo Y, Xue Y, Mao L, Lin Q, Tang G, Song H, et al. Activation Phenotype of *Mycobacterium tuberculosis*-Specific CD4(+) T Cells Promoting the Discrimination Between Active Tuberculosis and Latent Tuberculosis Infection. *Front Immunol.* 2021;12:721013.
 50. Dawa S, Menon D, Arumugam P, Bhaskar AK, Mondal M, Rao V, et al. Inhibition of Granuloma Triglyceride Synthesis Imparts Control of *Mycobacterium tuberculosis* Through Curtailed Inflammatory Responses. *Front Immunol.* 2021;12:722735.
 51. Bar-yam Y. *Dynamics Of Complex Systems* [Internet]. 1st Edition. CRC Press; 1999. Disponible en: <https://doi.org/10.1201/9780429034961>
 52. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. Complex networks: Structure and dynamics. *Phys Rep.* 1 de febrero de 2006;424(4):175-308.
 53. Albert R, Barabási AL. Statistical mechanics of complex networks. *Rev Mod Phys.* enero de 2002;74(1):47-97.
 54. Mendenhall W, Beaver RJ, Beaver BM. *Introduction to probability and statistics.* 12th ed. Belmont, CA: Thomson/Brooks/Cole Belmont, CA; 2006.
 55. Waddington CH. *The Strategy of the Genes.* London: Routledge; 2014. 274 p.

56. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. *J Theor Biol.* 1 de marzo de 1969;22(3):437-67.
57. Aldana M. Boolean dynamics of networks with scale-free topology. *Phys Nonlinear Phenom.* 15 de octubre de 2003;185(1):45-66.
58. Aldana M, Balleza E, Kauffman S, Resendiz O. Robustness and evolvability in genetic regulatory networks. *J Theor Biol.* 7 de abril de 2007;245(3):433-48.
59. Balleza E, Alvarez-Buylla ER, Chaos A, Kauffman S, Shmulevich I, Aldana M. Critical dynamics in genetic regulatory networks: examples from four kingdoms. *PLoS One.* 18 de junio de 2008;3(6):e2456.
60. Huang S, Ernberg I, Kauffman S. Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. *Semin Cell Dev Biol.* septiembre de 2009;20(7):869-76.
61. Martinez-Sanchez ME, Mendoza L, Villarreal C, Alvarez-Buylla ER. A Minimal Regulatory Network of Extrinsic and Intrinsic Factors Recovers Observed Patterns of CD4+ T Cell Differentiation and Plasticity. *PLoS Comput Biol.* junio de 2015;11(6):e1004324.
62. Ramirez R, Herrera AM, Ramirez J, Qian C, Melton DW, Shireman PK, et al. Deriving a Boolean dynamics to reveal macrophage activation with in vitro temporal cytokine expression profiles. *BMC Bioinformatics.* 18 de diciembre de 2019;20(1):725.
63. Hegde SR, Rajasingh H, Das C, Mande SS, Mande SC. Understanding communication signals during mycobacterial latency through predicted genome-wide protein interactions and boolean modeling. *PLoS One.* 2012;7(3):e33893.
64. Bose T, Das C, Dutta A, Mahamkali V, Sadhu S, Mande SS. Understanding the role of interactions between host and Mycobacterium tuberculosis under hypoxic condition: an in silico approach. *BMC Genomics.* 27 de julio de 2018;19(1):555.
65. Peterson EJ, Abidi AA, Arrieta-Ortiz ML, Aguilar B, Yurkovich JT, Kaur A, et al. Intricate Genetic Programs Controlling Dormancy in Mycobacterium tuberculosis. *Cell Rep.* 28 de abril de 2020;31(4):107577.
66. Ahmed A, Rakshit S, Vyakarnam A. HIV-TB co-infection: mechanisms that drive reactivation of Mycobacterium tuberculosis in HIV infection. *Oral Dis.* abril de 2016;22 Suppl 1:53-60.

67. Keane J, Bresnihan B. Tuberculosis reactivation during immunosuppressive therapy in rheumatic diseases: diagnostic and therapeutic strategies. *Curr Opin Rheumatol.* julio de 2008;20(4):443-9.
68. Jee B, Sharma P, Katoch K, Joshi B, Awasthi SK. IL-10 down-regulates the expression of survival associated gene hspX of *Mycobacterium tuberculosis* in murine macrophage. *Braz J Infect Dis Off Publ Braz Soc Infect Dis.* agosto de 2017;21(4):386-90.
69. Ahmed M, Mackenzie J, Tezera L, Krause R, Truebody B, Garay-Baquero D, et al. *Mycobacterium tuberculosis* senses host Interferon- γ via the membrane protein MmpL10. *Commun Biol.* 1 de diciembre de 2022;5(1):1317.
70. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
71. Celebi ME, Kingravi HA, Vela PA. A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm. *CoRR [Internet].* 2012;abs/1209.1960. Disponible en: <http://arxiv.org/abs/1209.1960>
72. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. *J Stat Softw [Internet].* 2008;025(i01). Disponible en: <https://EconPapers.repec.org/RePEc:jss:jstsof:v:025:i01>
73. Turkarlan S, Peterson EJR, Rustad TR, Minch KJ, Reiss DJ, Morrison R, et al. A comprehensive map of genome-wide gene regulation in *Mycobacterium tuberculosis*. *Sci Data.* 2015;2:150010.
74. Saadatpour A, Albert Ré, Reluga TC. A REDUCTION METHOD FOR BOOLEAN NETWORK MODELS PROVEN TO CONSERVE ATTRACTORS. *SIAM J Appl Dyn Syst.* 2013;12(4):1997-2011.
75. Majumdar SD, Sharma D, Vashist A, Kaur K, Taneja NK, Chauhan S, et al. Co-expression of DevR and DevR(N)-Aph proteins is associated with hypoxic adaptation defect and virulence attenuation of *Mycobacterium tuberculosis*. *PLoS One.* febrero de 2010;5(2):e9448.
76. Hu Y, Movahedzadeh F, Stoker NG, Coates ARM. Deletion of the *Mycobacterium tuberculosis* alpha-crystallin-like hspX gene causes increased bacterial growth in vivo. *Infect Immun.* febrero de 2006;74(2):861-8.
77. Gupta RK, Srivastava R. Resuscitation promoting factors: a family of microbial proteins in survival and resuscitation of dormant mycobacteria. *Indian J Microbiol.* junio de 2012;52(2):114-21.

78. Sharma D, Bose A, Shakila H, Das TK, Tyagi JS, Ramanathan VD. Expression of mycobacterial cell division protein, FtsZ, and dormancy proteins, DevR and Acr, within lung granulomas throughout guinea pig infection. *FEMS Immunol Med Microbiol.* diciembre de 2006;48(3):329-36.
79. Hett EC, Chao MC, Steyn AJ, Fortune SM, Deng LL, Rubin EJ. A partner for the resuscitation-promoting factors of *Mycobacterium tuberculosis*. *Mol Microbiol.* noviembre de 2007;66(3):658-68.
80. Verma A, Ghoshal A, Dwivedi VP, Bhaskar A. Tuberculosis: The success tale of less explored dormant *Mycobacterium tuberculosis*. *Front Cell Infect Microbiol.* 2022;12:1079569.
81. Ignatov DV, Salina EG, Fursov MV, Skvortsov TA, Azhikina TL, Kaprelyants AS. Dormant non-culturable *Mycobacterium tuberculosis* retains stable low-abundant mRNA. *BMC Genomics.* 16 de noviembre de 2015;16(1):954.
82. Kendall SL, Withers M, Soffair CN, Moreland NJ, Gurucha S, Sidders B, et al. A highly conserved transcriptional repressor controls a large regulon involved in lipid degradation in *Mycobacterium smegmatis* and *Mycobacterium tuberculosis*. *Mol Microbiol.* agosto de 2007;65(3):684-99.
83. Peterson EJ, Ma S, Sherman DR, Baliga NS. Network analysis identifies Rv0324 and Rv0880 as regulators of bedaquiline tolerance in *Mycobacterium tuberculosis*. *Nat Microbiol.* 6 de junio de 2016;1(8):16078.

XIII. ANEXOS

ANEXO 1. DISTRIBUCIONES DE POISSON Y BINOMIAL

Distribución de Poisson

Sea μ el número promedio de eventos que ocurre en un estado de tiempo, la probabilidad de que ocurran k eventos está dada por:

$$P(x=k) = \mu^k e^{-\mu} / k!$$

para valores de $k = 0, 1, 2, 3, \dots$

La media y la desviación estándar de la variable aleatoria de Poisson son:

Media: $m = \mu$

Desviación estándar: $\sigma = \sqrt{\mu}$

Calculando los valores de $p(x)$, se pueden usar para construir un histograma de probabilidad para la variable aleatoria x . Ver figura S1

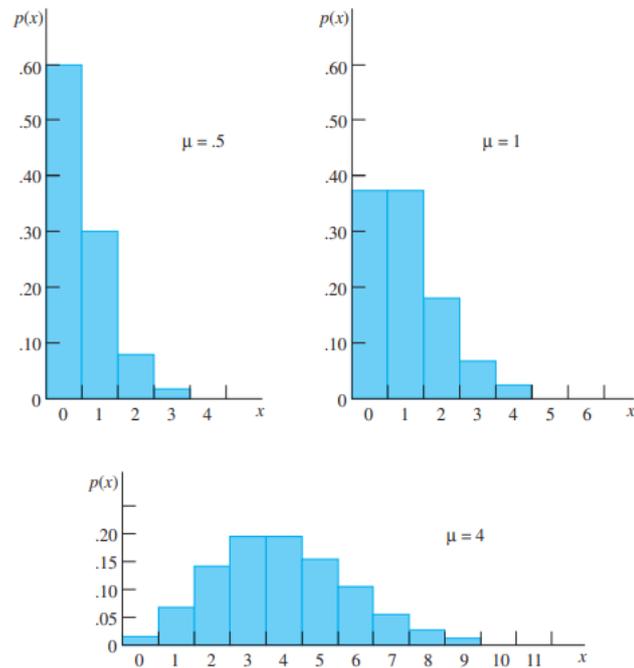


Fig.S1 Distribución de Poisson. Se muestra la distribución para $\mu = 0.5, 1$ y 4. Tomado de Introduction to Probability and Statistics (61).

Distribución Binomial:

Un experimento binomial consta de n intentos idénticos con probabilidad p de éxito en cada intento. La probabilidad de k éxitos en n intentos es:

$$P(x=k) = \binom{n}{k} p^k q^{n-k}$$

Dado que $q = 1-p$, se puede demostrar que el promedio μ está dado por $\mu = np$.

Al graficar estas distribuciones se comprueba que los valores promedios μ de estas distribuciones son representativas de la distribución. Ver figura S2.

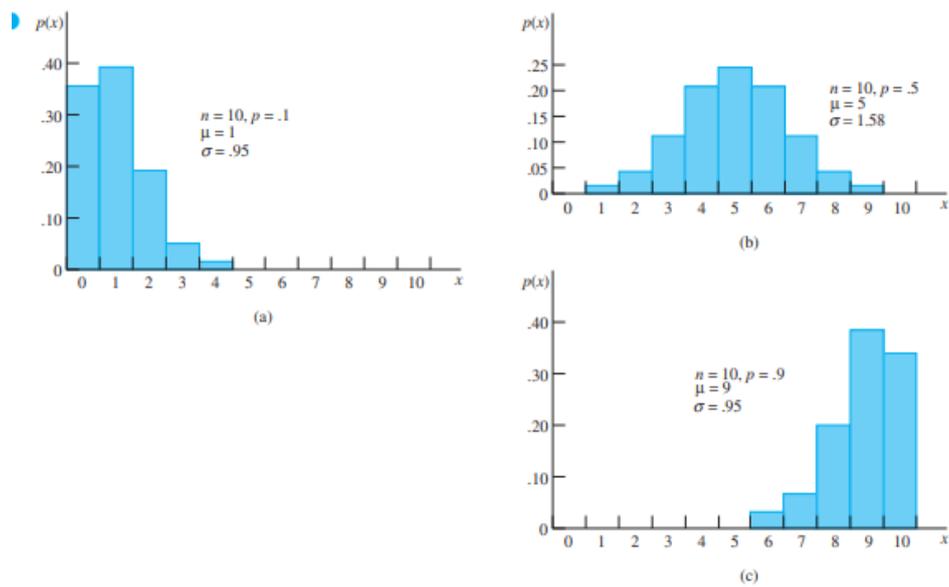


Fig.S2 Distribución Binomial. (a) El gráfico está sesgado a la derecha; es decir, la mayor parte del tiempo observará valores pequeños de x . La media o “punto de equilibrio” está alrededor de $x = 1$; (b) y (c) muestran otras dos distribuciones binomiales con $n = 10$, pero con diferentes valores de P . Cuando $p = 0.5$ la distribución es exactamente simétrica con respecto a la media, $m = np = 10(0.5) = 5$. Cuando $p = 0.9$, la distribución es la “imagen especular” de la distribución para $p = 0.1$ y es sesgado a la izquierda. Tomado de Introduction to Probability and Statistics (61).

ANEXO 2. VALORES DE EXPRESIÓN DIFERENCIAL

Las tablas que se presentan a continuación contienen los datos de expresión diferencial expresados en “log2ratio”, obtenidos como se indica la sección 7.1.4 con el algoritmo DESeq aplicado a las lecturas de RNAseq de 25 muestras tomadas de un cultivo de MTB sometido a las condiciones experimentales descritas en la tabla 2, como se describe en el trabajo de Peterson y colaboradores (65)

	MUESTRAS
--	----------

GENES	17h.- 57.6 % OD	23h.- 42.8 % OD	40h.- 11.5 % OD	41h.- 8.8 % OD	42h.- 7.49 % OD	43h.- 6.16 % OD	44h.- 4.31 % OD	45h.- 2.93 % OD	46h.- 0.41 % OD	47h.- 0.0 % OD	48.- 0.0 % OD	49h.- 0.0 % OD	50h.- 0.0 % OD
	1	2	3	4	5	6	7	8	9	10	11	12	13
Rv0081	0.72	-0.26	1.64	2.33	2.35	3.12	3.34	2.96	3.38	3.92	2.93	4.28	4.18
Rv0324	0.13	-0.79	-0.65	-1.51	-0.53	-0.41	-0.58	0.95	0.38	-1.08	1.07	-0.29	-0.82
Rv0348	0.09	0.34	0.21	0.22	0.16	-0.12	0.33	0.11	0.17	0.55	0.61	0.80	0.82
Rv0474	-1.75	-1.01	-0.28	-0.08	0.35	0.75	0.27	-0.18	0.16	0.71	0.21	1.14	1.10
Rv0494	-0.01	-2.43	-2.90	-2.96	-1.25	-3.10	-2.94	0.88	0.34	-2.72	0.93	-2.06	-2.95
Rv0653c	-0.08	-0.85	-1.23	-0.86	-0.74	-1.35	-1.76	-0.32	-0.25	-1.15	0.05	-0.75	-0.68
Rv0757	0.11	0.42	0.57	0.44	0.31	0.35	0.36	0.24	0.03	0.55	0.22	0.51	0.69
Rv0767c	-1.04	-0.89	-1.20	-1.36	-0.67	-0.82	-1.03	1.27	0.92	-1.20	1.16	-1.46	-0.56
Rv0827c	-0.23	0.20	-0.42	-0.31	0.24	-0.19	0.71	0.44	0.71	1.20	0.84	0.68	1.01
Rv0967	-0.87	-0.76	-0.64	-0.66	-0.28	-0.52	-0.44	-0.58	-0.69	-0.10	-1.10	-0.18	-0.11
Rv1255c	-0.58	-0.85	-0.89	-1.06	-1.15	-1.27	-1.08	-0.66	-1.07	-1.29	-1.05	-1.21	-1.19
Rv1985c	0.18	-0.13	-0.36	-0.14	-0.23	-0.70	-0.46	0.70	0.21	-0.30	0.96	-0.32	-0.17
Rv1990c	-0.67	-0.01	0.12	-0.30	0.18	-0.02	0.07	0.42	0.41	0.22	0.48	0.65	0.75
Rv1994c	0.65	1.84	1.27	0.79	0.93	0.94	1.00	0.65	0.90	1.55	0.53	1.44	1.42
Rv2011c	0.17	-0.64	-1.36	-1.65	-0.62	-1.32	-1.45	0.30	-0.14	-1.10	1.08	-1.01	-1.20
Rv2021c	-0.16	0.28	0.10	0.02	0.14	-0.02	0.02	0.34	-0.21	-0.07	-0.03	0.58	0.89
Rv2034	-0.74	-0.69	-0.63	-0.78	-0.12	-1.46	-0.73	0.60	0.48	-0.29	1.04	0.52	0.32

Rv2324	-0.84	-0.63	-0.98	-0.82	-0.52	-0.83	-0.61	0.03	-0.13	-0.89	0.22	-0.45	-0.18
Rv2374c	-0.16	0.20	0.24	0.30	0.12	0.39	0.40	-0.09	-0.17	0.20	-0.41	0.13	0.11
Rv3132c	-0.09	-0.11	0.27	0.69	0.89	1.42	2.17	2.32	3.29	3.86	3.33	4.36	4.34
Rv3133c	-0.10	0.20	0.27	0.99	1.43	2.06	2.88	2.90	4.13	4.68	4.51	5.41	5.35
Rv3574	-0.41	-0.59	-0.67	-0.34	-0.10	0.00	0.48	0.49	0.33	-0.43	0.74	-0.16	-0.19
Rv3597c	0.44	0.89	1.06	1.06	0.96	1.22	1.30	0.56	0.70	1.18	0.12	1.66	1.80
Rv3681c	-0.30	-0.28	-0.08	-0.28	-0.20	-0.13	-0.07	0.69	0.10	-0.05	0.66	-0.06	-0.04

Tabla S1 Resultados de expresión diferencial para las primeras 13 muestras. En el panel derecho se encuentran los genes de la red de dormancia de MTB. Los valores están expresados como “log2ratio”.

GENES	MUESTRAS											
	52h.- 0.0 % OD	54h.- 0.0 % OD	64h.- 0.0 % OD	70h.- 0.0 % OD	88h.- 0.0 % OD	94h.- 0.0 % OD	112h 0.0 % OD	114h 0.45 %OD	114.5 h.3.8 %OD	115h 6.0 % OD	116h 10.89 %OD	120h- 46.92 %OD
	14	15	16	17	18	19	20	21	22	23	24	25
Rv0081	3.82	3.35	2.98	2.90	3.58	1.94	2.09	2.68	3.56	3.35	3.52	2.82
Rv0324	-0.61	-0.24	0.65	0.21	-0.60	1.61	1.67	0.87	-1.12	-1.31	-1.39	-1.09
Rv0348	0.50	0.44	0.25	0.39	-0.03	-0.08	-0.04	0.30	0.48	0.62	0.40	0.18
Rv0474	1.69	1.46	1.07	1.08	1.53	-1.51	-1.97	0.60	1.36	0.58	-0.16	-0.75
Rv0494	-2.39	-0.87	0.59	-0.30	-1.50	1.70	1.74	0.74	-2.18	-2.70	-2.39	-2.28
Rv0653c	-0.51	-0.54	-0.32	-0.42	-0.77	0.45	0.99	-0.16	-1.30	-1.34	-0.90	-1.46

Rv0757	0.67	0.49	0.50	0.57	0.40	-0.32	-0.35	-0.09	0.11	0.10	-0.09	-0.38
Rv0767c	-0.48	0.09	1.26	0.58	0.78	2.14	1.72	1.63	0.34	-0.29	-0.51	-0.75
Rv0827c	1.36	1.08	1.57	1.93	2.15	1.05	1.16	1.09	1.11	0.70	0.61	0.36
Rv0967	0.66	0.09	0.44	0.86	1.81	-0.20	-0.23	1.12	2.21	2.01	1.70	0.84
Rv1255c	-0.91	-0.64	-0.87	-0.87	-0.55	-0.48	0.05	-0.12	0.35	0.34	0.09	-0.26
Rv1985c	0.27	0.34	0.65	0.44	0.65	1.47	1.36	1.06	-0.44	-0.89	-1.03	-1.15
Rv1990c	1.38	1.05	1.23	1.42	1.29	0.85	0.83	1.27	1.18	0.84	0.54	-0.17
Rv1994c	1.70	1.73	1.90	2.56	2.77	0.82	0.35	1.36	2.10	2.23	1.82	1.22
Rv2011c	-0.67	-0.27	0.56	0.00	-0.32	1.41	1.55	1.23	-0.36	-0.80	-0.71	-0.93
Rv2021c	1.62	1.39	1.92	2.23	2.24	0.37	0.21	1.43	2.09	1.91	1.55	0.76
Rv2034	1.38	1.25	1.74	1.89	1.90	1.48	1.54	1.75	2.06	1.67	0.93	0.07
Rv2324	0.47	-0.10	0.88	0.69	0.67	0.59	0.60	0.76	0.80	0.15	-0.06	-0.27
Rv2374c	0.08	-0.15	-0.21	-0.03	0.37	-0.58	-0.76	-0.43	-0.05	0.05	0.20	0.36
Rv3132c	4.05	3.81	3.54	3.64	3.86	1.02	0.58	1.46	2.16	2.23	2.12	1.17
Rv3133c	4.87	4.55	4.23	4.44	4.72	2.23	1.47	2.30	3.17	3.26	3.16	2.02
Rv3574	-0.18	-0.21	0.29	0.13	0.00	0.63	0.82	0.26	-0.63	-1.18	-1.16	-1.65
Rv3597c	2.20	1.72	0.48	0.63	0.72	-1.07	-1.58	-0.12	0.86	0.61	0.49	0.51
Rv3681c	0.44	0.29	0.57	0.68	0.86	0.47	0.50	0.53	0.35	0.31	0.32	0.20

Tabla S2 Resultados de expresión diferencial para las muestras 14 a 25.
 En el panel izquierdo se encuentran los genes de la red de dormancia de MTB.
 Los valores están expresados como “log₂ratio”.

Resultados del análisis del coeficiente de variación entre muestras

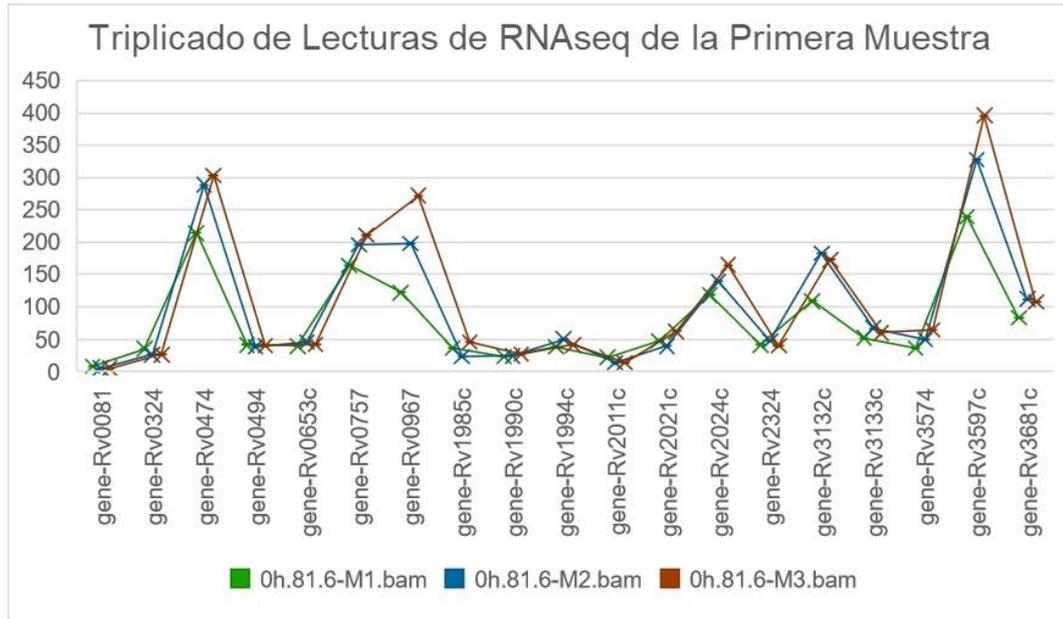


Fig.S3 Lecturas de RNAseq para la muestra 0 horas y 81.6 % de Oxígeno Disuelto. En el eje Y se muestran las lecturas de RNAseq, en el eje X se muestran los genes de la red de dormancia de MTB. En color rojo se muestran las lecturas de la muestra M3, en color azul las lecturas de la muestra M2 y en color verde las lecturas de la muestra M1.

	M1	M2	M3
	0horas.81.6 % OD	0horas.81.6 % OD	0horas.81.6 % OD
gene-Rv0081	8	5	5
gene-Rv0324	35	26	26
gene-Rv0474	214	289	303
gene-Rv0494	41	38	40
gene-Rv0653c	39	46	42
gene-Rv0757	164	196	211
gene-Rv0967	122	198	272

gene-Rv1985c	36	23	46
gene-Rv1990c	23	24	27
gene-Rv1994c	38	50	41
gene-Rv2011c	22	14	14
gene-Rv2021c	47	39	62
gene-Rv2024c	118	139	165
gene-Rv2324	41	47	40
gene-Rv3132c	109	183	173
gene-Rv3133c	52	67	60
gene-Rv3574	36	49	64
gene-Rv3597c	239	328	396
gene-Rv3681c	83	112	108
	COEF.COR. M1:M2	COEF.COR. M2:M3	COEF.COR. M1:M3
	0.993575283	0.992840604	0.986552489

Tabla S3 Coeficientes de correlación entre muestras. En el panel derecho se encuentran los genes de la red de dormancia de MTB. COEF.COR. M1:M2 representa el valor del coeficiente de correlación entre la muestra M1 y M2; COEF.COR. M2:M3 representa el valor del coeficiente de correlación entre la muestra M2 y M3; COEF.COR. M1:M3 representa el valor del coeficiente de correlación entre la muestra M1 y M3.

ANEXO 3. CÓDIGO PARA GENERAR LA MATRIZ “serie_tiempo”

Como se describe en la sección 7.2, la binarización de los valores de expresión genética, requiere crear como primer paso una matriz, las siguientes líneas de código en R crean la matriz “serie_tiempo”. Esta matriz está compuesta por los genes de la red de dormancia, organizados en vectores que contienen los valores de expresión diferencial expresados como “log2ratio” para cada una de las muestras analizadas.

CÓDIGO EN R PARA CARGAR LA MATRIZ “serie_tiempo” PARA 12 GENES DE LA RED DE DORMANCIA DE MTB

```
Rv1994c <- c(valores de expresión); Rv3133c <- c(valores de expression); Rv3132c <-  
c(valores de expression); Rv3574 <- c(valores de expression); Rv3597c <- c(valores de  
expression); Rv0757 <- c(valores de expression); Rv0081 <- c(valores de expression);  
Rv0324 <- c(valores de expression); Rv1985c <- c(valores de expression); Rv2034 <-  
c(valores de expression); Rv3681c <- c(valores de expression); Rv2011c <- c(valores de  
expression)
```

```
Serie_tiempo <- rbind(Rv1994c ,Rv3133c, Rv3132c, Rv3574,Rv3597c , Rv0757, Rv0081,  
Rv0324, Rv1985c, Rv2034, Rv3681c, Rv2011c)
```

ANEXO 4. FUNCIONES REGULADORAS

Las siguientes funciones reguladoras fueron obtenidas de siguiendo el algoritmo best-fit (24) descrito en la sección 7.4.4, usando un valor de K máximo de cuatro, a partir de los datos de expresión genética del trabajo de Peterson y colaboradores (65).

FUNCIONES PROBABILÍSTICAS PARA LOS 19 GENES: Rv0081 Rv0324 Rv0474 Rv0494 Rv0653c Rv0757 Rv0967 Rv1985c Rv1990c Rv1994c Rv2011c Rv2021c Rv2034 Rv2324 Rv3132c Rv3133c Rv3574 Rv3597c Rv3681c.

Funciones *in silico* para el gen Rv0081:

Rv0081 = (!Rv0653c & Rv3574 & !Rv3681c) | (!Rv0653c & Rv3133c & !Rv3681c) | (!Rv0653c & Rv3133c & !Rv3574) | (Rv0653c & Rv3574 & Rv3681c) | (Rv3133c & Rv3574 & !Rv3681c) (error: 1)Rv0081 = (!Rv0653c & !Rv3132c & !Rv3574 & Rv3681c) | (!Rv0653c & Rv3574 & !Rv3681c) | (!Rv0653c & Rv3132c & !Rv3681c) | (Rv0653c & Rv3574 & Rv3681c) | (Rv3132c & Rv3574 & !Rv3681c) (error: 1)Rv0081 = (!Rv0653c & !Rv2324 & Rv3574) | (!Rv0653c & Rv3133c & !Rv3574) | (!Rv2324 & Rv3133c & Rv3574) | (Rv0653c & Rv2324 & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv2034 & Rv3574) | (!Rv0653c & Rv3133c & !Rv3574) | (!Rv2034 & Rv3133c & Rv3574) | (Rv0653c & Rv2034 & Rv3574) (error: 1)Rv0081 = (!Rv2034 & Rv3574) | (!Rv0653c & !Rv2034 & Rv3132c) | (!Rv0653c & Rv2034 & !Rv3132c & !Rv3574) | (Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv2021c & Rv3574) | (!Rv0653c & Rv3133c & Rv3574) | (!Rv2021c & Rv3574) | (Rv0653c & Rv2021c & Rv3574) (error: 1)Rv0081 = (!Rv2021c & Rv3574) | (!Rv0653c & !Rv2021c & Rv3132c) | (!Rv0653c & Rv2021c & Rv3132c & !Rv3574) | (Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv1994c & Rv3574) | (!Rv0653c & Rv1994c & Rv3133c & !Rv3574) | (Rv0653c & Rv3133c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv1985c & Rv3574) | (!Rv0653c & !Rv1985c & Rv3133c) | (Rv0653c & Rv1985c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv0967 & Rv3574) | (!Rv0653c & Rv3133c & !Rv3574) | (!Rv0967 & Rv3133c & Rv3574) | (Rv0653c & Rv0967 & Rv3574) (error: 1)Rv0081 = (!Rv0967 & Rv3574) | (!Rv0653c & !Rv0967 & Rv3132c) | (!Rv0653c & Rv0967 & !Rv3132c & !Rv3574) | (Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv3574 & Rv3681c) | (Rv0653c & Rv3574 & Rv3681c) | (Rv0474 & !Rv0653c & !Rv3681c) | (Rv0474 & !Rv0653c & !Rv3681c) | (Rv0474 & Rv0653c & Rv3681c) (error: 1)Rv0081 = (!Rv0653c & Rv2034 & !Rv3574) | (Rv0653c & !Rv2034 & Rv3574) | (Rv0474 & !Rv0653c & !Rv3574) | (Rv0474 & !Rv2034 & Rv3574) | (Rv0474 & Rv0653c & !Rv3574) | (Rv0474 & !Rv2034 & Rv3574) (error: 1)Rv0081 = (!Rv0653c & Rv2021c & !Rv3574) | (Rv0474 & !Rv0653c & !Rv3574) | (Rv0474 & !Rv2021c & Rv3574) | (Rv0474 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv1985c & Rv3681c) | (Rv0653c & Rv1985c & Rv3681c) | (Rv0474 & !Rv0653c & !Rv1985c) | (Rv0474 & Rv0653c & !Rv1985c) (error: 1)Rv0081 = (!Rv0653c & !Rv1985c & Rv3133c) | (Rv0653c & Rv1985c & Rv3133c) | (Rv0474 & !Rv0653c & !Rv1985c) | (Rv0474 & Rv0653c & Rv1985c) (error: 1)Rv0081 = (!Rv0653c & Rv0967 & !Rv3574) | (Rv0653c & !Rv0967 & Rv3574) | (Rv0474 & !Rv0653c & !Rv3574) | (Rv0474 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & Rv0757 & Rv3681c) | (Rv0474 & !Rv0653c & Rv0757 & !Rv3681c) | (Rv0474 & Rv0653c & !Rv0757) (error: 1)Rv0081 = (!Rv0474 & !Rv0653c & !Rv2324) | (Rv0653c & Rv0757 & Rv2324) | (Rv0474 & !Rv0757 & Rv2324) | (Rv0474 & Rv0757 & !Rv2324) | (Rv0474 & Rv0653c) (error: 1)Rv0081 = (!Rv0653c & Rv2034) | (!Rv0474 & Rv0653c & Rv0757 & !Rv2034) | (Rv0474 & !Rv0653c & Rv0757 & !Rv2034) | (Rv0474 & Rv0653c & Rv0757 & !Rv2021c) | (Rv0474 & Rv0653c & !Rv0757) | (Rv0474 & Rv0653c & Rv2034) (error: 1)Rv0081 = (!Rv0653c & !Rv0757 & Rv2021c) | (Rv0653c & Rv0757 & !Rv2021c) | (Rv0474 & Rv0757 & !Rv2021c) | (Rv0474 & Rv0653c) (error: 1)Rv0081 = (!Rv0653c & !Rv1985c) | (Rv0653c & Rv1985c) | (Rv0474 & !Rv0653c & !Rv1985c) | (Rv0474 & Rv0653c & Rv1985c) (error: 1)Rv0081 = (!Rv0653c & !Rv0757 & Rv0967) | (Rv0653c & Rv0757 & !Rv0967) | (Rv0474 & Rv0757 & !Rv0967) | (Rv0474 & Rv0653c) (error: 1)Rv0081 = (!Rv0653c & Rv3574 & !Rv3681c) | (Rv0081 & !Rv0653c & !Rv3574) | (Rv0081 & Rv3574 & !Rv3681c) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv2324 & Rv3574) | (Rv0081 & !Rv0653c & !Rv3574) | (Rv0081 & !Rv2324 & Rv3574) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv2034 & Rv3574) | (Rv0081 & !Rv0653c & !Rv3574) | (Rv0081 & !Rv2034 & Rv3574) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv2021c & Rv3574) | (Rv0081 & !Rv0653c & !Rv3574) | (Rv0081 & !Rv2021c & Rv3574) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv1994c & Rv3574) | (Rv0081 & !Rv0653c & Rv1994c & !Rv3574) | (Rv0081 & !Rv1994c & Rv3574) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv1985c & Rv3574) | (Rv0081 & !Rv0653c & !Rv1985c & Rv3574) | (Rv0081 & !Rv1985c & Rv3574) | (Rv0081 & Rv0653c & !Rv1985c) | (Rv0081 & Rv0653c & Rv1985c & Rv3574) (error: 1)Rv0081 = (!Rv0653c & !Rv0967 & Rv3574) | (Rv0081 & !Rv0653c & !Rv0967 & Rv3574) | (Rv0081 & !Rv0967 & Rv3574) | (Rv0081 & Rv0653c & !Rv3574) | (Rv0081 & Rv0653c & Rv3574) (error: 1)Rv0081 = (Rv0474 & !Rv0653c & !Rv1985c) | (Rv0081 & !Rv0653c & !Rv1985c) | (Rv0081 & Rv0653c & Rv1985c) (error: 1)

Funciones *in silico* para el gen Rv0324:

Rv0324 = (!Rv0967 & !Rv1990c & !Rv3132c & Rv3133c) | (!Rv0653c & !Rv0967 & !Rv1990c & Rv3133c) | (!Rv0653c & Rv0967 & Rv1990c & Rv3132c & Rv3133c) | (Rv0653c & !Rv0967 & Rv1990c & !Rv3132c & !Rv3133c) | (Rv0653c & !Rv0967 & Rv1990c & Rv3132c & Rv3133c) (error: 0)

& !Rv0757 & Rv2324 & Rv3574) | (Rv0081 & !Rv0653c & Rv0757 & !Rv2324 & !Rv3574) | (Rv0081 & !Rv0653c & Rv0757 & Rv2324 & Rv3574) | (Rv0081 & Rv0653c & Rv0757 & !Rv2324 & Rv3574) (error: 1)Rv2011c = (!Rv0081 & Rv0653c & !Rv0757 & !Rv1994c & Rv2324) | (Rv0081 & Rv0757 & Rv1994c & !Rv2324) | (Rv0081 & !Rv0653c & Rv0757 & Rv1994c) (error: 1)Rv2011c = (!Rv0081 & Rv0653c & !Rv0757 & Rv1990c & Rv2324) | (Rv0081 & !Rv0653c & Rv0757 & !Rv1990c & !Rv2324) | (Rv0081 & !Rv0653c & Rv0757 & Rv1990c & Rv2324) | (Rv0081 & !Rv0653c & Rv0757 & Rv1990c & Rv2324) | (Rv0081 & Rv0653c & Rv0757 & Rv1990c & !Rv2324) (error: 1)Rv2011c = (!Rv0081 & Rv0324 & Rv0653c & !Rv0967 & Rv3574) | (Rv0081 & !Rv0324 & !Rv0653c & !Rv0967 & !Rv3574) | (Rv0081 & !Rv0324 & !Rv0653c & Rv0967 & Rv3574) | (Rv0081 & !Rv0324 & !Rv0653c & Rv0967 & Rv3574) | (Rv0081 & !Rv0324 & Rv0653c & !Rv0967 & Rv3574) (error: 1)

Funciones *in silico* para el gen Rv2021c:

Rv2021c = (!Rv0653c & Rv1990c & !Rv3574) | (Rv0653c & Rv1990c & Rv3574) (error: 2)Rv2021c = (Rv1990c & !Rv3132c) | (Rv0653c & Rv1990c) (error: 2)Rv2021c = (!Rv0653c & !Rv1985c & Rv1990c) | (Rv0653c & Rv1985c & Rv1990c) (error: 2)Rv2021c = (!Rv0757 & Rv1990c) | (Rv0653c & Rv1990c) (error: 2)

Funciones *in silico* para el gen Rv2034:

Rv2034 = (!Rv1990c & Rv3132c & !Rv3574) | (Rv1990c & !Rv3132c) | (Rv1990c & Rv3574) (error: 1)Rv2034 = (Rv1994c & Rv3132c) | (Rv1990c & !Rv3132c) (error: 1)Rv2034 = (!Rv1985c & !Rv1990c & Rv3132c) | (Rv1990c & !Rv3132c) | (Rv1985c & Rv1990c) (error: 1)Rv2034 = (!Rv0653c & Rv3132c) | (Rv1990c) (error: 1)Rv2034 = (!Rv0494 & Rv3132c) | (Rv1990c) (error: 1)Rv2034 = (!Rv1994c & !Rv3597c) | (Rv0474 & Rv1994c & Rv3597c) (error: 1)Rv2034 = (Rv1990c & Rv3574) | (Rv0474 & !Rv3574) (error: 1)Rv2034 = (Rv1990c & !Rv1994c) | (Rv0474 & Rv1994c) (error: 1)Rv2034 = (!Rv0324 & Rv3132c) | (Rv1990c) (error: 1)

Funciones *in silico* para el gen Rv2324:

Rv2324 = (!Rv0494 & !Rv0653c & !Rv2021c & Rv3133c) | (!Rv0494 & !Rv0653c & Rv3133c & Rv3574) | (Rv0494 & Rv0653c & !Rv3133c & Rv3574) | (Rv0494 & Rv0653c & Rv2021c & Rv3574) (error: 2)Rv2324 = (!Rv0494 & !Rv0653c & !Rv2021c & !Rv2324 & Rv3133c) | (!Rv0494 & !Rv0653c & Rv2021c & Rv2324 & Rv3133c) | (Rv0494 & Rv0653c & Rv2324 & !Rv3133c) | (Rv0494 & Rv0653c & Rv2021c & Rv3133c) (error: 2)Rv2324 = (!Rv0494 & !Rv0653c & !Rv1990c & Rv3133c & !Rv3681c) | (!Rv0494 & !Rv0653c & Rv1990c & Rv3133c & Rv3681c) | (Rv0494 & Rv0653c & Rv1990c & Rv3133c & Rv3681c) (error: 2)Rv2324 = (!Rv0494 & !Rv0653c & !Rv1990c & !Rv2034 & Rv3133c) | (!Rv0494 & !Rv0653c & Rv1990c & Rv2034 & Rv3133c) | (Rv0494 & Rv0653c & Rv1990c & Rv2034) (error: 2)Rv2324 = (!Rv0494 & !Rv0653c & !Rv1990c & !Rv2021c & Rv3133c) | (!Rv0494 & !Rv0653c & Rv1990c & Rv2021c & Rv3133c) | (Rv0494 & Rv0653c & Rv1990c & !Rv3133c) | (Rv0494 & Rv0653c & Rv1990c & Rv2021c) (error: 2)Rv2324 = (!Rv0494 & !Rv0653c & !Rv0967 & !Rv1990c & Rv3133c) | (!Rv0494 & !Rv0653c & Rv0967 & Rv1990c & Rv3133c) | (Rv0494 & Rv0653c & Rv1990c & Rv3133c) | (Rv0494 & Rv0653c & Rv0757 & Rv3133c) | (Rv0494 & Rv0653c & Rv1990c & !Rv3133c) | (Rv0474 & !Rv0494 & !Rv0653c & Rv3133c) | (Rv0474 & Rv0494 & Rv0653c & Rv1990c) (error: 2)Rv2324 = (!Rv0324 & !Rv0653c & !Rv2021c & !Rv2324 & Rv3133c) | (Rv0324 & Rv0653c & Rv2021c & !Rv2324 & Rv3133c) | (Rv0324 & Rv0653c & Rv2324 & !Rv3133c) | (Rv0324 & Rv0653c & Rv2021c & Rv2324) (error: 2)Rv2324 = (!Rv0324 & !Rv0653c & !Rv0967 & !Rv1990c & Rv3133c) | (!Rv0324 & !Rv0653c & Rv0967 & Rv1990c & Rv3133c) | (Rv0324 & Rv0653c & Rv1990c & Rv3133c) | (Rv0324 & Rv0653c & Rv1990c & Rv3133c) (error: 2)

Funciones *in silico* para el gen Rv3132c:

Rv3132c = (Rv3132c & Rv3597c & !Rv3681c) | (Rv0653c & Rv3597c & Rv3681c) (error: 0)Rv3132c = (Rv3132c & Rv3133c & !Rv3681c) | (Rv0653c & Rv3133c & Rv3681c) (error: 0)Rv3132c = (!Rv2324 & Rv3132c & Rv3597c) | (Rv0653c & Rv2324 & Rv3597c) (error: 0)Rv3132c = (!Rv2324 & Rv3132c & Rv3133c) | (Rv0653c & Rv2324 & Rv3133c) (error: 0)Rv3132c = (!Rv2034 & Rv3132c & !Rv3681c) | (Rv0653c & !Rv2034 & !Rv3132c & Rv3681c) | (Rv0653c & Rv2034 & Rv3132c & Rv3681c) (error: 0)Rv3132c = (!Rv0653c & !Rv2034 & Rv3132c) | (Rv0653c & !Rv2034 & Rv3132c) | (Rv0653c & !Rv2034 & Rv3574) | (Rv0653c & Rv3132c & Rv3574) (error: 0)Rv3132c = (!Rv2034 & Rv3132c & Rv3133c) | (Rv0653c & !Rv2034 & Rv3133c) | (Rv0653c & Rv3132c & Rv3133c) (error: 0)Rv3132c = (!Rv2034 & !Rv2324 & Rv3132c) | (Rv0653c & !Rv2034 & Rv2324 & !Rv3132c) | (Rv0653c & !Rv2324 & Rv3132c) | (Rv0653c & !Rv2324 & Rv3132c) | (Rv0653c & Rv2324 & Rv3132c) | (Rv0653c & Rv2034 & Rv3132c) (error: 0)Rv3132c = (!Rv2021c & Rv3132c & Rv3133c) | (Rv0653c & !Rv2021c & Rv3132c) | (Rv0653c & !Rv2021c & Rv3133c) | (Rv0653c & Rv2021c & Rv3132c) | (Rv0653c & Rv3132c & Rv3133c) (error: 0)Rv3132c = (!Rv1990c & Rv3132c & !Rv3681c) | (!Rv0653c & Rv3132c & !Rv3681c) | (Rv0653c & Rv3132c & !Rv3681c) | (Rv0653c & !Rv1990c & Rv3681c) | (Rv0653c & !Rv1990c & Rv3681c) (error: 0)Rv3132c = (!Rv2324 & Rv3132c) | (Rv0653c & !Rv1990c & Rv2324) | (Rv0653c & Rv3132c) (error: 0)Rv3132c = (!Rv0653c & !Rv1985c & Rv3132c & Rv3133c) | (Rv0653c & Rv1985c & Rv3133c) (error: 0)Rv3132c = (!Rv0967 & Rv3132c & Rv3133c) | (Rv0653c & !Rv0967 & Rv3133c) | (Rv0653c & Rv3132c & Rv3133c) (error: 0)Rv3132c = (Rv0757 & !Rv2324 & Rv3132c) | (Rv0653c & !Rv2324 & Rv3132c) | (Rv0653c & Rv0757 & Rv2324) (error: 0)Rv3132c = (!Rv0653c & Rv0757 & !Rv2034 & Rv3132c) | (Rv0653c & !Rv0757 & !Rv2034 & Rv3132c) | (Rv0653c & Rv0757 & !Rv2034 & !Rv3132c) | (Rv0653c & Rv0757 & Rv2034 & Rv3132c) (error: 0)Rv3132c = (Rv0757 & !Rv2021c & Rv3132c) | (Rv0653c & !Rv2021c & Rv3132c) | (Rv0653c & Rv0757 & !Rv2021c) | (Rv0653c & Rv0757 & Rv3132c) (error: 0)Rv3132c = (!Rv0653c & Rv0757 & !Rv1985c & Rv3132c) | (Rv0653c & Rv1985c & Rv3132c) | (Rv0653c & Rv0757 & Rv1985c) (error: 0)Rv3132c = (Rv0757 & !Rv0967 & Rv3132c) | (Rv0653c & !Rv0967 & Rv3132c) | (Rv0653c & Rv0757 & !Rv0967) | (Rv0653c & Rv0757 & Rv3132c) (error: 0)Rv3132c = (Rv0081 & !Rv2034 & Rv3132c) | (Rv0081 & Rv0653c & !Rv2034) | (Rv0081 & Rv0653c & Rv3132c) (error: 0)Rv3132c = (Rv0081 & !Rv2021c & Rv3132c) | (Rv0081 & Rv0653c & !Rv2021c) | (Rv0081 & Rv0653c & Rv3132c) (error: 0)Rv3132c = (Rv0081 & !Rv0967 & Rv3132c) | (Rv0081 & Rv0653c & !Rv0967) | (Rv0081 & Rv0653c & Rv3132c) (error: 0)

Funciones *in silico* para el gen Rv3133c:

Rv3133c = (Rv0081 & !Rv0757 & Rv2324) | (Rv0081 & Rv0757 & !Rv2324) | (Rv0081 & Rv0653c) (error: 0)Rv3133c = (Rv0081 & Rv0653c & Rv1985c) | (Rv0081 & Rv0474 & !Rv0653c & !Rv1985c) (error: 0)

ANEXO 5

Las tablas que se presentan a continuación contienen los datos de expresión diferencial expresados en “log2ratio”, obtenidos como se indica en la sección 7.1.4 con el algoritmo DESeq aplicado a las lecturas de RNAseq de 25 muestras para los genes asociados a la Reactivación.

GENES	MUESTRAS												
	17h.- 57.6 % OD	23h.- 42.8 % OD	40h.- 11.5 % OD	41h.- 8.8 % OD	42h.- 7.49 % OD	43h.- 6.16 % OD	44h.- 4.31 % OD	45h.- 2.9% OD	46h.- 0.41 % OD	47h.- 0.0 % OD	48h.- 0.0 % OD	49h.- 0.0 % OD	50h.- 0.0 % OD
	1	2	3	4	5	6	7	8	9	10	11	12	13
Rv1353 c	0.00	-0.85	-0.75	-0.90	-0.38	-0.36	-0.70	1.25	0.60	-0.67	1.28	-0.49	-0.76
Rv1884 c	0.30	1.03	1.07	1.33	1.21	1.41	1.57	0.40	0.79	1.30	0.21	1.05	1.05
Rv2150 c	0.96	1.13	1.37	1.22	1.02	1.27	1.36	0.38	0.65	1.27	0.21	1.37	1.48
Rv2389 c	-0.40	-0.05	-0.54	-0.61	-0.36	-0.31	-0.14	0.22	-0.17	-0.57	-1.05	-0.47	-0.34
Rv2450 c	0.81	0.44	0.67	1.01	0.84	1.33	0.76	0.33	0.33	0.46	-0.39	-0.59	-0.69
Rv1033 c	-0.19	-0.09	-0.31	-0.23	-0.33	-0.85	-0.33	0.30	-0.05	-0.45	0.73	-0.27	-0.58

Tabla S4 Resultados de expresión diferencial para las primeras 13 muestras. En el panel derecho se encuentran los genes de la red de dormancia de MTB. Los valores están expresados como “log2ratio”.

	MUESTRAS											
	52h- 0.0 % OD	54h- 0.0 % OD	64h- 0.0 % OD	70h- 0.0 % OD	88h- 0.0 % OD	94h- 0.0 % OD	112h0 .0 % OD	114h0 .45 % OD	114.5 h.3.8 %OD	115h6 .0 % OD	116h1 0.89 % OD	120h- 46.92 % OD
GENES	14	15	16	17	18	19	20	21	22	23	24	25
Rv1353c	-0.36	0.14	0.86	0.38	-0.18	1.79	1.94	1.29	-0.36	-0.40	-0.54	-0.87
Rv1884c	0.93	0.64	-0.60	-0.94	1.78	-2.91	-2.75	0.25	-2.43	-1.93	-1.70	-1.28
Rv2150c	1.45	1.23	1.09	1.07	-0.13	-0.76	-1.13	0.45	1.11	0.99	0.76	-0.09
Rv2389c	0.22	0.09	0.08	-0.09	0.21	0.58	0.50	-0.77	-0.06	-0.31	-0.02	-0.30
Rv2450c	-1.21	-1.39	-1.13	-0.69	0.72	-0.87	-1.03	1.02	1.34	2.35	2.81	2.77
Rv1033c	-0.20	-0.04	0.08	0.04	-0.38	0.81	0.90	0.73	-0.75	-0.67	-0.72	-1.05

Tabla S5 Resultados de expresión diferencial para las muestras 14 a 25.
 En el panel derecho se encuentran los genes de la red de dormancia de MTB.
 Los valores están expresados como “log2ratio”.

ANEXO 6

Código para la obtención de lecturas de RNA seq:

```
compressedFastq <- list.files(path="fastq/", pattern="*.fastq.gz",
full.names=TRUE)
for (t in 1:NROW (compressedFastq))
{
  gunzip ((compressedFastq[t]), remove=FALSE)
}
allFastq_1 <- list.files(path="fastq/" , pattern="*_1.fastq$", full.names=TRUE)
allFastq_2 <- list.files(path="fastq/" , pattern="*_2.fastq$", full.names=TRUE)
gunzip("cromosoma/GCF_000195955.2_ASM19595v2_genomic.fna.gz")
```

Código para el alineamiento de las lecturas al genoma de MTB:

```
bowtie2_build("cromosoma1/GCF_000195955.2_ASM19595v2_genomic.fna" ,
bt2Index = "cromosoma1/mtbee" )
PROJECT <- data.frame(fastq_1 = allFastq_1 , fastq_2 = allFastq_2)
write.table(PROJECT, file="myproject.tsv", sep="\t", row.names=FALSE)
PROJECT <- read.table("myproject.tsv" , header=TRUE, sep="\t")
dir.create("aligned")
for(t in 1:NROW(PROJECT))
{
  myfq1 <- PROJECT$fastq_1[t]
  myfq2 <- PROJECT$fastq_2[t]
  mysam <- paste0("aligned/" , PROJECT1$conditions[t], "-",
PROJECT1$replicate[t], ".sam")
  bowtie2(bt2Index = "cromosoma/mtbee",samOutput = mysam, seq1 = myfq1,
seq2=myfq2, "--threads 3")
  asBam(mysam)
}
data4 <- featureCounts(
files = bamFiles,
annot.ext = "genes/sequence.gff3",
isGTFAnnotation = TRUE,
GTF.featureType = "gene",
GTF.attrType = "ID",
isPairedEnd = TRUE,
```

```

requireBothEndsMapped = TRUE,
      minMQS = 20,
      strandSpecific = 2,
      nthreads = 3
    )
write.table(data4$counts, file = "raw_counts.tsv", sep = "\t")

```

Código para el análisis de expresión diferencial entre la muestra de condición 17 horas (con 57.62 % OD) y la muestra de 0 horas (con 81.6 % OD):

```

dds <- DESeqDataSetFromMatrix(data4$counts, colData=PROJECT1, design =
~ conditions)
dds <- DESeq(dds)

res17h.57.62vs0h.81.6 <- results(dds, contrast=c("conditions",
"17h.57.6", "0h.81.6"))
final_17h.57.62vs0h.81.6 <- data.frame(
      GENEID =
row.names(res17h.57.62vs0h.81.6),
      log2BaseMean17h.57.62vs0h.81.6 =
log2(res17h.57.62vs0h.81.6$baseMean),
      los2Ratio17h.57.62vs0h.81.6 =
res17h.57.62vs0h.81.6$log2FoldChange,
      STDERR_log2Ratio17h.57.62vs0h.81.6 =
res17h.57.62vs0h.81.6$lfcSE,
      pvalue = res17h.57.62vs0h.81.6$pvalue,
      padjust = res17h.57.62vs0h.81.6$padj,
      assays(dds)$counts
    )
write.table(final_17h.57.62vs0h.81.6, file="final2_17h.57.62vs0h.81.6.tsv",
sep="\t", row.names=FALSE)

```

Código para binarización de los niveles de expresión:

```

Rv0081 <- c (0.719928385558296,-0.26114378301079,,,)
Rv0324 <- c (0.131328725085413,-0.78866772832949,,,)

```

```
serie_tiempo ←rbind(Rv0081,Rv0324,Rv0348,Rv0474,Rv0494,Rv0653c,  
Rv0757,Rv0767c,Rv0827c,Rv0967,Rv1255c,Rv1985c,Rv1990c,Rv1994c  
,Rv2011c,Rv2021c,Rv2031c,Rv2034,Rv2324,Rv2374c,Rv3132c,Rv3133c,Rv357  
4,Rv3597c,Rv3681c).  
binSeries <- binarizeTimeSeries(serie_tiempo)  
write.table(binSeries, file="serie_tiempo_25.tsv", sep="\t", row.names=TRUE)
```

Código para realizar el agrupamiento de los estados:

```
library(FactoMiner)  
estados_hipoxia read.csv (file ='estados_hipoxia.csv' , header= TRUE, sep="," ,  
row.names=1)  
acp_estados_hipoxia <- PCA (data, graph = FALSE)  
plot ( acp_estados_hipoxia, choix = "ind")  
plot ( acp_estados_hipoxia, choix = "var")
```

Código para obtener la gráfica de la red:

```
library(BoolNet)  
library(downloader)  
getwd()  
net <- loadNetwork("redmtbe.24.txt")  
plotNetworkWiring(net)
```

Código para obtener los atractores de la red:

```
net <- loadNetwork("redmtbereducida.txt")  
attr <- getAttractors(net)  
plotAttractors(attr)
```

Código para obtener las funciones reguladoras *in silico*:

```
net.data <- binSeries$binarizedMeasurements  
network.size<- 12  
booleannet <- reconstructNetwork(net.data, method = "bestfit", maxK = 4,  
readableFunctions=T,  
returnPBN = F)  
print (booleannet)
```

Código para el análisis de trayectorias hacia el atractor:

```
startState <- generateState(net,
specs=c("Rv3133c"=1,"Rv3132c"=1,"Rv3574"=0,"Rv3597c"=0,"Rv0757"=0,"Rv
0081"=0, "Rv0324"=0, "Rv2034"=0, "Rv0474"=0, "Rv1990c"=0, "Rv0967"=0,
"Rv1994c"=0, "Rv2021c"=0, "Rv3681c"=0, "Rv2324" =0, "Rv0494"=0,
"Rv0653c"=0, "Rv1985c"=0, "Rv2011c"=0,"Rv2031c"=0 ))
stateTransition(net,startState)
path <- getPathToAttractor(network=net3, state=startState)
plotSequence(sequence=path)
```

Código para la mutación de activación:

```
overExpressed <- fixGenes(net, "Rv3133c", 1)attr <- attr <-
getAttractors(overExpressed)
plotAttractors(attr)
```

Código para la mutación de inactivación:

```
knockedOut <- fixGenes(net3, "Rv0081", 0)
knockedOut
```

Código para generar funciones aleatorias:

```
net_random <- generateRandomNKNNetwork(n=20,
k=c(4,2,4,6,5,3,4,3,3,4,2,3,3,3,2,2,3,6,2,1))
net_random
```