



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
INSTITUTO DE INVESTIGACIONES BIOMÉDICAS

# DINÁMICA DE LA DIFERENCIACIÓN DE CÉLULAS TH: MODELACIÓN CON REDES BOOLEANAS

TESIS PARA OBTENER EL TÍTULO DE:  
LICENCIADA EN INVESTIGACIÓN BIOMÉDICA BÁSICA

PRESENTA:  
MARIANA ESTHER MARTÍNEZ SÁNCHEZ

DIRECTORES DE TESIS:  
DR. GERMINAL COCHO GIL  
DR. CARLOS VILLARREAL LUJÁN

México, D.F. 2011



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

---

# Agradecimientos

Gracias, muchas gracias.

---

# Índice general

<b>Agradecimientos</b>	<b>2</b>
<b>Resumen</b>	<b>9</b>
<b>1. Introducción</b>	<b>11</b>
<b>2. Del gen a la proteína</b>	<b>14</b>
2.1. Regulación génica . . . . .	15
2.1.1. Transcripción . . . . .	16
2.1.2. Traducción . . . . .	18
2.1.3. Post-traducción . . . . .	19
2.2. Interacciones proteína-proteína . . . . .	21
2.2.1. Fosforilación . . . . .	22
2.2.2. Dimerización . . . . .	22
2.2.3. Competencia sitio-específica . . . . .	22
<b>3. El sistema inmunológico</b>	<b>24</b>
3.1. Inmunidad innata . . . . .	24
3.2. Inmunidad adaptativa . . . . .	26

---

3.2.1.	Linfocitos . . . . .	27
3.2.2.	Células presentadoras de antígeno . . . . .	29
3.3.	Diferenciación de células Th . . . . .	29
3.3.1.	Ontogenia de células Th . . . . .	30
3.3.2.	Th0 . . . . .	32
3.3.3.	Th1 . . . . .	33
3.3.4.	Th2 . . . . .	34
3.3.5.	Th17 . . . . .	35
3.3.6.	Treg . . . . .	36
3.3.7.	Regulación cruzada . . . . .	37
3.3.8.	Plasticidad . . . . .	40
3.4.	Sistema inmune y enfermedad . . . . .	42
3.4.1.	VIH/SIDA . . . . .	43
3.4.2.	Diabetes tipo 2 . . . . .	43
<b>4.</b>	<b>Redes</b>	<b>45</b>
4.1.	Funciones booleanas . . . . .	48
4.1.1.	Funciones sigmoidales . . . . .	49
4.2.	Elementos de las redes . . . . .	50
4.2.1.	Dirección y conectividad . . . . .	50
4.2.2.	Trayectorias . . . . .	51
4.2.3.	Centralidad . . . . .	51
4.2.4.	Estructuras y motivos . . . . .	52
4.3.	Red de estados . . . . .	53
4.4.	Fases dinámicas . . . . .	54
4.4.1.	Efectos de las perturbaciones . . . . .	56

---

<b>5. Diseño de la red</b>	<b>57</b>
5.1. Reglas lógicas y tablas de verdad . . . . .	60
5.1.1. Regulación epigenética . . . . .	60
5.1.2. Regulación por proteínas SOCS . . . . .	61
5.1.3. El caso $TGF\beta$ . . . . .	61
5.2. Simplificación de la red . . . . .	62
5.3. Análisis de la red . . . . .	62
5.3.1. Criterios de clasificación . . . . .	64
<b>6. Soluciones</b>	<b>67</b>
6.1. Soluciones puntuales . . . . .	68
6.1.1. Ciclos . . . . .	68
6.2. Modularidad . . . . .	71
6.3. Centralidad . . . . .	73
<b>7. Atractores</b>	<b>74</b>
7.1. Efectos de las perturbaciones . . . . .	76
<b>8. Conclusiones</b>	<b>79</b>
<b>Bibliografía</b>	<b>81</b>

---

# Índice de figuras

2.1. Dogma central de la biología . . . . .	14
2.2. Mecanismos de regulación genética . . . . .	15
2.3. Estructura del DNA . . . . .	16
2.4. Interacciones proteicas . . . . .	23
3.1. Inmunidad innata y adaptativa . . . . .	25
3.2. Linajes de células Th . . . . .	30
3.3. Ontogenia de los linfocitos . . . . .	31
3.4. Regulación cruzada de células Th . . . . .	38
3.5. Regulación mediada por SOCS . . . . .	39
3.6. Plasticidad en células Th . . . . .	41
3.7. Transiciones entre linajes de células Th . . . . .	42
4.1. Ejemplo de red booleana . . . . .	47
4.2. Función escalón . . . . .	49
4.3. Función de Hill . . . . .	49
4.4. Función logística . . . . .	49
4.5. Ejemplo de una red de estados . . . . .	53

---

4.6. Diagrama de fases dinámicas de redes booleanas . . . . .	54
4.7. Mapa de Derrida con redes aleatorias . . . . .	56
5.1. Imagen de la red de regulación de células Th . . . . .	58
5.2. Regulación diferencial por $TGF\beta$ . . . . .	62
5.3. Red booleana de diferenciación de células Th . . . . .	63
6.1. Ejemplo de ciclos de tamaño dos y tres . . . . .	70
6.2. Módulos de la red . . . . .	72
7.1. Tamaño de la cuenca de atracción de los distintos linajes . . . . .	74
7.2. Iteraciones necesarias para llegar a una solución . . . . .	75
7.3. Cantidad de aristas de entrada por nodo. . . . .	76
7.4. Sensibilidad a las perturbaciones . . . . .	76
7.5. Efectos del ruido en la transición entre linajes . . . . .	77
7.6. Conservación de los atractores . . . . .	78



---

# Índice de cuadros

4.1. Funciones and y or . . . . .	48
4.2. Función not . . . . .	48
5.1. Funciones de la red de regulación de células Th . . . . .	59
5.2. Tabla de verdad de regulación epigenética . . . . .	60
5.3. Tabla de verdad de transducción de señales . . . . .	61
5.4. Red reducida de regulación de células Th . . . . .	64
6.1. Clasificación de los estados . . . . .	67
6.2. Atractores de Th1, Th2, Th17 y Treg . . . . .	68
6.3. Criterios de clasificación de linajes . . . . .	69
6.4. Matriz de adyacencia de la red . . . . .	71
6.5. Módulos de la red . . . . .	72
6.6. Centralidad . . . . .	73

---

# Resumen

Las células T CD4<sup>+</sup> coadyuvantes vírgenes (Th0) se diferencian en distintos linajes incluyendo Th1, Th2, Th17 y T reguladoras (Treg). El destino celular de las células Th depende del tipo y fuerza de los estímulos que éstas reciben y de los subsecuentes patrones de expresión génica. Cada uno de estos linajes Th tiene funciones específicas en la homeostasis y defensa inmune. La incorrecta regulación de la diferenciación de las células Th se ha visto implicada en diversas enfermedades autoinmunes, la inflamación crónica, el VIH y el síndrome metabólico.

Con base en la literatura se construyó una red de regulación booleana de 43 nodos y 84 interacciones, donde los nodos representan factores de regulación transcripcional y citocinas. Para facilitar el análisis, la red fue simplificada usando reglas de simplificación de funciones booleanas, quedando una red de 17 nodos y 51 interacciones. Los nodos son: IL-2, IL-4, GATA3, IL-12, IL-18, IFN $\gamma$ , T-bet, IL-1, IL-6, IL-17, IL-21, IL-23, STAT3, ROR $\gamma$ t, TGF $\beta$ lo, TGF $\beta$ hi y Foxp3. Se generaron dos nodos TGF $\beta$  correspondientes a los diferentes niveles de expresión de la proteína para lograr reflejar la dinámica del sistema.

Se generó una serie de programas usando el lenguaje de programación Python para analizar el comportamiento de la red. De los 131072 ( $2^{17}$ ) estados posibles de la red 256 (0.19%) son estados estacionarios, además hay 60 (0.05%) estados que corresponden a doce ciclos de tamaño dos y doce ciclos de tamaño tres. De los restantes 4355 (3.32%) son estados transitivos y 126401 (96.44%) son cúspides, estados que no pueden ser alcanzados desde ningún estado.

El análisis para encontrar los estados estacionarios se efectuó sobre la red simplificada. Se recuperaron los atractores correspondientes a Th0, Th1, Th2, Th17 y Treg. La cuenca de mayor tamaño corresponde a Th0 (51890 estados), le sigue Th2 (37184 estados), Th1 (20574 estados), Th17 (10310 estados) y finalmente Treg (448 estados).

Un análisis de adyacencia (utilizando la biblioteca Networkx de Python) revela que la red puede ser reducida a seis módulos: Th1-Th2, Th17(A) y (B), Treg, IL-2 e IL-1. Los módulos corresponden parcialmente a los genes asociados a cada linaje. La centralidad es una medida de que tan conectado está un nodo. Los nodos con mayor centralidad de grado son: Tbet, GATA-3, IL-4 y RORgt. Los nodos con mayor centralidad de intermediación son: Tbet, IFN $\gamma$ , RORgt e IL-6. Tbet y RORgt son nodos fundamentales en la transmisión de la información dentro de la red.

El comportamiento dinámico de la red puede estar en fase ordenada, caótica o crítica. En la fase ordenada el ruido no se propaga, mientras que en fase caótica el ruido se propaga por toda la red. La zona entre estas dos se conoce como fase crítica y representa un compromiso entre la robustez y la plasticidad. Un análisis de Derrida muestra que la pendiente de la avalancha de perturbaciones es cercana a la recta identidad en el origen. También, el coeficiente de conectividad de la red original es  $K = 1,95$  y su coeficiente de criticalidad es  $K_c = 2,04$ : esto hace que la diferencia entre ambas sea de 4.1 %, lo indica que la red se encuentra cerca de la fase crítica.

---

# Capítulo 1

## Introducción

Entre los problemas de salud de este siglo se encuentran los ocasionados por el virus de inmunodeficiencia humana (VIH) y la diabetes tipo 2. En 2004, el VIH/SIDA y la diabetes mellitus causaron el 3.5 % y el 1.9 % de las muertes a nivel mundial respectivamente [1]. La disminución de la incidencia de ambos se encuentra dentro de los objetivos de la Organización Mundial de la Salud. Sin embargo, a pesar de los recursos dedicados a combatirlos todavía no se tiene una cura.

Ambos problemas de salud han sido relacionados con anomalías en las células Th. En el caso del VIH hay una disminución en las células Th1 y una polarización hacia Th2 [2, 3]. La diabetes tipo 2 se ha relacionado con un ambiente inflamatorio, un incremento de las células Th1 y un decremento en las células Treg [4-8].

El SIDA y la diabetes son complejos dado que sus causas y su progreso dependen de múltiples factores. Por esto es de vital importancia entender el funcionamiento del sistema inmune en general y de las células Th en particular. Entender los mecanismos de polarización de los distintos linajes Th puede ayudar a generar blancos terapéuticos y tratar estos problemas de salud.

Las células T CD4+ coadyuvantes (Th) carecen de capacidad citotóxica, fagocítica o de inducir apoptosis en células infectadas. Sin embargo, estas células se encargan de activar y dirigir a otras células del sistema inmune, regulando la respuesta in-

munológica. Inicialmente las células Th se encuentran en un estado virgen (Th0). Al presentarse un reto inmunológico estas células son activadas en respuesta a la presentación de antígenos específicos. Dependiendo de interacciones célula-célula, la afinidad por el MHC y el tipo de citocinas en el entorno, las células Th0 se diferenciarán en distintos linajes como Th1, Th2, Th17 o T reguladoras (Treg).

Cada linaje expresa reguladores transcripcionales y citocinas características y regula respuestas inmunes particulares. Las células Th1 son las responsables de la defensa contra patógenos intracelulares y la autoinmunidad y expresan T-bet e  $IFN\gamma$ . Las células Th2 dirigen la defensa contra parásitos extracelulares, alergias y asma y expresan GATA-3 e IL-4. Las células Th17 median la respuesta a bacterias extracelulares y hongos además de causar autoinmunidad y expresan  $ROR\gamma_t$  e IL-17. Las células Treg mantienen la tolerancia inmune y expresan Foxp3 y  $TGF\beta$  [9]. Sin embargo, la regulación de la diferenciación de estas células no es completamente conocida.

Con el propósito de analizar la diferenciación de los linajes Th se generó una red booleana a partir de los datos de regulación genética e interacciones proteicas disponibles en la literatura. Diversos autores han utilizado redes booleanas para analizar procesos de regulación genética. Estas redes han sido de gran utilidad para entender la diferenciación celular en *Drosophila* [10] y *Arabidopsis thaliana* [11].

En el caso de las células Th, varios autores han estudiado su regulación y diferenciación utilizando redes. Originalmente se publicaron trabajos sobre la diferenciación Th1/Th2 [12], en fechas más recientes se han agregado citocinas y factores transcripcionales para poder estudiar también a las células Th17 y Treg [13, 14]. El enfoque de redes también ha permeado a las revisiones [15, 16].

Originalmente se construyó una red de 43 nodos y 84 interacciones, que fue simplificada a una red de 17 nodos y 51 interacciones. Los nodos son IL-2, IL-4, GATA-3, IL-12, IL-18,  $IFN\gamma$ , T-bet, IL-1, IL-6, IL-17, IL-21, IL-23, STAT-3,  $ROR\gamma_t$ ,  $TGF\beta$  y Foxp3. En el caso de  $TGF\beta$  se generaron dos nodos, uno  $TGF\beta_{low}$  y otro  $TGF\beta_{high}$ , correspondiendo al nivel de expresión. Esto es debido a que al generar un solo nodo  $TGF\beta$  fue imposible recuperar los estados correspondientes a Th17.

Una vez descrita la red se evaluaron los  $2^{17}$  (131072) estados posibles. De estos estados el 96.44% son cúspides. De los estados restantes 4355 (3.32%) son transitivos, hay 256 (0.19%) soluciones puntuales y 60 (0.05%) estados que corresponden a ciclos de tamaño dos y tres. La red tiene una  $K = 1,95$  y una  $K_c = 2,04$  lo cual la coloca cerca de la fase crítica. Se necesitaron en promedio 2.53 iteraciones para llegar a una solución puntual o ciclo, siendo el mínimo una y el máximo seis.

Dentro de las soluciones de la red se encontraron las correspondientes a los linajes Th0, Th1, Th2, Th17 y Treg, además de otros estados reportados en la literatura. Se definió que una solución pertenece a un linaje si expresa su regulador transcripcional maestro y sus citocinas características. A partir de estos criterios se determinó que la cuenca de atracción de mayor tamaño corresponde a Th0 (51890 estados), seguida por Th2 (37184 estados), luego Th1 (20574 estados), después Th17 (10310 estados) y finalmente Treg (448 estados).

La red puede ser reducida a seis módulos: Th1-Th2, Th17(A) y (B), Treg, IL-2 e IL-1. Los módulos corresponden parcialmente a los genes de cada linaje. Los nodos con mayor centralidad de grado son: Tbet, GATA-3, IL-4 y ROR $\gamma$ t. Los nodos con mayor centralidad de intermediación son: Tbet, IFN $\gamma$ , ROR $\gamma$ t e IL-6. Esto hace que Tbet y ROR $\gamma$ t sean nodos fundamentales en la transmisión de la información dentro de la red.

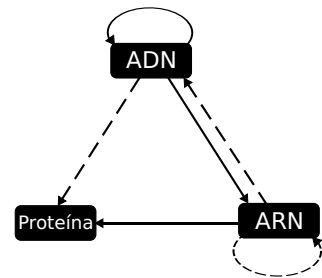
Como resultado de los distintos análisis se determinó la dinámica, las cuencas de atracción y la modularidad de la red. Además, las soluciones obtenidas reflejan las observaciones experimentales reportadas en la literatura. Sin embargo, aún falta mucho por hacer ya que falta incluir algunos grupos de genes, citocinas y otros factores que afectan la diferenciación de las células Th.

---

## Capítulo 2

# Del gen a la proteína <sup>[17]</sup>

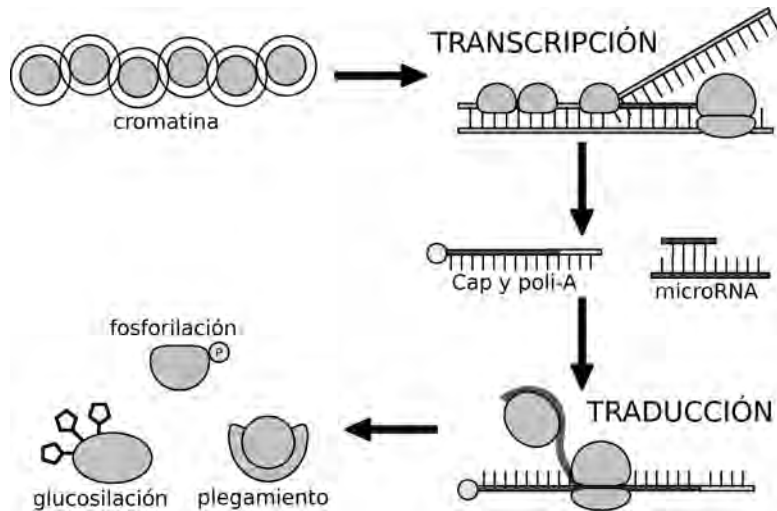
Los seres vivos poseen una organización interna que les permite mantener un estado de homeostasis, reaccionar ante el ambiente, obtener energía de los nutrientes, crecer y reproducirse. El como se hereda y mantiene la organización de los seres vivos ha sido ampliamente estudiado. En 1952 A. D. Hershey y M. Chase realizaron una serie de experimentos con el bacteriófago T2 donde demostraron que el ADN transporta la información genética [19]. Esto motivó el estudio de la estructura del ADN y el código genético y demás mecanismos de transmisión de la información.



**Figura 2.1:** El dogma central de la biología modificado (tomada de Wikimedia [18])

El dogma central de la biología molecular [Fig: 2.1] propone que el ADN puede replicarse, transmitiendo la información genética a próximas generaciones; el ADN puede también transcribirse a ARN mensajero (mARN) y este ser traducido a proteína. El esquema ha sido modificado para expresar los mecanismos de los virus de ARN (ARN → ARN), retrovirus (ARN → ADN) y sistemas *in vitro* de traducción de ADN (ADN → proteína)

Los diferentes niveles de expresión de las proteínas son ocasionados por mecanismos de regulación [Fig: 2.2] como la epigenética, la iniciación de la transcripción, procesamiento del transcrito, regulación por microARNs, traducción del ARN, plegamiento, transporte al citoplasma o modificaciones post-traduccionales. Aun cuando se tiene una proteína funcional puede ser regulada por otras proteínas por medio de interacciones proteína-proteína o de fosforilación.



**Figura 2.2:** Existen diversos mecanismos que regulan la transmisión de la información. Estos incluyen la epigenética, factores transcripcionales, adición de caps y colas de poli-A, miARNs, plegamiento, fosforilación y modificaciones post-transcripcionales.

## 2.1. Regulación génica

Los seres vivos regulan de manera diferencial la expresión de genes y proteínas. Los organismos necesitan activar o reprimir la expresión de las proteínas en respuesta a los estímulos del medio. En el caso de los organismos pluricelulares cada tipo de célula diferenciada produce un conjunto de proteínas específico bajo el principio de equivalencia nuclear: los núcleos de todas las células adultas contienen la misma información genética pero la expresan de forma distinta. Algunas proteínas

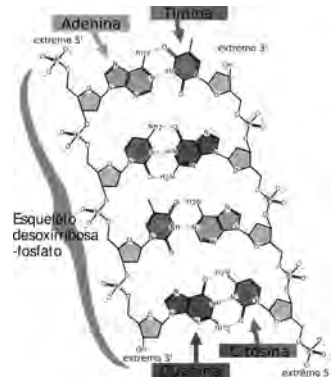


son transcritas constitutivamente, ya que son necesarias para el mantenimiento de la célula. Otras sólo son necesarias en algunos tipos celulares, por lo que se expresan de forma selectiva. No todas las células siguen el principio de equivalencia nuclear. Los gametos sexuales son haploides, contienen sólo una copia de cada cromosoma. La recombinación V(D)J de linfocitos B y T asegura una alta variabilidad en las inmunoglobulinas y el MHC.

Los procesos que controlan el nivel de expresión de genes en cada tipo de célula diferenciada dependen de señales provenientes de la célula y su entorno, que interactúan con el ADN, ARN y proteínas. La regulación actúa de forma diferente sobre los genes dependiendo de las necesidades de cada tipo celular. Un gen puede ser regulado a nivel cromatina, por factores transcripcionales. Una vez que se ha convertido en mRNA puede ser regulado por proteínas que reconocen los extremos 5' y 3' además de por microARNs, entre otros.

### 2.1.1. Transcripción

El ácido desoxirribonucleico (ADN) es una macromolécula que transmite la información genética de las células. Es un polinucleótido conformado por dos hebras de nucleótidos: una desoxirribosa unida a una base nitrogenada (adenina, timina, citosina o guanina) y un grupo fosfato. Las dos hebras están unidas por puentes de hidrógeno entre las bases complementarias (A-T, C-G) [Fig: 2.3]. Dentro de las células eucariontes el ADN está organizado en estructuras formadas por la adición de proteínas conocida como cromatina.



**Figura 2.3:** Estructura del ADN (tomada de Wikimedia [20])

El ácido ribonucleico (ARN) es una cadena de ribonucleótidos lineal y de hebra sencilla. La hebra consta de nucleótidos: una ribosa unida a una base nitrogenada (adenina, uracilo, citosina o guanina) y un grupo fosfato. Tiene diversas funciones en las células: participa en las etapas intermedias de la síntesis proteica, regula la expresión génica y algunos tienen actividad catalítica.

La primera fase para la expresión de genes es el proceso de transcripción, en el cual la información contenida en el ADN, se copia en mARN dentro del núcleo. En la transcripción la ARN polimerasa crea una secuencia complementaria de mARN a partir de la hebra original de ADN. Se inicia cuando la ARN polimerasa reconoce una secuencia promotora al inicio de un gen y termina con el reconocimiento de una secuencia de paro. La regulación de esta etapa está determinada por: la secuencia del ADN, la cromatina, ARNs, factores transcripcionales entre otros.

Una vez que el ADN ha sido transcrito por la ARN polimerasa en mARN está sujeto a una serie de modificaciones. Éstas incluyen el corte de intrones, la adición de caps, otros ARNs y el propio plegamiento del mARN. Cuando el mARN ha sido correctamente procesado puede ser traducido en proteína por el ribosoma.

### **Factores Transcripcionales**

Los factores transcripcionales son proteínas que se unen a secuencias específicas de ADN. Algunas de estas secuencias son promotores ascendentes, intensificadores y sitios de inicio de transcripción. Las proteínas pueden ser activadoras, ayudando al reclutamiento de la polimerasa, o represoras, evitando la transcripción del ADN. La expresión y activación de factores transcripcionales varía según las señales externas o internas, además de la posición espacial y temporal. Algunos sólo se encuentran en ciertos tipos celulares, otros se expresan en respuesta a señales del entorno, en ciertos periodos del ciclo celular o dependen de la ubicación espacial.

### **Epigenética**

El ADN se encuentra acompañado de una serie de proteínas que sirven para su empaquetamiento, organización y protección. La cromatina consiste de ADN y las proteínas unidas a él. Para que un gen sea transcrito la cromatina debe permitir el acceso a los factores transcripcionales que realizan cambios adicionales para permitir el acceso a la polimerasa. Aproximadamente el 99 % del genoma son nucleosomas: octámeros de histonas rodeados de 146 pb. El 1 % restante está accesible donde los complejos regulatorios han abierto el ADN y desplazado los nucleosomas [21]. Los cambios en la estructura de la cromatina se conocen como regulación epigenética.

## Intrones

En eucariontes los genes contienen intrones, secuencias que no codifican una proteína introducidas dentro del ADN. Los genes, incluyendo intrones, se copian en el primer transcripto, llamado pre-mARN. Los intrones son removidos, uniendo los exones por medio de secuencias de bases especiales dentro y alrededor de los intrones reconocidos por los complejos ribonucleoproteínicos nucleares pequeños (snRNP) o por ribozimas. No siempre se unen los mismos exones en el mismo orden, modificaciones en el uso de exones pueden generar distintos péptidos.

## Cap y cola

Para ser traducido, el mARN debe modificarse para poder viajar del núcleo al citoplasma. Estas modificaciones incluyen la adición de un cap y una cola poli-A. La adición de un cap ocurre al principio de la síntesis. Consiste en la unión de una molécula de 7-metilguanilato a la cola 5' del transcripto, con una unión tipo 5' a 5'. La adición de la cola de poli-A ocurre posteriormente a la síntesis. Consiste en unir al extremo 3' una secuencia de 100 a 250 adeninas por medio de la poli-A polimerasa. Estas modificaciones protegen contra la degradación de RNAsas, mejoran la estabilidad y señalizan la exportación por medio de los poros nucleares.

### 2.1.2. Traducción

Las proteínas son macromoléculas formadas por cadenas lineales de aminoácidos. Tienen funciones estructurales, reguladoras, transportadoras, enzimáticas entre otras. Las proteínas se pliegan obteniendo una estructura que les da sus propiedades funcionales. El plegamiento de una proteína puede ser afectado por otras proteínas, la temperatura, el pH, etc. La traducción es la fase intermedia para la expresión de genes. El ribosoma traduce el mARN del extremo 5' al 3' leyendo un codón a la vez. Los codones constan de tres bases nitrogenadas y cada uno codifica uno de los 20 aminoácidos. La traducción se divide en tres etapas: iniciación, alargamiento o elongación, y terminación. En la iniciación se recluta al ribosoma y éste se coloca en

el codón de inicio AUG. En la elongación se añaden otros aminoácidos al polipéptido mediante enlaces peptídicos. En la terminación se reconocen los codones de paro UAA, UGA o UAG y se disocia el ribosoma.

### **Plegamiento del ARN**

Al igual que en el ADN, las bases del ARN son complementarias entre si (A-U, C-G). Esto puede promover que el ARN se pliegue sobre si mismo adquiriendo estructuras secundarias y terciarias. Este plegamiento impide el reclutamiento del ribosoma, inhibiendo la traducción de ARN. Esto permite que el ARN sea regulado tanto por proteínas como por temperatura, ya que incrementos en la temperatura pueden desplegar el ARN. Cabe destacar que el plegamiento del ARN puede conferirle actividad catalítica, como en el caso de las ribozimas.

### **ARNs pequeños**

Los ARNs pequeños tienen menos de 30 nucleótidos y suprimen la expresión de genes específicos. Dentro de estos se encuentran los miARNs, siARNs y piwiARNs. Los miARNs son de 21-25 nucleótidos derivados de precursores de mayor tamaño que forman estructuras tallo-rizo. En animales los miARNs se unen al UTR 3' inhibiendo la traducción y facilitando la degradación. Los siARNs son de 21-25 nucleótidos, sin embargo son originados por el corte de un ARN de doble cadena de 500 pb. Estos siARNs son incorporados al RISC (ARN-induced silencing complex) el cual corta el mARN complementario [22].

#### **2.1.3. Post-traducción**

Que el mARN haya sido traducido a proteína no implica que ésta sea capaz de realizar sus funciones. La proteína debe unirse a cofactores y proteínas que le permitirán realizar su función y sufrir una serie de modificaciones como plegamiento, cortes por proteasas, glicosilaciones y otras modificaciones post-traduccionales. También necesita ser transportada al sitio donde llevará a cabo sus funciones. Todo esto hace que las proteínas estén ampliamente reguladas.

## **Plegamiento proteico**

Para ser funcional una proteína debe plegarse en una estructura tridimensional específica. La información necesaria para la maduración de una proteína está determinada en su secuencia de aminoácidos. Una proteína tiende a plegarse en la estructura compacta con la menor energía libre posible. Para asegurar que la proteína adquiera una conformación adecuada se utilizan chaperonas. Las chaperonas hidrolizan ATP mientras se unen y despegan a la proteína blanco, de esta forma cambian su conformación estructural favoreciendo una estructura funcional.

## **Transporte**

La mayoría de las proteínas son sintetizadas por ribosomas en el citosol, sin embargo, algunas proteínas necesitan ser transportadas a otros organelos. El transporte de proteínas se logra gracias a señales en la secuencia de aminoácidos que se utilizan para dirigir a las proteínas. Existen dos tipos de señales: secuencias de 15-60 a.a. que son removidas al llegar al compartimiento y parches generados por la estructura tridimensional de la proteína. El transporte puede realizarse a través de poros entre compartimientos que permiten sólo el paso de proteínas con las señales correctas. Existen también translocadores transmembranales que transportan directamente proteínas específicas. Finalmente las proteínas pueden ser transportadas en vesículas.

## **Adición de grupos funcionales**

Después de la traducción, a las proteínas se les puede agregar una serie de grupos químicos funcionales, los cuales cambian las características químicas, estructurales y funcionales de las proteínas. Estas modificaciones incluyen la adición de grupos acetato, lípidos o carbohidratos y/o la modificación de ciertos aminoácidos. Muchas de estas reacciones se llevan a cabo en el aparato de Golgi.

## 2.2. Interacciones proteína-proteína

Desde la publicación en 1976 del genoma completo del ARN del virus del bacteriófago MS2 [23] se ha incrementado la cantidad de genomas secuenciados y esta cantidad crece día con día. Estos incluyen los genomas de organismos modelo como *Escherichia coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana* y el propio genoma del ser humano. Gracias a estos proyectos de secuenciación se cuenta con una lista casi completa de los genes presentes en un organismo.

Sin embargo, para entender un sistema no es suficiente conocer sus partes, también es necesario conocer las interacciones entre ellas. Para entender las vías metabólicas, de señalización o los circuitos genéticos es necesario entender las interacciones metabolito-proteína, proteína-proteína y proteína-ADN. Esto es especialmente relevante en la construcción de redes de interacción genética. Muchas redes se construyen basadas en datos provenientes de CHIP y microarreglos, los cuales expresan las proteínas unidas a ADN y la cantidad de ARN transcrito. Sin embargo, el transcriptoma no necesariamente predice la cantidad de proteína ni su estado dentro del ser vivo [24].

Las relaciones entre proteínas proveen la base mecánica para la mayor parte de la fisiología y función de los organismos [25]. Las proteínas necesitan interactuar entre ellas de manera específica y ordenada para mantener los procesos metabólicos. Las proteínas no sólo alteran a los metabolitos, si no a las mismas proteínas. Es por eso que al construir una red de regulación es importante no sólo incluir la regulación genética, sino las interacciones proteína-proteína.

La diferenciación de las células Th no es una excepción. Para lograr la expresión de un fenotipo específico es necesario sensar las señales del medio, transmitir las al núcleo, alterar los patrones de expresión genética y posteriormente producir una serie de proteínas efectoras. Para llevar a cabo la expresión del fenotipo las proteínas tienen una gran cantidad de interacciones entre ellas como la fosforilación, la formación de dímeros y la competencia por el sitio.

### 2.2.1. Fosforilación

La fosforilación es una modificación post-transcripcional en la cual se le agregan grupos fosfato a las proteínas, esto cambia su estructura tridimensional y la unión de ligandos. La fosforilación es extensamente usada en cadenas de transducción de señales, de hecho una tercera parte de las aproximadamente 10000 proteínas de un mamífero se encuentra fosforilada en un momento u otro [17].

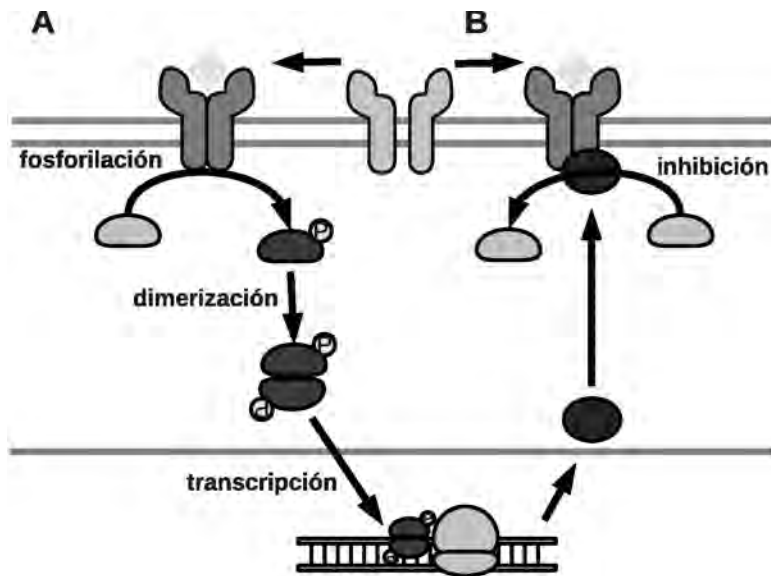
### 2.2.2. Dimerización

La mayoría de las proteínas necesitan unirse a otras para llevar a cabo sus funciones formando complejos proteicos. Un dímero es un complejo formado por dos proteínas no unidas covalentemente. Se dice que es un homodímero si las dos proteínas son iguales y un heterodímero si las dos proteínas son diferentes.

Un ejemplo particular de dimerización es el de los receptores transmembranales. Muchos receptores transmembranales están conformados por dos o más subunidades diferentes. Al unirse el ligando al receptor estabiliza cierta conformación, permitiendo que el dominio catalítico del receptor se active (o inactivo). A continuación la proteína mensajera se une al receptor y es modificada por este, lo que motiva la disociación de la proteína mensajera que procede a transmitir la señal. [Fig: 2.4]. Esto permite regular el receptor a través de la presencia de sus subunidades. También es posible utilizar ligandos que no fomenten los cambios conformacionales adecuados, secuestrando los receptores en un estado inactivo.

### 2.2.3. Competencia sitio-específica

La competencia sitio-específica sucede cuando dos proteínas se pueden unir al mismo sitio de una proteína catalizadora, bloqueando la unión de la otra proteína. Generalmente una de las proteínas es un mensajero y la otra actúa como inhibidor de la señalización. La proteína inhibidora tiende a mantenerse más tiempo unida a la proteína catalizadora, ya que no sufre cambios conformacionales que alteren su afinidad. Este mecanismo puede ser activado por la misma cascada de transducción de señales, sirviendo como un mecanismo de auto-regulación [Fig: 2.4].



**Figura 2.4:** Los receptores están sujetos a diversas interacciones que regulan su actividad. A) Al unirse el ligando al receptor estabiliza estructuras conformacionales que le permiten activar factores transcripcionales a través de reacciones químicas. Estos factores pueden dimerizar y posteriormente migrar al núcleo donde ejercen su acción. B) Proteínas reguladoras impiden la activación de los factores transcripcionales, inhibiendo la acción del receptor.



---

## Capítulo 3

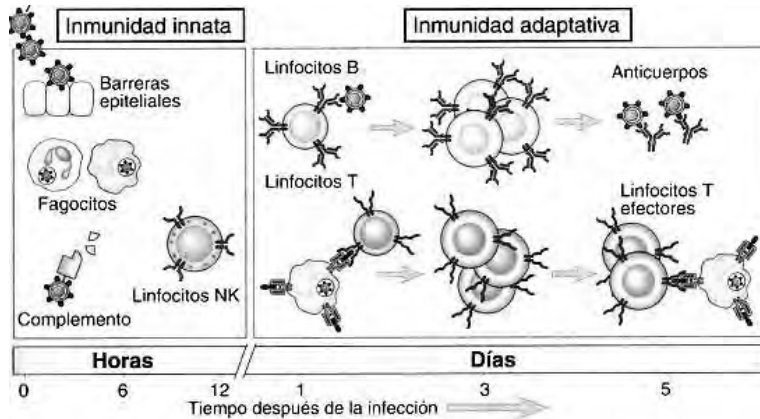
# El sistema inmunológico [9, 26, 27]

El sistema inmune es el encargado de proteger al organismo de infecciones por medio de la identificación y eliminación de patógenos. Los retos inmunológicos abarcan desde virus hasta gusanos parásitos intestinales o células cancerosas, por lo que existen múltiples mecanismos celulares y moleculares que protegen al individuo. La respuesta inmune consta de dos funciones principales: reconocimiento y respuesta. Es necesario detectar con absoluta especificidad a los patógenos y distinguirlos de las células normales del organismo. Una vez que el patógeno ha sido reconocido es necesario montar una respuesta adecuada que permita eliminarlo o neutralizarlo.

El sistema inmune ha desarrollado una variedad de respuestas apropiadas para combatir a cada tipo de reto inmunológico. Estos mecanismos se dividen en dos tipos principales: innatos y adaptativos [Fig: 3.1], cuya diferencia principal reside en las estructuras de reconocimiento de los patógenos y su tiempo de acción.

### 3.1. Inmunidad innata

El sistema inmune innato o no específico se refiere a los mecanismos básicos de resistencia de un organismo. En lugar de reconocer específicamente al patógeno, la respuesta inmune innata reconoce un grupo de patrones moleculares altamente



**Figura 3.1:** La respuesta inmune puede ser innata y adaptativa (tomada de Abbas 2004 [28])

conservados comunes a un grupo o familia entera de patógenos. Estos mecanismos conforman la respuesta inmediata a los patógenos y se mantienen a lo largo de toda la reacción inmunitaria.

Existen cuatro tipos de barreras: anatómicas, fisiológicas, endocíticas y fagocíticas e inflamatorias.

- Las barreras anatómicas impiden físicamente la entrada de patógenos al organismo. Éstas incluyen la piel y las membranas mucosas.
- Las barreras fisiológicas incluyen temperatura, pH, oxígeno y otros factores solubles. Éstas actúan generando ambientes poco favorables para el patógeno.
- Las barreras endocíticas y fagocíticas internalizan y degradan macromoléculas extracelulares dentro de compartimientos específicos de la célula. La mayoría de las células son capaces de endocitosis; los monocitos, neutrófilos y macrófagos se especializan en la fagocitosis.
- Las respuestas inflamatorias incrementan el flujo sanguíneo, la permeabilidad capilar y la cantidad de células fagocíticas y proteínas de fase aguda.

## 3.2. Inmunidad adaptativa

La inmunidad adaptativa o específica difiere de la inmunidad innata en que tiene especificidad, diversidad, memoria y reconocimiento de lo propio y ajeno. La especificidad puede reconocer y responder a una amplia gama de moléculas blanco de manera individual. Su diversidad le permite reaccionar a distintos antígenos. La memoria puede alterar las respuestas a partir de experiencias previas de manera que encuentros subsecuentes con un patógeno originan respuestas más rápidas y vigorosas. Finalmente el reconocimiento de lo propio y ajeno le permite no atacar a los tejidos del cuerpo o reconocer cuando éstos se encuentran infectados.

La respuesta inmune adaptativa no es inmediata, necesita tiempo para madurar y combatir los retos inmunológicos. Por lo anterior no es independiente de la respuesta innata. La respuesta innata desencadena las primeras señales de la inmunidad adaptativa y la inmunidad adaptativa regula y dirige factores específicos de la inmunidad innata aumentando su eficiencia. La generación de una respuesta inmune adaptativa depende en gran parte de dos tipos de células: linfocitos y células presentadoras de antígeno (APC). Los linfocitos se dividen a su vez en linfocitos B y T.

Al entrar un patógeno al cuerpo es reconocido en primera instancia por la respuesta innata. Posteriormente los antígenos son transportados al sistema linfático. El sistema linfático es rico en APCs como células dendríticas, macrófagos o células B. Las células dendríticas y los macrófagos fagocitan a los patógenos, para posteriormente procesarlos y expresar los antígenos en membrana junto con moléculas co-estimuladoras. Los antígenos solubles pueden adherirse a los anticuerpos de las células B. Adicionalmente las células T reconocen los antígenos unidos a moléculas coestimuladoras como el MHC expresado por las APCs.

La unión del antígeno a los receptores de los linfocitos B y T es la señal de activación de éstos. Una vez activados los linfocitos se multiplican clonalmente, generando dos tipos de células: efectoras y de memoria. Las células efectoras tienen un corto tiempo de vida y efectúan y regulan la respuesta al patógeno. Las células de memoria viven largo tiempo y sirven para incrementar la velocidad e intensidad de la respuesta inmune si se vuelve a encontrar al mismo patógeno.

### 3.2.1. Linfocitos

Los linfocitos son generados en la médula durante la hematopoiesis. Los linfocitos B y T tienen receptores de membrana que reconocen antígenos específicos. Una vez que han madurado viajan por el sistema circulatorio y linfático en un estado inactivo. Si una APC les presenta un antígeno que coincida con su receptor de membrana los linfocitos son activados, montando una respuesta inmune. La alta variabilidad de los linfocitos provoca que haya pocos que sean específicos para cada patógeno. Por ello, una vez activados, los linfocitos se multiplican para adquirir el suficiente número para eliminar el patógeno. Este periodo de proliferación requiere al menos una semana para desarrollarse, durante la cual el control de la infección por los mecanismos innatos es esencial para la supervivencia del individuo.

#### Células B

Los linfocitos B maduran en la médula y la dejan expresando un receptor de membrana antígeno-específico. Este receptor se conoce como anticuerpo o inmunoglobulina y es una glicoproteína que puede encontrarse soluble o en membrana. Está conformado de dos cadenas pesadas y dos cadenas ligeras. Cambios en la cadena pesada determinan el isotipo del antígeno, en mamíferos estos son IgA, IgD, IgE, IgG e IgM y desempeñan funciones diferentes. Las cadenas ligeras además del dominio que las une a la cadena pesada cambian de célula a célula, es esta región la que reconoce al antígeno, confiriéndole diversidad y especificidad al antígeno.

Una vez que el anticuerpo ha reconocido a su antígeno específico la célula B empieza a dividirse rápidamente. Sus células hijas se diferencian en células B de memoria y en células plasmáticas. Las células de memoria viven por un largo tiempo de tal manera que si se vuelve a encontrar al patógeno sea posible reaccionar de manera efectiva contra él. Las células plasmáticas viven solo unos días, pero durante ese tiempo expresan una gran cantidad de anticuerpos. Estos anticuerpos no eliminan al patógeno de manera directa sino que facilitan su destrucción por los mecanismos de la inmunidad innata. Los anticuerpos pueden neutralizar a los virus y bacterias evitando la unión a moléculas específicas de la célula blanco y su posterior entrada. También son capaces de aumentar la fagocitosis (opsonización) ya que

son reconocidos por las células fagocíticas. Además pueden activar las proteínas del sistema del complemento.

## Células T

Los linfocitos T se generan en la médula, pero maduran en el timo expresando un receptor llamado TCR. Está compuesto por una cadena  $\alpha$  y una  $\beta$  o por una cadena  $\gamma$  y una  $\delta$ . Las células T no reconocen a los antígenos de forma soluble, sino pequeños fragmentos de ellos asociados a las proteínas de membrana codificadas por el complejo principal de histocompatibilidad, (MHC).

Existen dos poblaciones de linfocitos T: citotóxicas (Tc) y coadyuvantes (Th). Experimentalmente se pueden diferenciar porque las células Tc expresan CD8 en membrana y las células Th expresan CD4. Las células Th reconocen el MHC II de las APC y liberan una serie de factores de crecimiento conocidos como citocinas. Estas citocinas activan a las células B, Tc, fagocíticas y otras células de la respuesta inmune. Cambios en el patrón de expresión de citocinas causan cambios en el tipo de respuesta inmune desarrollada. Las células Tc reconocen el MHC I presente en la mayoría de las células, expresan pocas citocinas pero son capaces de inducir apoptosis en células infectadas o cancerosas.

Durante el proceso de maduración y diferenciación las células T son expuestas a mecanismos de regulación para asegurarse de que no maduren las que reaccionan a moléculas propias del cuerpo. Una vez que las células T han madurado viajan a través del sistema linfático y circulatorio donde son expuestas a diversos antígenos. La mayoría de los antígenos son transportados al sistema linfático donde son procesados por las APCs, que expresan los antígenos junto con el MHC en membrana.

Una vez que una célula T reconoce el antígeno se activa y empieza la respuesta inmune consecuente. El reconocimiento del MHC por el TCR activa la expresión de factores transcripcionales como NFAT y NF- $\kappa$ B, que son vitales en la activación de las células Th. Estos factores en conjunción con señales del APC y citocinas presentes en el ambiente determinan la diferenciación de la célula Th.

### 3.2.2. Células presentadoras de antígeno

Las APC son células especializadas en la presentación de antígenos a las células Th. Las APC incluyen a macrófagos, células B y células dendríticas y se distinguen por expresar el MHC I. Las células B son altamente eficientes en presentar antígenos que se han unido a sus anticuerpos de superficie. Las células dendríticas internalizan a los patógenos por fagocitosis o endocitosis. Una vez que los patógenos han sido procesados estas células expresan parte de los antígenos junto con el MHC I.

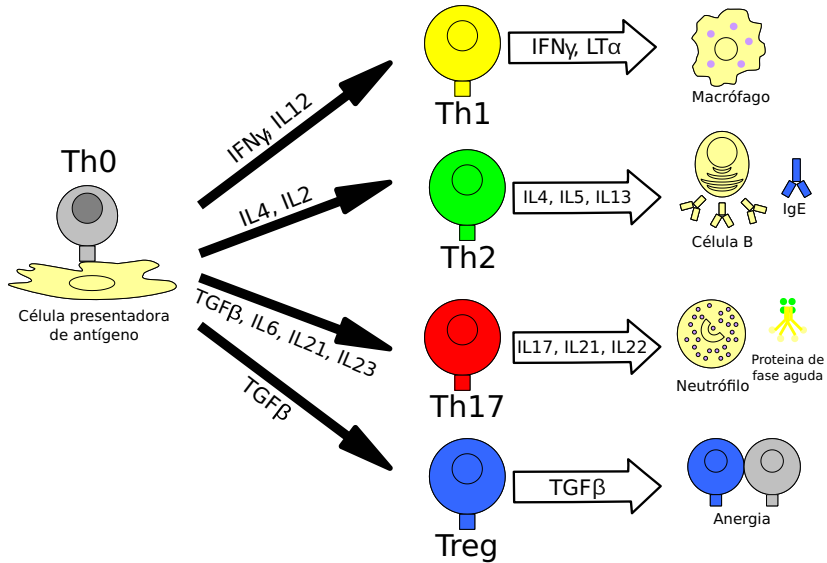
### 3.3. Diferenciación de células Th

Las células Th CD4+ tienen un papel fundamental en el sistema inmune. Esto es debido a su capacidad de promover la generación de anticuerpos por las células B, inducir la actividad microbicida de los macrófagos, reclutar neutrófilos, eosinófilos y basófilos al sitio de la infección y fomentar la respuesta inflamatoria. Todo esto es logrado a través de la producción de citocinas y otros factores de crecimiento.

Sin embargo, no todos los retos inmunológicos requieren el mismo tipo de respuesta inmune. Cada patógeno tiene ciclos de vida particulares y es importante elegir la respuesta inmune que actúe más eficientemente contra el patógeno y cause el menor daño al organismo. En 1986 Mossmann y Coffman mostraron que las células Th vírgenes podían diferenciarse en dos grupos: Th1 y Th2. Cada grupo tenía un patrón particular de producción de citocinas y de funciones [29]. Curiosamente, las células Th vírgenes salían del timo maduras, pero la diferenciación en Th1 y Th2 no se daba hasta que eran activadas por estimulación del TCR.

Durante mucho tiempo se creyó que la respuesta inmune era dicotómica, con las células Th1 encargadas de la respuesta inmune celular y las Th2 de la respuesta humoral. En 1995 Sakaguchi *et.al.* definieron a las células reguladoras Treg como encargadas de controlar las respuestas autoinmunes [30] (había reportes anteriores de este tipo de células, pero no fueron contundentes). En 2006 el grupo de Stockinger definió un grupo de células Th que expresa IL-17 y es diferente a Th1 y Th2, estas fueron después nombradas las células Th17 [31].

Las células Th pueden diferenciarse en cuatro tipos celulares: Th1, Th2, Th17, y Treg [Fig: 3.2] aunque posiblemente existan más. Cada uno de estos linajes es determinado por las señales que recibe la célula Th cuando interacciona inicialmente con el antígeno. Cada tipo celular tiene funciones específicas y expresa un conjunto definido de citocinas y factores de transcripción.



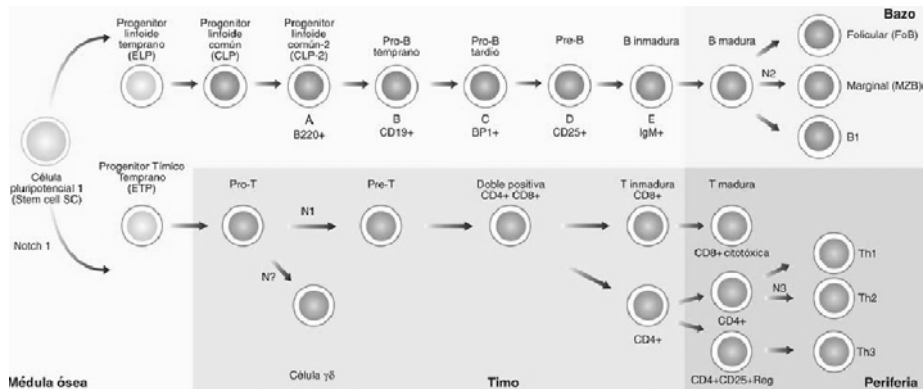
**Figura 3.2:** Las células Th0 pueden diferenciarse en cuatro distintos linajes: Th1, Th2, Th17 y Treg. Cada linaje es inducido por citocinas específicas en el entorno y expresa un patrón de factores transcripcionales y citocinas particular.

### 3.3.1. Ontogenia de células Th

Las primeras células linfoides de un embrión humano aparecen en el tejido hematopoyético, aproximadamente a los dos meses de vida intrauterina. Las células B y T se generan a partir de un subgrupo de células totipotenciales de la médula ósea que expresan Notch. Notch-1 induce al progenitor tímico temprano (ETP), a partir del cual derivarán los linfocitos T. En ausencia de Notch1 se induce al progenitor linfóide temprano (ELP), derivándose los linfocitos B [32].

Una vez en el timo, ETP inicia el rearreglo y la expresión de los genes  $\beta$ ,  $\gamma$  y  $\delta$  del receptor de células T (TCR). En este momento los linfocitos inmaduros no expresan a CD4 ni a CD8, por lo que se les llama linfocitos pro-T o T CD4-CD8-. Una minoría de estas células expresa los receptores  $\gamma\delta$  y se diferencia en linfocitos T  $\gamma\delta$ . El resto de los linfocitos expresan las cadenas  $\alpha\beta$  del TCR y se diferencian en linfocitos pre-T o en células doblemente positivas CD4+CD8+.

A continuación los linfocitos CD4+CD8+ pasan por la selección positiva, que consiste en unirse con baja afinidad a los péptidos presentados por el MHC. Los linfocitos con afinidad al MHC I se diferenciarán en células T o CD4-CD8+. Los linfocitos con afinidad por el MHC II se diferenciarán en células Th o CD4+CD8-. Posteriormente los linfocitos pasan por la selección negativa, donde las células que reconocen con alta avidez o afinidad los péptidos propios son eliminadas, asegurándose así la destrucción de células potencialmente reactivas. Las células T  $\alpha\beta$  con un solo marcador CD4+ o CD8+ salen a los tejidos linfoides periféricos donde podrán ser activadas, proliferar y diferenciarse [28] [Fig: 3.3].



**Figura 3.3:** Ontogenia de los linfocitos T y B (tomada de Anaya 2005 [32])

Una vez que las células T han completado su desarrollo en el timo, entran al torrente sanguíneo. Al llegar a un órgano linfocitario periférico dejan la sangre y migran a través de los tejidos linfocitarios, retornando vía linfática al torrente sanguíneo para recircular entre la sangre y tejidos linfocitarios periféricos. Las células T recirculantes maduras que no han tenido contacto con su antígeno específico se conocen como



células T vírgenes o Th0. Para participar en una respuesta inmune adaptativa estas células deben reconocer a su antígeno específico presentado como un péptido asociado a la molécula del MHC en la superficie de la célula presentadora de antígenos, lo que la induce a diferenciarse y proliferar.

### 3.3.2. Th0

Las células Th0 o vírgenes son células Th CD4+ a las que aún no se les ha presentado un péptido y por lo tanto no han sido activadas ni se han diferenciado. La cantidad y calidad de células Th0 se ven alteradas en síndromes como el VIH y CVID (Common Variable Immunodeficiency) [33].

#### Producción de citocinas

Las células Th0 expresan una amplia gama de citocinas. Para la mayoría de los ensayos *in vitro* se supone que las células Th0 no producen citocinas características como IL-4, IFN $\gamma$ , IL-17, IL-10 y TGF $\beta$ . Sin embargo, se han reportado células Th0 que expresan combinaciones de IFN $\gamma$ , IL-4, IL-5, IL-6, IL-10, IL-17 y GM-CSF [34] las cuales se relacionan generalmente con células Th1, Th2, Th17 y Treg.

Este patrón de expresión de citocinas podría deberse a la regulación de las proteínas SOCS. En linfocitos Th0 la expresión de SOCS3 disminuye, mientras que la expresión de SOCS1 y SOCS2 aumenta. Esto sugiere que la regulación de las proteínas SOCS es diferente entre los linfocitos Th0 y los linfocitos activos, llevando a diferentes patrones de represión de citocinas [35]. Además, SOCS1 regula la sensibilidad a IL-7 e IL-15 en células Th0, pero no en células Th activadas [36].

#### Diferenciación

Las células Th0 expresan una serie de moléculas de membrana como CD45RA y CD62L. También carecen de marcadores de activación como CD25, CD44, CD62L y CD69. IL-7 y su transductor STAT5 es necesario para la supervivencia de las células Th0, IL-2 favorece la diferenciación hacia células efectoras e IL-15 hacia células de memoria [36].

### 3.3.3. Th1

Las células Th1 median las respuestas contra patógenos intracelulares como *mycobacteria*, favorecen la proliferación de células Tc, el reclutamiento de macrófagos y el incremento en su actividad microbicida. Además, está asociado con una serie de enfermedades autoinmunes como la hipersensibilidad tipo 4, esclerosis múltiple y artritis reumatoide.

#### Producción de citocinas

Los principales productos de las células Th1 son  $\text{IFN}\gamma$ , linfotoxina  $\alpha$  ( $\text{LT}\alpha$ ) e IL-2.  $\text{IFN}\gamma$  fomenta la activación de macrófagos.  $\text{LT}\alpha$  está relacionada con respuestas inflamatorias, antivirales y autoinmunes como la progresión de la esclerosis múltiple. IL-2 es importante en la maduración de células de memoria CD4 y CD8, además de la diferenciación de las células Th2.

#### Diferenciación

T-bet es el regulador maestro de las células Th1, fomentando la expresión de  $\text{IFN}\gamma$  y  $\text{LT}\alpha$ . T-bet puede ser activado por NFAT y  $\text{NF-}\kappa\text{B}$  junto con STAT1, transductor de  $\text{IFN}\gamma$  o por si mismo [37], formando un ciclo de autoinducción.  $\text{IFN}\gamma$  es fomentado por T-bet y en menor medida por Eomes. STAT4, el transductor de IL-12 sinergiza con la señal de IL-18 para producir  $\text{IFN}\gamma$ . T-bet es un remodelador de la cromatina de  $\text{IFN}\gamma$  y se necesita de la señalización del TCR para que  $\text{IFN}\gamma$  sea expresado [37].

Las células Th1 sobrerregulan IL-12R $\beta$ 2 en respuesta a la activación del TCR y de  $\text{IFN}\gamma$ . (La cadena IL-12R $\beta$ 1 es expresada constitutivamente por las células Th). IL-18R $\alpha$  es también sobreexpresada en la diferenciación de células Th1. IL-18 e IL-12 sinergizan para expresar  $\text{IFN}\gamma$ , lo cual podría explicar la sobrerregulación de sus receptores.

### 3.3.4. Th2

Las células Th2 median las respuestas contra patógenos extracelulares como helmintos. Favorecen la proliferación de células B y el cambio de isotipo a IgE. El tipo está asociado con una serie de enfermedades autoinmunes como la hipersensibilidad tipo 1 o alergia mediada por anticuerpos y el asma.

#### Producción de citocinas

El principal producto de las células Th2 es IL-4, aunque también producen IL-5, IL-9, IL-13, IL-25 y anfiregulina. IL-4 favorece el cambio de isotipo a IgE y la diferenciación en Th2. IgE se une a basófilos y mastocitos a través de FcεRI, provocando que secreten histaminas, serotonina, IL-4, IL-12 y TNF-α. IL-5 recluta eosinófilos. IL-9 induce la producción de mucinas durante las reacciones alérgicas. IL-13 fomenta la expulsión de helmintos y la hipersensibilidad de vías respiratorias. IL-25 pertenece a la familia de IL-17 y participa en la amplificación de la respuesta Th2.

#### Diferenciación

GATA-3 es el regulador maestro de las células Th2 y es inducido por STAT6, el transductor de IL-4. También IL-4 incrementa la expresión de GATA-3. A su vez, GATA-3 es capaz de fomentar su propia expresión, además de estar asociado con la elección de linaje a nivel del timo, que está especialmente expresado en la población transicional CD4+CD8*bajo*. La sobreexpresión de esta proteína en el timo fetal favorece la polarización hacia CD4, mientras que su eliminación favorece la diferenciación a CD8.

IL-4 puede activarse a si mismo a través de su transductor STAT6. GATA-3 es necesario pero no suficiente para fomentar la inducción de IL-4, ya que sólo modifica la cromatina del locus de IL-4 [37]. Es necesaria la ayuda de NFAT, NF-κB, c-Maf o el transductor de IL-2, STAT5 [38, 39] para lograr la expresión de IL-4. Por otro lado GATA-3 es necesario y suficiente para lograr la expresión de IL-5 e IL-13. IL-4 induce a Gfi-1.

IL-4R $\alpha$  es sobrerregulado por IL-4. CD25 o IL-2R $\alpha$  es también sobrerregulado en las células Th2, posiblemente por la acción de c-Maf. Esto puede incrementar la sensibilidad de las células Th2 a IL-2. IL-33R $\alpha$  es el principal marcador de superficie de las células Th, su función podría ser análoga a la de IL-18R $\alpha$  en las células Th1.

### 3.3.5. Th17

Las células Th17 median las respuestas contra bacterias extracelulares y hongos. Favorecen la respuesta inflamatoria y el reclutamiento de neutrófilos. Están asociadas con una serie de enfermedades autoinmunes como la encefalomiелitis, psoriasis, artritis reumatoide e inflamaciones gástricas [40].

#### Producción de citocinas

Los principales productos de las células Th17 son IL-17a, IL-17f, IL-6, IL-21 e IL-22. IL-17a y IL-17f se encuentran en el mismo locus y generalmente son coexpresados y se unen al mismo receptor IL-17RA, aunque IL-17a tiene una mayor afinidad, en general se refiere a ambos como IL-17. Ambos inducen diversas citocinas inflamatorias como IL-6 además de reclutar y activar neutrófilos. IL-21 activa células Tc, B, NK, dendríticas y fomenta la diferenciación de Th17. IL-22 fomenta la inflamación dérmica, y la defensa contra bacterias como *K. pneumoniae*, aunque esta última podría ser inducida por IL-23.

#### Diferenciación

ROR $\gamma$ t es el regulador maestro de las células Th17 e induce IL-17 y a IL-21. ROR $\gamma$ t es inducido por STAT3 y por el transductor de IL-1. También es inducido por bajas concentraciones de TGF $\beta$ , pero concentraciones altas de este lo inhiben, aunque no es claro si es directamente o indirectamente a través de Foxp3. ROR $\alpha$  también es capaz de inducir IL-17 aunque en menor medida que ROR $\gamma$ t. IL-6, IL-21, e IL-23 son todas citocinas pro-inflamatorias que señalizan por medio de STAT3. Además, STAT3 es indispensable para la inducción de IL-17 e IL-23R. Otro factor

que interviene en la diferenciación de las células Th17 es el receptor de hidrocarburos de arilo (AHR), el cual parece participar en la regulación de IL-22.

Las células Th17 expresan altos niveles de IL-23R, e IL-1R. En humanos IL-1 es necesario para la diferenciación de las células Th17, aunque en modelos murinos e *in vitro* esto no es tan claro. IL-17A activa quinasas reguladas por señales extracelulares (ERK1 y ERK2), Jun y p38 en diversos tipos celulares [41]. IL-1 activa una gran variedad de vías proteicas como Myd88, IRAKs, TRAF6, NF $\kappa$ B y MAPK [42]

### 3.3.6. Treg

Las células Treg median la auto-tolerancia. El incremento de células Treg en ratón previene el rechazo de implantes y combate enfermedades autoinmunes. La falta de células Treg está fuertemente relacionada con enfermedades autoinmunes, aunque también incrementa la inmunidad contra tumores e infecciones crónicas.

### Producción de citocinas

Los principales productos de las células Treg son TGF- $\beta$ , IL-10 e IL-35. TGF- $\beta$  participa en la anergia, la tolerancia oral, la diferenciación de las células Treg y Th17. TGF- $\beta$  regula múltiples procesos, las proteínas SMAD constituyen los componentes básicos de la cascada de señalización intracelular. Estudios recientes revelan que las proteínas SMAD y los receptores transmembranales de TGF $\beta$  están regulados [43]. IL-10 tiene funciones antiinflamatorias y anérgicas. Sin embargo no todas las células Treg producen IL-10 necesariamente. De la misma manera, IL-10 puede ser producido por células Th1, Th2 y Th17. Esto hace que sea posible que IL-10 sea una citocina de autoregulación propia de las Th. IL-35 comparte la subunidad ED13 con IL-12 e IL-27 y participa en la actividad supresora. Las Treg pueden además promover la anergia de las células Th a través del contacto célula-célula.

### Diferenciación

Foxp3 es el regulador maestro de la célula Treg. TGF- $\beta$  induce la expresión de Foxp3 y viceversa. El nivel de TGF- $\beta$  necesario para inducir Foxp3 es mayor al

necesario para inducir  $ROR\gamma_t$ .  $STAT5$  participa en la inducción de  $Foxp3$  uniéndose a su promotor.  $TGF-\beta$  y  $Foxp3$  no son necesarios ni suficientes para expresar  $IL-10$ . La mayoría de las Tregs expresan  $IL-2R\alpha$  (CD25) lo cual apunta a la importancia de  $IL-2$ . También expresan  $CTLA-4$ ,  $GITR$  y  $Folr4$  aunque no se conoce su función o relevancia.

### 3.3.7. Regulación cruzada

La diferenciación de células Th no solo está regulada por ciclos de autoinducción como los generados por  $IFN\gamma$ ,  $IL-4$ ,  $IL-21$  o  $TGF\beta$ . Los diferentes linajes también se inhiben entre ellos. Esto se logra a través de la represión por factores transcripcionales, por competencia sitio-específica en los receptores de citocinas por las proteínas SOCS y por factores epigenéticos.

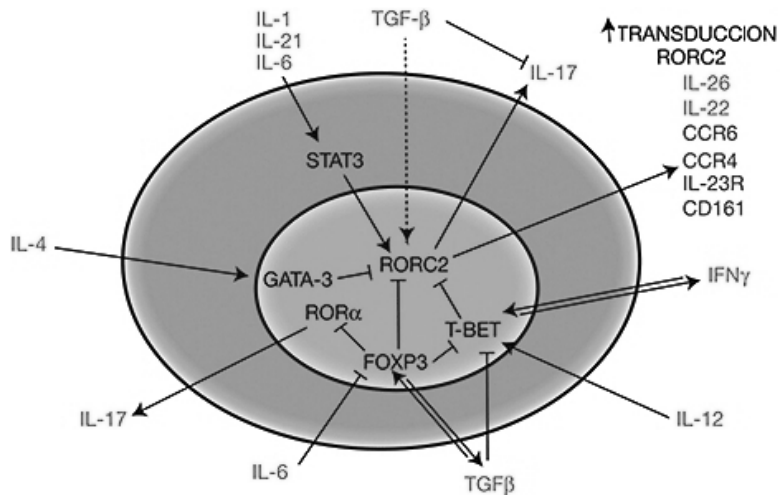
#### Factores transcripcionales

Existen varias vías de inhibición dentro de las células Th [Fig: 3.4]. La regulación entre Th1 y Th2 es mediada por la competencia entre sus citocinas  $IFN\gamma$  e  $IL-4$  y entre sus factores transcripcionales T-bet y GATA-3. T-bet inhibe la diferenciación en Th2 mediante la fosforilación de tirosinas de GATA3, impidiendo que éste se una al ADN [44]. T-bet y GATA3 también son inhibidos por  $Foxp3$  y  $TGF\beta$ .  $IFN\gamma$  e  $IL-4$  se inhiben indirectamente a través de sus factores transcripcionales y de su activación de T-bet y GATA3. Curiosamente la competencia entre estas dos citocinas podría ser también directa, ya que se han reportado interacciones intercromosomales entre los genes de ambas citocinas [9].

La diferenciación de Th17 está también regulada. La adición de  $IFN\gamma$  o  $IL-4$  inhibe la expresión de  $ROR\gamma_t$ , probablemente a través de T-bet y GATA3.  $TGF\beta$  favorece la expresión de  $ROR\gamma_t$  (Th17) y de  $Foxp3$  (Treg) esto ha generado un interesante mecanismo de regulación.  $Foxp3$  se une a  $ROR\alpha$  y  $ROR\gamma_t$  e inhibe su activación en una manera dosis-dependiente.  $TGF\beta$  lleva a la rápida inducción de  $ROR\gamma_t$ , pero la unión de  $ROR\gamma_t$  al promotor de  $IL-17$  es suprimida por su interacción con  $Foxp3$ . Al añadir  $IL-6$  o  $IL-21$  los niveles de  $Foxp3$  disminuyen,

impidiendo su acción inhibitoria. IL-6 no altera la expresión de ROR $\gamma$ t, pero reprime a Foxp3 permitiendo la expresión de IL-17 [45]. TGF $\beta$  sólo en concentraciones de 1 ng/ml no induce a IL-17. Sin embargo, TGF $\beta$  a 10 ng/ml en combinación con IL-1, IL-21 o IL-23 promueve ROR $\gamma$ t e IL-17. Mientras tanto 50 ng/ml suprimen a IL-17, probablemente a través de la activación de Foxp3 [46].

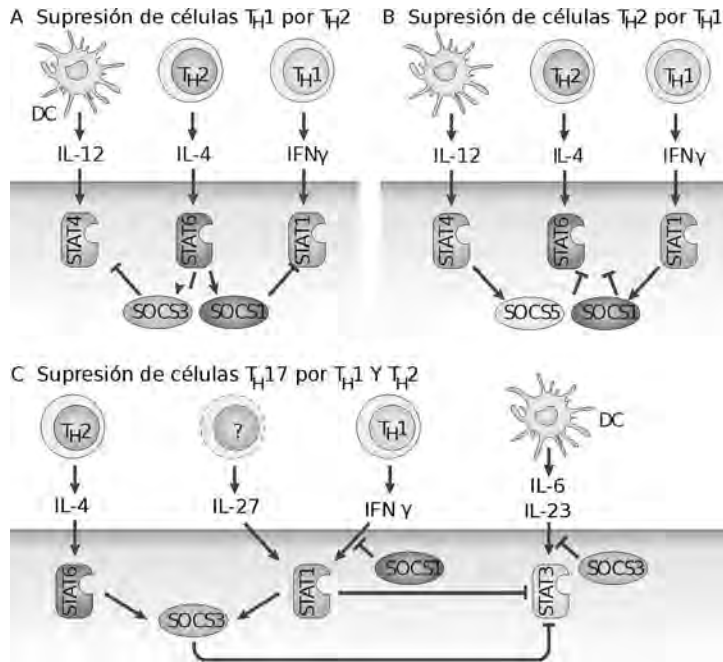
El caso de IL-10 es interesante. IL-10 es una citocina antiinflamatoria. Está principalmente relacionada con las células Treg, sin embargo puede ser producida por Th1, Th2 y Th17. IL-10 señala a través de STAT-3, al igual que IL-6, IL-21 e IL-23, lo cual es paradójico, dado que estas son citocinas pro-inflamatorias. En las células Treg la expresión de Foxp3 y TGF $\beta$  no es necesaria ni suficiente para la expresión de IL-10. También es la citocina característica de las células Th3, sin embargo no se sabe si este es un verdadero linaje o es el producto de la expresión de IL-10 por Th1, Th2 o Th17 [47]. Esto lleva a suponer que IL-10 es una citocina propia de todas las células Th que sirve como mecanismo de autoregulación.



**Figura 3.4:** Regulación cruzada de los factores transcripcionales de las células Th (tomada de Crome 2010 [40])

### Proteínas SOCS

Las proteínas supresoras de señalización de citocinas (SOCS) inhiben las vías de señalización de citocinas. Las citocinas activan las vías de señalización JAK-STAT de fosforilación al unirse a los receptores transmembranales. Las proteínas SOCS se unen al dominio de fosforilación impidiendo la activación de la vía JAK-STAT. La mayoría de estas proteínas SOCS son inducidas por citocinas y sirven como ciclos de autoinhibición y mecanismos de diferenciación celular. SOCS1 es inducida por STAT6 (IL-4) y por STAT1 (IFN $\gamma$ ), esta proteína se une a IL-4R, IFN $\alpha$ R, IFN $\gamma$ R y MyD88. SOCS1 también se une a IL-2 sobretodo en el timo durante la maduración de las células Th. SOCS3 es inducida por STAT6 (IL-4) y por IL-10, se une a IL-6R e IL-12R. SOCS5 es inducida por IL-12 y se une a IL-4R [48] [Fig: 3.5].



**Figura 3.5:** Las proteínas SOCS regulan la transducción de señales por parte de los receptores de citocinas uniéndose a sus dominios catalíticos e impidiendo que estos activen la vía JAK-STAT (tomada de Yoshimura 2007 [48])



## Epigenética

La expresión preferencial de genes es heredable gracias a diversos mecanismos epigenéticos y de miARN. Los genes de las citocinas *ifn $\gamma$* , *il4* e *il17* tienen marcas epigenéticas permisivas y represivas. *ifn $\gamma$*  e *il4* muestran desmetilación CpG, asociada con expresión génica, durante el proceso de diferenciación en Th1 y Th2. La región río arriba de Foxp3 es metilada en células no Treg pero desmetilada en las Tregs. Los micro ARNs son otro factor que afecta la estabilidad de Foxp3. La delección de las enzimas Dicer y Drosha (implicadas en el procesamiento de miARNs) resulta en la pérdida de expresión por Foxp3 [49].

### 3.3.8. Plasticidad

Las células Th se especializan en distintos subtipos y expresan patrones característicos de citocinas y factores transcripcionales. Tradicionalmente los linajes son Th1, Th2, Th17 y Treg, sin embargo estudios recientes ponen esto en duda. El estudio de las células Th empezó con la caracterización de las células productoras de IFN $\gamma$  (Th1) y las productoras de IL-4 (Th2). Este modelo fue cuestionado con el descubrimiento de la diferenciación de células reguladoras o Tregs. Posteriormente se definió aun otro linaje de células productoras de IL-17 (Th17).

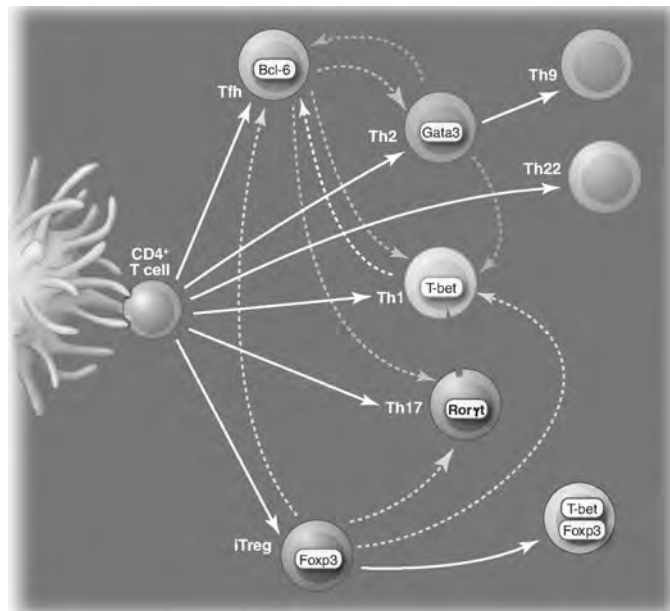
Desde entonces se ha continuado con el descubrimiento de posibles linajes con patrones de citocinas característicos. Ejemplos de esto son las células  $T_{FH}$  implicadas en la producción de anticuerpos, las cuales expresan el represor transcripcional Bcl6 y producen IL-21 [49]. Otro ejemplo son las células Th3, asociadas con la regulación de la respuesta inmune, caracterizadas por producir IL-10. Sin embargo, las células  $T_{FH}$  y Th3 pueden originarse de los demás subtipos (Th1, Th2, Th17 y Treg) y expresar algunas de sus citocinas, lo que lleva a cuestionarse si son verdaderos linajes o estados particulares como las células de memoria [Fig: 3.6].

Además de estos subtipos se ha caracterizado una serie de células Th con patrones de citocinas únicos. Las células Th2 pueden producir IL-9 en presencia de TGF $\beta$ , pero también pueden hacerlo las Th17. Es común que las células Th17 produzcan IL-17 e IFN $\gamma$ , sobretodo *in vivo*. Las células Th17 producen IL-22, pero

células que producen IL-22 y no IL-17 han sido caracterizadas. Incluso se han reportado células productoras de IL-4 en ratón que adquieren la capacidad de producir  $IFN\gamma$  al ser transferidas a otros ratones [49].

La expresión de los factores transcripcionales en las células Th no está fija, lo cual afecta su diferenciación. Al parecer la expresión transitoria de Foxp3 es un efecto común. También se han detectado células que expresan Foxp3 y T-bet o Foxp3 y ROR $\gamma$ t, aunque esto último podría deberse a la acción competitiva entre estos dos factores [45].

La mayoría de los estudios de diferenciación de células Th son realizados *in vitro* y las células son estimuladas durante un largo tiempo. Es cuestionable si los mismos patrones de diferenciación que vemos *in vitro* son reproducibles *in vivo*, dado que las muestras *in vivo* muestran gran heterogeneidad y es difícil separar las células para determinar sus patrones de expresión de citocinas y factores transcripcionales.

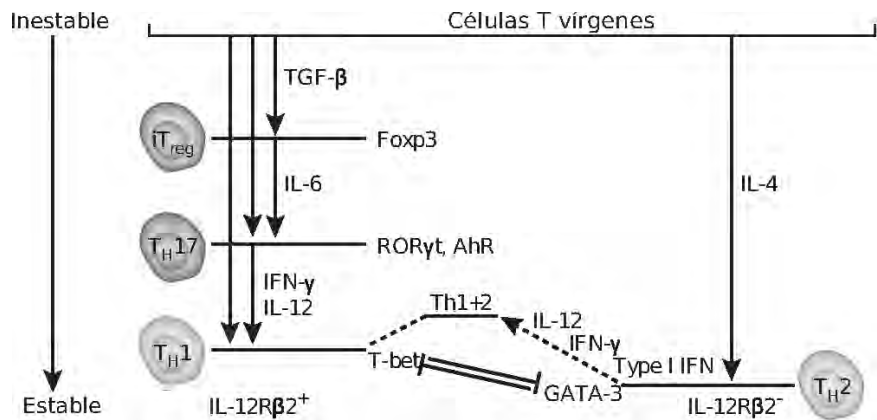


**Figura 3.6:** Las células Th pueden cambiar su patrón de producción de citocinas y en algunos casos su regulador transcripcional maestro (tomado de O'Shea 2010 [49])

Recientemente, se han descrito diferentes formas de plasticidad entre los linajes de células Th [Fig: 3.7]. La transición entre células Treg y Th17 depende del entorno inflamatorio, dado que Foxp3 es inhibido por las citocinas inflamatorias. Es posible polarizar las células Th17 en células Th1 en presencia de IFN $\gamma$  e IL-12. Se puede promover un estado híbrido Th1+2 al exponer a células Th2 a IFN $\gamma$  e IL-12. Sin embargo, no se han observado algunas transiciones, como de Th2 a Th17 o Treg [15]. La estabilidad de las células es mantenida gracias a modificaciones epigenéticas y ciclos de autoinducción.

### 3.4. Sistema inmune y enfermedad

Los sistemas de defensa contra amenazas externas son comunes a todos los organismos. Ejemplos son los "sistemas de restricción-modificación" de las bacterias, las defencinas de los invertebrados, el "sistema de inmunidad adquirida" de las plantas [26] o el sistema inmune de los vertebrados. De ahí se deriva la importancia de que el sistema inmune sea capaz de funcionar de manera adecuada ante los distintos retos inmunológicos. La falla del sistema inmune está relacionada con varias enfermedades



**Figura 3.7:** Transiciones entre los distintos linajes de células Th. Los diferentes tipos celulares pueden cambiar de linaje dependiendo de la estabilidad de su propio linaje y de las señales del entorno. (tomado de Murphy 2010 [15])

autoinmunes como la artritis reumatoide, la diabetes tipo 1, la esclerosis múltiple, el asma, el lupus, la encefalomielitis o la psoriasis. Éstas no son las únicas enfermedades relacionadas con el sistema inmune. El virus de inmunodeficiencia humana causa una falla del sistema inmune que lleva a la muerte. Además, existe evidencia de que la diabetes tipo 2 está relacionada con una desregulación del sistema inmune [4–7, 50].

### 3.4.1. VIH/SIDA

El virus de inmunodeficiencia humano (VIH) es un retrovirus que infecta las células del sistema inmune como Th, macrófagos y dendríticas destruyéndolas o impidiendo su correcto funcionamiento. Conforme la enfermedad avanza el sistema inmune se debilita y la persona se vuelve más susceptible a infecciones. El estado más avanzado de la infección es el síndrome de inmunodeficiencia adquirida (SIDA). Se calcula que 33.3 millones de personas están infectadas por el virus, la mayoría en países del tercer mundo. Cada año 2 millones de personas mueren de VIH/SIDA, la mayor parte a causa de infecciones por tuberculosis [1].

Desde hace mucho tiempo se sabe que el VIH se caracteriza por la desregulación de los linajes de las células Th. Se ha observado una reducción de las células Th1 y un incremento de replicación de células Th0 y Th2 [2, 3]. En el caso del SIV hay una disminución de las células Th17 y una predominancia de Th1 en los intestinos de simios infectados. En seres humanos se ha reportado que la cantidad de células Th17 en mucosas correlaciona negativamente con la viremia en plasma [51]. Las células Treg son elevadas en la fase aguda de la infección, lo cual puede prevenir la eliminación del virus, por otro lado, en la fase crónica, una elevada cantidad de células Treg previene la falla inmune [52].

### 3.4.2. Diabetes tipo 2

La diabetes es una enfermedad en la que la persona tiene altos niveles de azúcar. Esto puede deberse a que el páncreas no produce suficiente insulina (Tipo 1), que las

células no responden a la insulina producida (Tipo 2) o a desordenes durante el embarazo (gestacional). La obesidad y el síndrome metabólico incrementan el riesgo de desarrollar la enfermedad. La diabetes correlaciona fuertemente con enfermedades cardiacas, retinopatía, fallo de los riñones y del hígado y neuropatía.

Actualmente 220 millones de personas viven con diabetes en todo el mundo. En 2005 se estimó que 1.1 millones de personas murieron de diabetes, aunque el número podría ser mayor, ya que las muertes son atribuidas a enfermedades del corazón o fallo de los riñones. El 80 % de las muertes causadas por diabetes ocurren en países del tercer mundo y se espera que la cantidad de muertes se duplique para 2030 [1].

Se ha visto que hay una correlación entre la obesidad y la diabetes, esto podría deberse a la inflamación del tejido adiposo. El tejido adiposo de individuos obesos tiene un contexto inflamatorio y tiene altas cantidades de citocinas como IL-6, IL-1 y  $TGF\beta$ , macrófagos, células T citotóxicas y Th1, mientras que los niveles de Treg han disminuido [50].

Estudios en ratones han mostrado que los ratones delgados tienen una mayor cantidad de Tregs, las cuales desaparecen en ratones obesos [7]. Esto podría deberse a la infiltración de células Th1 en ratones obesos, lo cual disminuye las Tregs y favorece la infiltración por células T citotóxicas y macrófagos [6].

Las células T citóxicas infiltradas en el tejido adiposo favorecen un contexto inflamatorio y el reclutamiento de macrófagos [5]. Curiosamente, la inflamación inicial podría ser causada por los mastocitos, ya que inhibir la activación de estos redujo el incremento de peso y la resistencia a la insulina en ratones [4].

Estos estudios plantean la posibilidad de que la diabetes tipo 2 tenga causas autoinmunes en la constante inflamación del tejido adiposo. Además, estos estudios dan una esperanza terapéutica. El tratamiento de ratones con anticuerpos contra el co-receptor CD3 depletó las células Th1, pero dejó a las células Treg. También se mejoró la tolerancia a la glucosa y la sensibilidad a la insulina.

---

## Capítulo 4

# Redes [53, 54]

La mayoría de los sistemas complejos de la biología, como la regulación genética, el sistema nervioso, el sistema inmune o los ecosistemas, están compuestos de múltiples elementos que interactúan entre sí. Estos sistemas pueden ser modelados con redes. Una red es una colección de elementos conectados. Puede ser visualizada como un conjunto de nodos que corresponden a los elementos de la red y un conjunto de aristas que representan las interacciones entre los elementos. Cada nodo tiene un valor de salida y varios valores de entrada, además de una regla que define el valor de la salida en función de las entradas. Las salidas y las entradas están definidas por las interacciones con otros elementos de la red.

Cuando el valor de los nodos cambia con respecto al tiempo se dice que la red es dinámica. Entonces la función de salida de un nodo puede ser expresada como:  $A_{t+\delta t} = f_A(A_t, B_t, C_t, \dots)$  donde  $A_{t+1}$  será el valor de la salida del nodo,  $A_t, B_t, C_t$  son las entradas y  $f_A$  la función que las relaciona. Cabe destacar que un nodo puede ser entrada de sí mismo.

Los sistemas biológicos pueden ser modelados como redes. Los nodos representan elementos como genes, proteínas, células u organismos. Mientras tanto las interacciones entre ellos corresponden a las aristas. Sin embargo, dada la cantidad de elementos y la complejidad de las interacciones en los sistemas biológicos es difícil

generar una descripción exacta del sistema. Es por esto que se recurre a simplificaciones como suponer tiempos discretos o que los elementos sólo pueden tener un número limitado de valores. Una de estas simplificaciones son las redes booleanas.

En una red genética cada nodo representa un gen o proteína. El nivel de activación del nodo es la cantidad de proteína correspondiente al gen que ha sido sintetizada. En muchos sistemas biológicos el nivel de expresión de un gen se mantiene en un nivel basal constitutivamente. Cuando los factores activadores sobrepasan un umbral, entonces la cantidad de la proteína correspondiente se incrementa rápidamente hasta alcanzar un máximo. Esto permite suponer que la expresión del gen puede ser expresada en base a funciones booleanas, con el gen inactivo (expresión basal) o activo (incremento de la función).

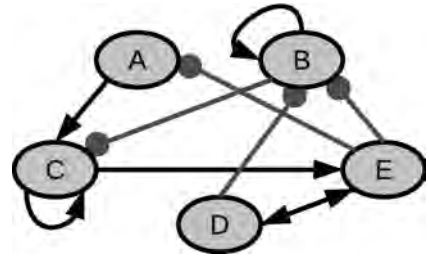
Las aristas de una red genética representan la forma en que otros genes ayudan a la activación o represión de la transcripción del gen. Una arista de activación que une a dos nodos implica que uno de los nodos es un factor transcripcional del otro, fomentando la síntesis de la proteína correspondiente. Por otro lado, si la arista representa una inhibición entonces uno de los genes es un inhibidor de otro, evitando que el gen blanco sea sintetizado.

En 1969 Kauffman propuso a las redes booleanas como modelos de regulación genética [55]. Partió de la idea de que un gen puede ser expresado (1) o no expresado (0). Dado que la expresión de cada gen es controlada por la expresión de otros genes Kauffman supuso que el genoma es una red donde una arista del gen A al gen B representa que A regula la expresión de B ya sea activándolo o inhibiéndolo. Dado que no existían suficientes datos experimentales Kauffman generó redes aleatorias. Sin embargo, su generalidad ha motivado interés más allá de su propósito original [56, 57].

Las redes booleanas son un tipo de red dinámica discreta, finita y determinista. Una red booleana consiste de  $N$  nodos unidos por  $K$  conexiones cada uno. Cada nodo tiene  $k$  aristas de entrada y  $l$  aristas de salida. Los nodos son booleanos, es decir su estado es "prendido" (1) o "apagado" (0). El estado de cada nodo se actualiza en un tiempo discreto, el estado en el tiempo  $t + 1$  depende de los estados de sus  $K$  entradas en el tiempo  $t$  a través de una función booleana [Fig: 4.1]. Las redes

aleatorias han sido útiles en el estudio de las características de las redes booleanas y en determinar el efecto de la elección de funciones, características de los atractores y las fases de la dinámica de la red.

Las redes booleanas aleatorias de Kauffman siguen un modelo N-K (donde  $N$  es el número de nodos de la red y  $K$  la conectividad). Originalmente Kauffman usó una  $K$  homogénea, de tal forma que cada nodo de la red tiene exactamente  $K$  conexiones de entrada. Otros autores han usado diferentes distribuciones de probabilidad para el número de conexiones  $K$  de cada nodo, como las distribuciones libres de escala.



**Figura 4.1:** Ejemplo de red booleana. En esta red la  $K$  de cada nodo es dos.

En una topología libre de escala, la probabilidad  $P(k)$  de que un nodo arbitrario de la red esté conectado con exactamente otros  $k$  nodos es:

$$P(k) = Ck^{-\gamma}$$

donde  $\gamma$  es un exponente libre de escala. En las redes libres de escala una pequeña fracción de los nodos está altamente conectada, mientras que la mayoría de los nodos están poco conectados.

Según Kauffman, los estados estacionarios de las redes de regulación genética pueden representar diferentes tipos celulares dependiendo del estado de los nodos o genes. Dada una configuración inicial y las relaciones entre los genes, se puede dejar evolucionar el sistema y encontrar que los elementos de la red tienen estados de activación cíclicos o de punto fijo, que corresponden a los atractores. Estos estados estacionarios o atractores, no son estados intrínsecos de cada elemento, sino un fenómeno emergente de la complejidad del sistema.

En base a los datos experimentales disponibles se han podido inferir redes de regulación genética de varios organismos como *E. coli*, *S. cerevisiae*, *B. subtilis* [58], la morfogénesis de las flores de *A. thaliana* [59], los genes de polaridad de segmentación de *D. melanogaster* [60] y la diferenciación de células Th [12–14].



## 4.1. Funciones booleanas

La salida de un nodo está dada por una función booleana que toma como argumentos sus entradas. Si un nodo tiene  $E$  entradas estas pueden estar en  $2^E$  estados diferentes. Cada regla le asigna una sola salida a cada posible estado de las entradas. Sin embargo, todas estas funciones se pueden expresar en base a tres funciones booleanas: *and*, *or* y *not* [Cuadro: 4.1 y 4.2].

$t$		$t + 1$	
A	B	A <i>and</i> B	A <i>or</i> B
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	1

**Cuadro 4.1:** Funciones and y or

$t$	$t + 1$
A	<i>not</i> A
0	1
1	0

**Cuadro 4.2:** Función not

No siempre se conocen las funciones booleanas de una red. Cuando es así se pueden inferir las funciones identificando los elementos del sistema, sus relaciones y tipo de interacciones. La formulación y análisis de redes permite encontrar nuevo conocimiento y la posibilidad de hacer predicciones. Un grafo puede tener varias funciones booleanas diferentes; al inferir sus reglas es necesario buscar aquellas que resultan más parsimoniosas y consistentes con las observaciones experimentales.

En el caso de los sistemas biológicos existen varias formas de inferir las funciones de una red de regulación. Se puede utilizar la información ya conocida de la expresión de genes y proteínas, incluyendo la regulación transcripcional, postranscripcional, traduccional y postraduccional. También se pueden realizar inferencias de la función de una proteína analizando las diferencias en las respuestas a diversos estímulos en organismos silvestres y mutantes. Una vez propuesta la relación de regulación entre dos genes, ésta puede ser confirmada experimentalmente.

Otro tipo de redes booleanas son las redes booleanas aleatorias (RBN). Constan de  $N$  nodos con  $K$  conexiones en promedio. Las conexiones y las funciones son elegidas aleatoriamente cuando la red es generada. En la red cada gen es expresado (valor 1) con probabilidad  $p$  y no expresado (valor 0) con probabilidad  $p - 1$ .

### 4.1.1. Funciones sigmoidales

Las funciones booleanas se caracterizan porque el estado de un nodo es inactivo o activo, hasta que sus entradas alcanzan un umbral donde el nodo cambia de estado. La función escalón [Fig: 4.2] es una función discreta, tal que:

$$f(x) = \begin{cases} 0 & \text{si } x < \alpha \\ 1 & \text{si } x \geq \alpha \end{cases}$$

donde  $\alpha$  es el umbral de la función.

Un modelo en redes booleanas puede ser transformado a un modelo continuo utilizando funciones sigmoidales. Una función con comportamiento sigmoidal se caracteriza por comenzar en un valor e incrementar suavemente a otro valor cuando  $x$  tiende infinito, con un solo punto de inflexión. Hay varias funciones sigmoidales como la función de Hill y la logística.

La función de Hill [Fig: 4.4] se utiliza para describir las asociaciones cooperativas de moléculas y proteínas. La función es:

$$f(x) = \frac{x^n}{\theta^n + x^n}$$

donde  $\theta$  es la constante de disociación de las moléculas y la  $n$  es el coeficiente de Hill o cooperatividad.

La logística [Fig: 4.3] modela crecimientos con un valor máximo fijo. Su solución esta dada por:

$$f(x) = \frac{k \cdot x_0 \cdot e^{rt}}{k + x_0(e^{rt} - 1)}$$

donde  $r$  es la constante de crecimiento y  $k$  determina la capacidad de carga del sistema.

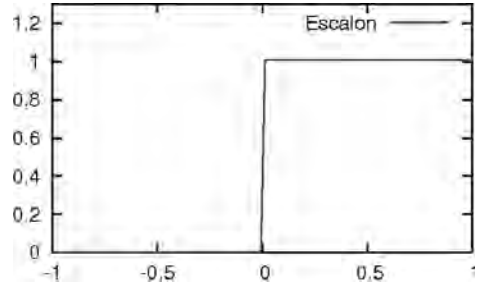


Figura 4.2: Función escalón,  $\alpha = 0$ .

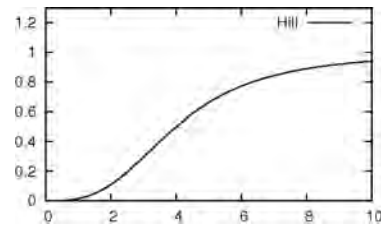


Figura 4.3: Función de Hill,  $n = 3$  y  $\theta = 4$ .

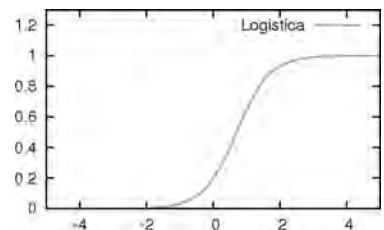


Figura 4.4: Función logística,  $b = 4$  y  $k = 2$

## 4.2. Elementos de las redes

Una red es un conjunto de elementos conectados. Sin embargo, la forma en la que la red se conecta y el tipo de conexiones afectan a la red. Una red puede ser clasificada según su tipo de aristas, ya sean dirigidas o no dirigidas. Las trayectorias formadas por las aristas también generan estructuras y patrones dentro de la red que determinan su dinámica.

### 4.2.1. Dirección y conectividad

Una red puede ser dirigida o no dirigida. En una red no dirigida las aristas indican una interacción entre nodos, mientras que en una red dirigida existe dirección en la comunicación de los nodos. Las redes no dirigidas se utilizan cuando la fuente y el objetivo no están bien identificados o cuando la relación es simétrica. Estas aristas sirven para expresar, por ejemplo, relaciones de unión proteína-proteína. Las redes dirigidas se utilizan cuando un nodo tienen un efecto directo y direccional sobre otro. Estas aristas sirven para expresar, por ejemplo, activación o inhibición de genes.

Dos nodos en una red no dirigida están conectados si existe un camino de interacciones entre los dos. Si sólo se necesita transitar una arista entre uno y otro los nodos son adyacentes. Cuando todos los nodos de la red están conectados se dice que es conexa. Si la red es dirigida se dice que es débilmente conexa si al convertir sus aristas a aristas no dirigidas la red es conexa. La red es fuertemente conexa si existe un camino dirigido entre todos los nodos de la red.

La matriz adyacente  $A_{N \times N}$  resume la estructura de una red con  $N$  nodos, pues indica si cada par de nodos  $x_i$  y  $x_j$  están conectados. Si la matriz es simétrica la red es no dirigida. Las componentes de la matriz están dadas por:

$$A_{ij} = \begin{cases} 1 & \text{si } (i, j) \in U \\ 0 & \text{si } (i, j) \notin U \end{cases}$$

donde  $U$  es el conjunto de las conexiones de la matriz. Si la conexión entre  $x_i$  y  $x_j$  es ponderada el valor de la componente será la ponderación. Para inhibición es -1.

El número total de nodos de la red define su orden. El número de aristas que convergen a un nodo es el grado de entrada y el número de aristas que salen el de salida. El grado de entrada más el grado de salida, da el grado de cada nodo. La densidad de una red es la proporción entre el número de aristas  $k$  de la red y el número de aristas posibles entre los  $N$  nodos:

$$\rho = \frac{k}{N(N-1)}$$

### 4.2.2. Trayectorias

Una trayectoria es la secuencia de aristas adyacentes, es decir, que se puede ir de una a otra en un solo paso. Se dice que una conexión es básica si al removerla se interrumpen todas las trayectorias entre dos nodos. La distancia de la trayectoria entre dos nodos es el número de aristas en uno de los caminos que los conecta, la distancia más corta se conoce como distancia geodésica. La matriz de distancias geodésicas  $G_{ij}$  tiene como entradas el número de aristas que componen la distancia geodésica entre dos nodos  $x_i$  y  $x_j$ .

Una trayectoria puede representar rutas metabólicas o de transducción de señales. El promedio de la longitud de trayectoria para múltiples redes metabólicas, transcripcionales, de proteínas y de transducción, es menor a cuatro [61], por lo que estas redes responden rápidamente a perturbaciones. De igual manera, estas redes suelen tener redundancia de trayectorias y múltiples rutas entre nodos, lo cual las hace robustas ante perturbaciones, permitiéndoles recuperar la trayectoria preferente.

### 4.2.3. Centralidad

La importancia relativa de un nodo dentro de la red está dada por su centralidad, que considera el grado de entrada y salida, así como la conectividad. La centralidad de grado  $C_D$  de la red, da la fracción de nodos de grado  $k$  en una red. Para una gráfica con  $N$  nodos, la centralidad de grado para un nodo de grado  $k$ , es:

$$C_D(k) = \frac{k}{N-1}$$

La centralidad de intermediación mide la frecuencia con que un nodo aparece en la trayectoria geodésica (la de menor distancia) entre dos nodos, y es grande para los nodos que funcionan como intermediarios y controladores de la comunicación. Para una red  $G(X, U)$  con  $N$  nodos, la centralidad de intermediación de un nodo  $x_v$ , es:

$$C_B(k) = \frac{1}{B} \sum_{\substack{s \neq v \neq t \in v \\ s \neq t}}^N \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Donde  $\sigma_{st}$  es el número de trayectorias geodésicas que unen el nodo  $x_s$  con el nodo  $x_t$ ,  $\sigma_{st}(v)$  es el número de trayectorias geodésicas que unen  $x_s$  con  $x_t$  y que pasan por  $x_v$ .  $B$  es una cantidad que normaliza respecto al número de pares de vértices que no incluyen a  $x_v$ , es decir  $B = (N - 1)(N - 2)$  para gráficas dirigidas y  $B = \frac{(N-1)(N-2)}{2}$  para gráficas no dirigidas.

#### 4.2.4. Estructuras y motivos

Las redes pueden tener estructura de cadena, de ciclos cerrados o combinaciones de estas según las conexiones de sus aristas. En las cadenas un estado A se conecta a un estado B, éste a un estado C y así sucesivamente. Dado que el nodo inicial no tiene una entrada se supone que su valor es constante. Esto hace que el valor de todos los nodos dependa del valor del nodo inicial. En un ciclo todos los nodos están conectados sucesivamente y la salida del último es la entrada del primero. Este tipo de estructuras puede presentar ciclos, donde los nodos regresan a su estado inicial. La mayoría de las redes tienen combinaciones de estas dos estructuras.

Además de grandes cúmulos de nodos conectados de forma bidireccional, en la red hay motivos de interacción, que son pequeñas subgráficas con topología bien definida, que forman bloques de redes celulares o módulos, robustos y separables o autónomos. Estos motivos son especialmente prevalentes en redes biológicas, donde su cantidad es mayor a la que se encuentra en redes aleatorias [62].

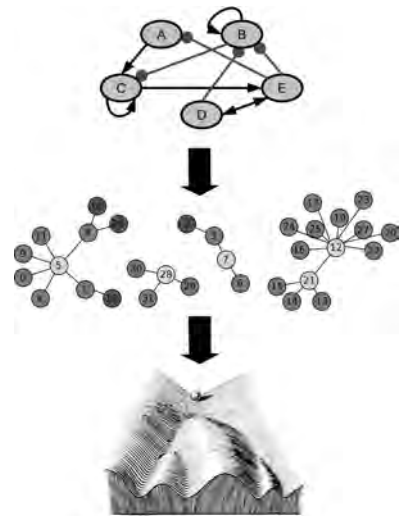
### 4.3. Red de estados

Las redes booleanas son deterministas y finitas. Una red de  $N$  nodos tiene  $2^N$  estados posibles. Cada estado corresponde a un nodo de la red de estados de la red original. Las aristas de esta red se construyen tomando uno de los estados, evaluándolo de acuerdo a las funciones de la red y viendo cual es el nuevo estado obtenido. De esta forma, en una red booleana determinista, cada estado tiene una arista de salida que lleva a su sucesor.

Los valores de los nodos cuando  $t = 0$  se conocen como condición inicial. Cuando un estado es revisitado la red ha alcanzado una solución o atractor. El número de estados que tarda en reencontrarse un estado específico determina el periodo de la solución. Las soluciones puntuales tienen un periodo de uno (un solo estado), mientras que los cíclicos tienen periodos más grandes que uno (estados múltiples). Una red booleana puede tener uno o más atractores.

El conjunto de estados visitados hasta que un atractor es alcanzado se llama transitorio. El conjunto de estados llevando hacia un atractor forma su cuenca de atracción. La cuenca de los diferentes atractores divide el espacio de estados [Fig: 4.5]. Las redes booleanas son disipativas, un estado puede tener múltiples predecesores, pero sólo un sucesor, dado que su transición es determinista. Los estados sin un predecesor se llaman cúspides, dado que sólo pueden ser alcanzados si son la condición inicial y en la siguiente iteración se pierden.

Es importante notar la diferencia entre la topología de una red booleana y su red de estados. La topología de una red booleana representa como los  $N$  nodos de la red se afectan unos a los otros. En la red de estados cada nodo representa un estado



**Figura 4.5:** Ejemplo de la red de estados de una red booleana. Los nodos claros representan las soluciones de la red, dos son soluciones puntuales y otros dos un ciclo. La red de estados corresponde al paisaje epigenético del sistema.

de la red booleana, donde hay  $2^N$  estados, y la red representa las transiciones de todo el espacio de estados. Uno de los principales temas en el estudio de las redes booleanas es entender como cambios en la topología de la red (escala baja) afectan la red de estados (escala alta), lo cual no es trivial.

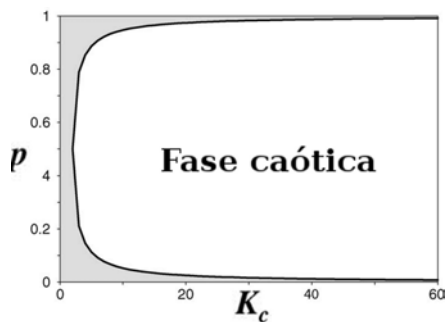
En los sistemas biológicos, la red booleana representa las interacciones entre los genes y proteínas del sistema. En contraste, la red de estados representa las posibles combinaciones de activación de genes y proteínas. Se supone que esta red es equivalente al paisaje epigenético del sistema, donde cada estado estacionario corresponde a un estado de diferenciación de la célula.

#### 4.4. Fases dinámicas

Las redes booleanas pueden tener diferentes comportamientos dinámicos. Existen dos fases dinámicas: caótica y ordenada [Fig: 4.6]. La transición de fase es caracterizada por su criticalidad y se conoce como el límite del caos.

En la fase ordenada la mayoría de los nodos son estáticos, es decir, no cambian su estado. En promedio los estados tienen muchos predecesores y estados similares tienden a los mismos atractores. Esto lleva a una alta convergencia, poco tiempo de transitivo y una alta densidad de cúspides. En esta fase la red tiende a ser robusta ante la eliminación de nodos.

En la fase caótica la mayoría de los nodos cambian de estado constantemente. En esta fase las redes son frágiles ya que el daño se expande a través de la red. Los estados similares tienden a diferentes atractores. En promedio los estados tienen pocos predecesores lo que lleva a una baja convergencia, largo tiempo de transitivo y una baja densidad de cúspides.



**Figura 4.6:** Gráfica de fases dinámicas de las redes booleanas en función de  $p$  contra  $K_c$ . En gris se encuentra la fase estática y en blanco la caótica (tomada de Aldana 2003 [57])

El régimen crítico se encuentra cerca de la transición entre las fases caótica y ordenada. En esta fase algunos nodos cambian y otros son estáticos, por lo tanto el daño es transmitido pero no necesariamente a toda la red. Estados similares tienden a encontrarse en trayectorias que no convergen o divergen en el espacio fase [63]. La cantidad de predecesores de un nodo se aproxima a una ley de potencias de la forma  $y = ax^k$  y la convergencia es media. Recientemente se ha visto que el régimen crítico maximiza el almacenamiento y transmisión de información [64], además de la información de Fisher [65].

Uno de los factores que determina la dinámica de una red booleana es la probabilidad  $p$  de que un gen se encuentre activo dependiendo de las características de los  $K$  genes que lo controlan. Si  $p = 1$  todas las posibles salidas de las funciones booleanas son 1, si  $p = 0$  todas las salidas son 0. Cuando  $p = 0,5$  existe la mayor variabilidad posible en las funciones booleanas.

La transición de fase está dada analíticamente por los parámetros  $K$  y  $p$ . La sensibilidad promedio de la red  $S$  se puede definir como:

$$S = 2p(1 - p)K$$

Si  $S > 1$  la red se encuentra en el régimen caótico, si  $S < 1$  en el régimen ordenado. La transición de fase ocurre en  $S = 1$  lo cual es conocido como el régimen crítico.

Para cada  $p$  existe un valor crítico de conectividad  $K_c$  cuando  $S = 1$  dado por:

$$K_c(p) = \frac{1}{2p(1 - p)}$$

Si  $K < K_c$  el sistema se encuentra en la fase ordenada, si  $K > K_c$  el sistema se encuentra en fase caótica. Cuando  $K = K_c$  el sistema se encuentra en fase crítica, es entonces que la red tiene las propiedades necesarias para simular las características de las redes biológicas como las redes de regulación genética o ecológicas.

Se ha dicho que la computación y la vida yacen necesariamente en el límite del caos [63]. Existe evidencia experimental de que las redes genéticas de organismos de al menos cuatro reinos están cerca del régimen crítico [58]. Esta tendencia hacia la criticalidad puede deberse a que aporta suficiente robustez para mantener la información y suficiente variación para permitir la evolución de los seres vivos.



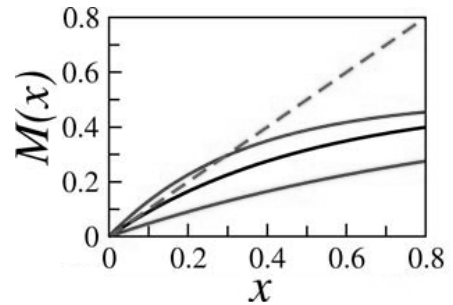
### 4.4.1. Efectos de las perturbaciones

Una manera de verificar la robustez de una red es caracterizar el tamaño de la avalancha de cambios en la trayectoria dinámica de un estado en respuesta a las perturbaciones utilizando un mapa de Derrida. La distancia de Hamming  $x(t)$  mide el número de nodos diferentes entre dos estados. Desde un punto de vista biológico, la distancia de Hamming es el promedio normalizado en el tiempo  $t$  de la avalancha de perturbaciones en el perfil de expresión genético producida por la perturbación de una fracción  $x(0)$  de genes en el tiempo  $t$ .

El factor  $M(x)$  relaciona el tamaño  $x(t)$  de la avalancha de perturbaciones en el tiempo  $t$  con respecto a la avalancha de perturbaciones en el tiempo  $t + 1$  de tal forma que  $x(t + 1) = M(x(t))$ . Dada una perturbación inicial  $x(0)$  en  $t = 0$ , las iteraciones sucesivas de  $M(x)$  convergerán a un valor  $x(\infty)$ , el cual es el tamaño final de la perturbación.

En los sistemas biológicos esto equivale a tener un patrón de expresión génica que eventualmente llegará a una solución dada con una expresión de genes específica. Al alterarse la expresión de algunos genes se llegará a una nueva solución eventualmente. La distancia  $x(\infty)$  mide el porcentaje de genes distintos entre las dos soluciones.

La transición dinámica entre la fase ordenada y la fase caótica puede ser caracterizada por la evolución temporal de la distancia de Hamming  $x(t)$  entre dos trayectorias dinámicas producidas por dos condiciones iniciales ligeramente diferentes [Fig: 4.7]. Si  $x(\infty) = 0$  entonces la red está en la fase ordenada y eventualmente todas las perturbaciones desaparecerán. Si  $x(\infty) > 0$  la red se encuentra en fase caótica y eventualmente una fracción  $x(\infty)$  de los genes de la red serán alterados en respuesta a las perturbaciones.



**Figura 4.7:** Mapa de Derrida utilizando tres redes aleatorias en diferentes fases dinámicas: ordenada (abajo), crítica (enmedio) y caótica (arriba). Tomada de Balleza 2008 [58]

---

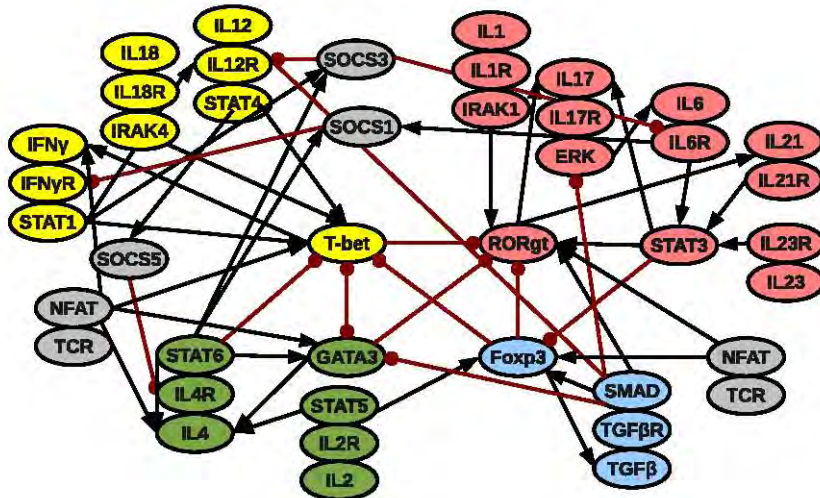
## Capítulo 5

# Diseño de la red

En la construcción de la red booleana de diferenciación de células Th se consideró que el estado de un nodo  $X$  en el tiempo  $t$  es una función booleana de sus  $N$  reguladores en el tiempo anterior tal que:  $X_{t+1} = f(A_t, C_t, \dots, N_t)$  Si el nodo depende de varios reguladores diferentes se les relacionó usando **or**, excepto en casos reportados de sinergia donde se utilizó **and**. La inhibición se representó usando **and not**. Un nodo está activo (1) cuando la proteína o factor transcripcional que representa existe en cantidad suficiente para activar o reprimir otros nodos. Si no existe en cantidad suficiente el nodo está inactivo (0).

En base a la revisión de la literatura sobre la regulación de células Th [9, 15, 33–40, 44–49], se trazó una red de regulación para el sistema, donde las interacciones de activación se indican con flechas y las de inhibición con barras [Fig: 5.1]. A partir de ésta y tomando en cuenta lo antes expuesto se construyeron las funciones booleanas de la red [Cuadro: 5.1]. Esta red contiene 43 nodos y 84 interacciones entre ellos.

Al construir la red se supuso que se encuentran presentes todas las subunidades de los receptores de citocinas y los factores transcripcionales, aunque su conformación no esté activa. Además, no existen limitaciones en el transporte, plegamiento y dimerización y activación de las proteínas. A continuación se explica en detalle como se construyeron algunas de las tablas de verdad.



**Figura 5.1:** Red de regulación genética de células Th. Las interacciones de activación se indican con flechas negras y las de inhibición con barras rojas. Nodos contiguos representan vías de señalización. Los nodos correspondientes a Th1 son amarillos, Th2 verdes, Th17 rojos y Treg azules, las proteínas SOCS, el TCR y NFAT se encuentran en gris.

Gen ( $t + 1$ )	Función ( $t$ )
TBET	((STAT4 and IRAK4) or STAT1 or (TBET)) and NFAT and not (STAT6 or GATA3 or FOXP3)
IL12	IL12
IL12R	IL12 and not (SMADlo or SMADhi or (SOCS3 and not STAT1))
STAT4	IL12R
IL18	IL18
IL18R	IL18
IRAK4	IL18R
IFN $\gamma$	TBET and NFAT
IFN $\gamma$ R	IFN $\gamma$ and not SOCS1
STAT1	IFN $\gamma$ R
GATA3	(STAT6 or GATA3) and NFAT and not (TBET or SMADlo or SMADhi)
IL4	((STAT5 and GATA3) or STAT6) and NFAT
IL4R	IL4 and not SOCS5
STAT6	IL4R

Gen ( $t + 1$ )	Función ( $t$ )
IL2 IL2R STAT5	IL2 IL2 IL2R
ROR $\gamma$ t IL1 IL1R IRAK1 IL17 IL17R ERK IL6 IL6R IL21 IL21R IL23 IL23R STAT3	(STAT3 or IRAK1) and SMADlo and NFAT and not (FOXP3 or TBET or GATA3 or SMADhi) IL1 IL1 IL1R ROR $\gamma$ t and STAT3 IL17 and not SMADhi IL17R IL6 or ERK IL6 and not SOCS3 ROR $\gamma$ t IL21 IL21 IL23 IL23 IL6R or IL21R or IL23R
FOXP3 TGFBlo TGFBhi TGFBloR TGFBhiR SMADlo SMADhi	SMADhi and STAT5 and NFAT and not STAT3 SMADlo and not SMADhi FOXP3 TGFBlo TGFBhi TGFBloR TGFBhiR
SOCS1 SOCS3 SOCS5 TCR NFAT	STAT6 or STAT3 STAT6 or STAT1 STAT4 TCR TCR

**Cuadro 5.1:** Red de regulación genética

## 5.1. Reglas lógicas y tablas de verdad

### 5.1.1. Regulación epigenética

Los genes no están sólo regulados por factores transcripcionales, sino también por factores epigenéticos. Un gen sólo puede ser activado (1), si su cromatina se encuentra abierta de tal forma que la polimerasa pueda transcribir el gen. El estado de la cromatina puede ser abierto o cerrado por reguladores epigenéticos.

Este es el caso de proteínas como  $\text{IFN}\gamma$  e IL-4. Los reguladores transcripcionales T-bet y GATA-3 no activan directamente a  $\text{IFN}\gamma$  e IL-4, sino que abren su cromatina permitiendo la entrada de NFAT,  $\text{NF}\kappa\text{B}$  o STAT5 que reclutan la polimerasa, permitiendo la expresión de las citocinas. Una vez que la cromatina ha sido modificada permanece en el nuevo estado durante varias generaciones. Además, GATA-3 es necesario para la diferenciación de las células T CD4+ [66] por lo que se encuentra en un nivel basal suficiente para fomentar la diferenciación en células Th, pero no en células Th2. Se consideró que este nivel basal es (0).

Para que la citocina sea expresada es necesario que la cromatina esté abierta y que exista un factor transcripcional que active la expresión. La tabla de verdad [Cuadro: 5.2] corresponde a la función **and**.

$t$		$t + 1$
Remodelador cromatina	Factor transcripcional	Citocina
0	0	0
0	1	0
1	0	0
1	1	1

**Cuadro 5.2:** Tabla de verdad de regulación epigenética.

### 5.1.2. Regulación por proteínas SOCS

Las proteínas SOCS compiten con los factores transcripcionales STAT por el sitio de unión al receptor de las citocinas, impidiendo la fosforilación e inhibiendo la transducción de señales. La inhibición de las proteínas SOCS depende de que haya suficientes para evitar bloquear una parte importante de los receptores. Sin embargo, esto puede ser evitado si algún otro factor incrementa la cantidad de receptores, permitiendo la señalización.

En la señalización por IL-12, SOCS3 se une al receptor de IL-12R, impidiendo que STAT4 sea fosforilado.  $\text{IFN}\gamma$  fomenta la expresión de IL-12R, superando la inhibición de SOCS3. Esto puede ser expresado como una tabla de verdad [Cuadro: 5.3], obteniéndose la expresión:  $\text{STAT}_{t+1} = \text{Citocina}_t \text{ and not } (\text{SOCS}_t \text{ and not Activador}_t) = \text{Citocina}_t \text{ and } (\text{not SOCS}_t \text{ or Citocina}_t)$

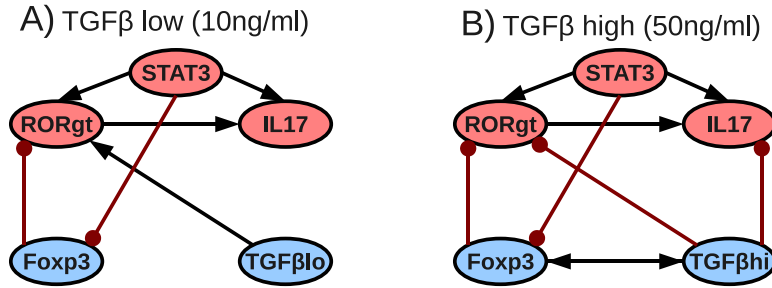
$t$			$t + 1$
IL-12	SOCS3	$\text{IFN}\gamma$	STAT4
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

**Cuadro 5.3:** Tabla de verdad de transducción de señales.

### 5.1.3. El caso $\text{TGF}\beta$

Originalmente se diseñó la red con un solo nodo de  $\text{TGF}\beta$ . Esta red no lograba expresar correctamente la regulación de las células Th, principalmente en el caso de  $\text{ROR}\gamma\text{t}$ . Además, al analizar la red se encontró que no era posible recuperar los atractores de Th17. Esto se debe a que el circuito de regulación de  $\text{ROR}\gamma\text{t}$ , IL-17, STAT3, Foxp3 y  $\text{TGF}\beta$  es diferente si el nivel de  $\text{TGF}\beta$  es bajo (10 ng/ml) a si

el nivel de  $TGF\beta$  es alto (50 ng/ml) [46]. De esta forma, se pueden construir dos circuitos diferentes para  $TGF\beta$ hi y  $TGF\beta$ lo [Fig 5.2]. Tomando en cuenta estos circuitos se construyó la red [Cuadro: 5.1].



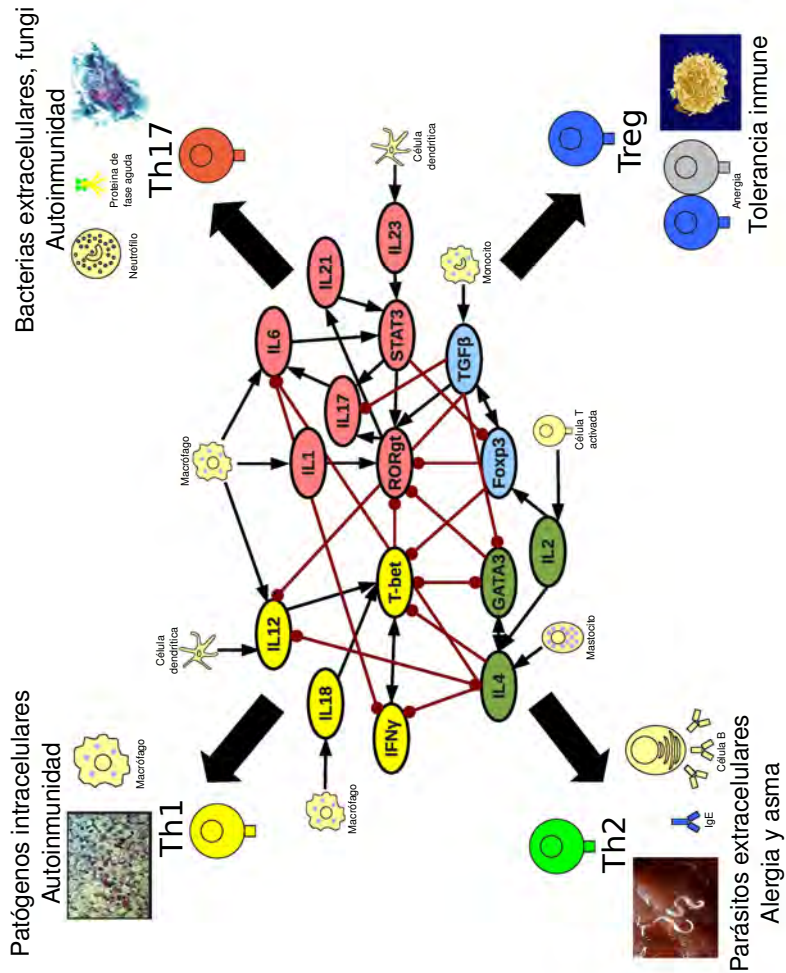
**Figura 5.2:** Regulación diferencial por  $TGF\beta$ . A) Si el nivel de  $TGF\beta$  es bajo este fomenta la expresión de ROR $\gamma$ t. B) Al incrementar  $TGF\beta$  este activa a Foxp3 y empieza a inhibir a ROR $\gamma$ t e IL-17.

## 5.2. Simplificación de la red

Para facilitar el análisis de la red se simplificó usando reglas booleanas. Si un nodo depende únicamente del valor de otro, los dos pueden ser reducidos a uno solo, simplificando la red. También se supuso que el TCR está activado (1) y por lo tanto se encuentran presentes NFAT y  $NF\kappa B$ . La red resultante consta de 17 nodos y 51 interacciones, teniendo  $2^{17}$  estados posibles [Cuadro: 5.4 y Fig: 5.3].

## 5.3. Análisis de la red

Con el fin de analizar la red se construyó un programa que itera síncronamente los 17 nodos de la red reducida hasta encontrar una solución o llegar a un número máximo de iteraciones. A partir de los resultados obtenidos se determinaron las soluciones puntuales, la cantidad y tamaño de los ciclos, los estados transitivos, las cúspides y el tamaño de la cuenca de atracción correspondiente a cada solución.



**Figura 5.3:** Red booleana de diferenciación de células Th. Los nodos correspondientes a Th1 son de color amarillo, los de Th2 verdes, los de Th17 rojos y los de Treg azules. Se indican las células productoras de citocinas.



Gen (t+1)	Función (t)
TBET	((IL12 AND IL18) or IFNg or TBET) and not (IL4 or GATA3 or FOXP3)
IL12	IL12 and not (TGFBlo or TGFBhi or (IL4 and not IFNg))
IL18	IL18
IFNg	TBET and not (IL4 or IL6)
GATA3	(IL4 or GATA3) and not (TBET or TGFBlo or TGFBhi)
IL4	((IL2 and GATA3) or IL4) and not (IL12)
IL2	IL2
ROR $\gamma$ t	(STAT3 or IL1) and TGFBlo and not (FOXP3 or TBET or GATA3 or TGFBhi)
IL1	IL1
IL6	(IL6 or IL17) and not (IL4 or IFNg)
IL17	(ROR $\gamma$ t and STAT3) and not TGFBhi
IL21	ROR $\gamma$ t
IL23	IL23
STAT3	IL6 or IL21 or IL23
FOXP3	TGFBhi and IL2 and not STAT3
TGFBlo	TGFBlo and not TGFBhi
TGFBhi	FOXP3

Cuadro 5.4: Red reducida de regulación genética

Posteriormente se analizó el efecto del ruido en la red. Para esto se tomaron estados de la red y se cambió al azar de uno a ocho nodos y se vio cual es la nueva solución. A partir de esto se calculó la distancia de Hamming y el tamaño de la avalancha de perturbaciones, viendo la respuesta de la red al ruido y como afecta los distintos linajes. Para el cálculo de las matrices de adyacencia y geodésica, además de la centralidad de los nodos y la modularidad se utilizó la librería Networkx [67].

### 5.3.1. Criterios de clasificación

Se definió que un estado estacionario corresponde a un linaje si expresaba su factor transcripcional maestro y sus citocinas características [Cuadro: 6.3]. En el caso de Th0 se consideró que una célula era Th0 si no expresa ninguno de los factores transcripcionales maestros.

## Th0

Se consideró que una célula es Th0 si no expresa ninguno de los factor transcripcionales maestros como T-bet, GATA-3, ROR $\gamma$ t o Foxp3. Muchas veces, se clasifica una célula como Th0 si no expresa ninguna citocina [12], sin embargo también se han reportado células Th0 con expresión de IFN $\gamma$ , IL-4, IL-5, IL-6, IL-10, IL-17 y GM-CSF y TGF $\beta$  [34]. Al analizar la red se encontró la solución sin expresión de ninguna citocina, pero también se encontraron soluciones con expresión de IL-6, TGF $\beta$ lo e IL-4 [Cuadro: 6.3]. Sin embargo, no se encontraron soluciones que expresen IFN $\gamma$  o IL-17, que también han sido reportados. Esto podría deberse a que la regulación de las proteínas SOCS cambia según el estado de activación de las células [35, 36].

## Th1

Se consideró que una célula es Th1 si expresa T-bet e IFN $\gamma$ . Ignorando las citocinas exógenas se encontraron 48 soluciones que expresen T-bet e IFN $\gamma$ . De estas, 8 expresan todas las citocinas vinculadas con la diferenciación en Th1: IFN $\gamma$ , IL-12 e IL-18 [Cuadro: 6.3]. Curiosamente, también se encontraron 8 soluciones que expresan T-bet, IL-12 e IL-18, pero no IFN $\gamma$ , todas estas soluciones tienen STAT3 activo. Además de 40 soluciones que solo expresan T-bet, las cuales no es posible clasificar. En la literatura se han reportado células T-bet+IFN $\gamma$ -, generalmente relacionándolas con niveles de expresión insuficientes de T-bet o inhibición por vías alternas [68].

## Th2

Se consideró que una célula es Th2 si expresa GATA-3 e IL-4. Ignorando las citocinas exógenas se encontraron 16 soluciones que expresen GATA-3 e IL-4. De estas 8 expresan todas las citocinas vinculadas con la diferenciación en Th2: IL-4 e IL-2. Se supuso que la expresión de GATA-3 implica la presencia de IL-5 e IL-13 [Cuadro:6.3]. También se encontraron 48 soluciones que expresan GATA-3 (y por lo tanto IL-5 e IL-13) pero no IL-4. Se ha reportado en la literatura que una vez que el

fenotipo Th2 ha sido establecido IL-4 no es necesario [39]. Además, se encontraron 24 soluciones que expresan IL-4 pero no GATA-3. Estas podrían ser clasificadas como Th0, aunque estudios de knock-out condicional de GATA-3 parecen indicar que una vez que el fenotipo ha sido establecido es posible mantener las citocinas características de Th2 sin necesidad de GATA-3 [38].

### **Th17**

Se consideró que una célula es Th17 si expresa  $ROR\gamma_t$  e IL-17. Ignorando las citocinas exógenas se encontraron 32 soluciones que expresen  $ROR\gamma_t$  e IL-17. De estas sólo cuatro expresan todas las citocinas vinculadas con la diferenciación en Th17:  $ROR\gamma_t$ , IL-1, IL-6, IL-17, IL-21, IL-23 y  $TGF\beta_{lo}$  [Cuadro: 6.3].

### **Treg**

Se consideró que una célula es Treg si expresa Foxp3 y  $TGF\beta_{hi}$ . Ignorando las citocinas exógenas se encontraron ocho soluciones que expresen Foxp3,  $TGF\beta_{hi}$  [Cuadro: 6.3].

---

## Capítulo 6

# Soluciones

Para determinar las soluciones del sistema se analizó la red simplificada de 17 nodos (incluyendo IL-2, IL-4, GATA3, IL-12, IL-18, IFN $\gamma$ , T-bet, IL-1, IL-6, IL-17, IL-21, IL-23, STAT3, ROR $\gamma$ t, TGF $\beta$ lo, TGF $\beta$ hi y Foxp3). Se evaluaron los  $2^{17}$  (131072) estados posibles. De estos estados el 96.44% son cúspides. De los estados restantes 4355 (3.32%) son transitivos, hay 256 (0.19%) puntos fijos y 60 (0.05%) estados que corresponden a ciclos de tamaño dos y tres [Cuadro: 6.1]. No todos los estados iniciales se encuentran necesariamente presentes en los seres vivos. La mayor parte de los estados de la red son cúspides, lo cual implica que esas combinaciones de genes son inestables y probablemente no se encuentren en los seres vivos.

	Cantidad	Porcentaje
Puntos fijos	256	0.19
Ciclos	60	0.05
Transitivos	4355	3.32
Cúspides	126401	96.44
Estados totales	131072	100

**Cuadro 6.1:** Clasificación de los  $2^{17}$  estados generados por la red.

## 6.1. Soluciones puntuales

Fue posible recuperar los atractores clásicos de Th0, Th1, Th2, Th17 y Treg [Cuadro: 6.2]. Esto muestra que la red refleja el comportamiento observado en la diferenciación de las células Th. También fue posible recuperar el estado sin expresión de ninguna citocina o gen.

	TBET	IL-12	IL-18	IFG $\gamma$	GATA3	IL-4	IL-2	ROR $\gamma$ t	IL-1	IL-6	IL-17	IL-21	IL23	STAT3	Foxp3	TGF $\beta$ lo	TGF $\beta$ hi
Th1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Th2	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
Th17	0	0	0	0	0	0	0	1	1	1	1	1	1	1	0	1	0
Treg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1

**Cuadro 6.2:** Atractores de Th1, Th2, Th17 y Treg

Varias de las citocinas incluidas en la red no son producidas por las células Th, sino por otras células del sistema inmune. Este es el caso de IL-1, IL-2, IL-12, IL-18 e IL-23. Por lo tanto, se definió que una solución puntual correspondía a un linaje si expresaba su factor transcripcional maestro y sus citocinas características [Cuadro: 6.3]. En el caso de Th0 se consideró que una célula era Th0 si no expresa ninguno de los factores transcripcionales maestros. Utilizando estos criterios se encontró que a cada linaje le corresponden varias soluciones puntuales.

### 6.1.1. Ciclos

Se encontraron 60 estados correspondientes a ciclos: doce de estos ciclos son de tamaño dos y otros doce de tamaño tres. Biológicamente, la presencia de ciclos está relacionada con oscilaciones periódicas en la expresión de genes específicos, por ejemplo, los relojes circadianos. Otras oscilaciones en sistemas biológicos están asociadas con el espacio como los genes de la regla par en *Drosophila*. La mayoría de estos ciclos están relacionados con mecanismos de autoregulación. Otra posibilidad es que la producción de proteínas no alcance el umbral necesario para estabilizarse y oscile cerca de éste. Sin embargo, es imposible asegurar que los ciclos obtenidos

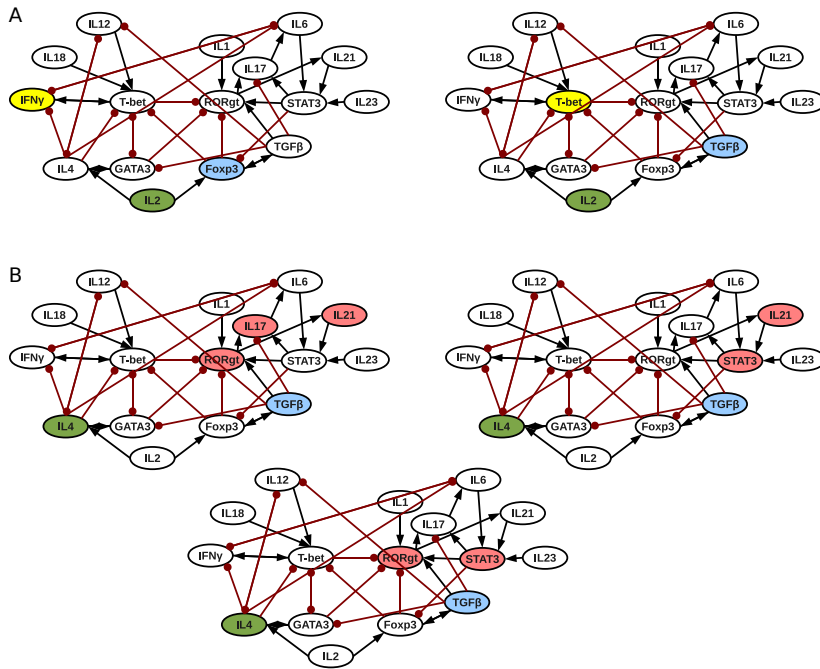
Linaje	Citocinas	Soluciones
Th0	ninguna	24
	IL-6	24
	TGF $\beta$ lo	4
	IL-4, TGF $\beta$ lo	4
	<b>Total Th0</b>	<b>56</b>
Th1	Tbet, IFN $\gamma$ , IL-12, IL-18	8
	Tbet, IFN $\gamma$	48
	Tbet*	40*
	Tbet, IL-12, IL-18	8
	<b>Total Th1</b>	<b>56</b>
Th2	GATA-3, IL-4	16
	GATA-3, IL-4, IL-2	8
	GATA3+ IL-4	48
	<b>Total Th2</b>	<b>64</b>
Th17	ROR $\gamma$ t, IL-17	32
	ROR $\gamma$ t, IL-1, 6, 17, 21, 23 y TGF $\beta$ lo	4
	<b>Total Th17</b>	<b>32</b>
Treg	Foxp3, TGF $\beta$ hi	8
	<b>Total Treg</b>	<b>8</b>

**Cuadro 6.3:** Criterios de clasificación de linajes de células Th.

no sean un resultado de la metodología usada. Sería necesario llevar a cabo análisis con otras metodologías como redes continuas o estocásticas y buscar evidencias experimentales antes de poder asegurar que tengan relevancia biológica.

Los ciclos de tamaño dos se caracterizan por que todos tienen oscilaciones de Foxp3 y TGF $\beta$ hi, además de expresar IL-2 y no expresar IL-12, TGF $\beta$ lo y los factores de Th17: ROR $\gamma$ t, IL-6, IL-17, IL-21, IL-23 y STAT3. En cuatro de los ciclos se observan oscilaciones adicionales de T-bet con IFN $\gamma$ , correspondientes a Th1 [Fig: 6.1]. Existen otros cuatro ciclos con oscilaciones de GATA-3, correspondientes a Th2.

Una posibilidad es que estos ciclos correspondan a estados indiferenciados de células Th0. Otra posibilidad es que sean estados intermedios de diferenciación de las células Treg, sobretodo cuando éstas se ven expuestas a ruido. Las células Treg son el linaje menos estable [15] y es fácil polarizarlas en Th17 o Th1, estos ciclos



**Figura 6.1:** Ejemplo de ciclos de tamaño dos y tres. A) En el ciclo de tamaño dos se puede ver la oscilación de Foxp3 con TGF $\beta$ hi y de Tbet con IFN $\gamma$ . B) En el ciclo de tamaño tres se puede ver la oscilación de ROR $\gamma$ t con IL-17, IL-21 y STAT3. Los nodos inactivos (0) se encuentran en blanco, los nodos activos (1) se encuentran coloreados.

podrían estar implicados en esta polarización. Cabe destacar la oscilación Foxp3-TGF $\beta$ hi, ya que se ha visto que la expresión transitiva de Foxp3 es común y las células Treg pueden fácilmente diferenciarse en otros linajes de células Th [49].

Los ciclos de tamaño tres se caracterizan por que todos presentan oscilaciones de ROR $\gamma$ t, IL-21 y STAT3, además de expresar TGF $\beta$ lo. Cuatro de los ciclos también presentan oscilaciones en la expresión de IL-17 [Fig: 6.1]. Al igual que los ciclos de tamaño dos, los ciclos de tamaño tres podrían considerarse parte del atractor Th0. Por otro lado Th17 es el linaje que más fácilmente se transforma en otros linajes después de Treg [15]. La presencia de estos ciclos podría explicar su plasticidad al servir de un estado intermedio anterior a la polarización en Th1.

## 6.2. Modularidad

Para ver la modularidad de la red se calculó la matriz de adyacencia  $A$  [Cuadro: 6.4]. Las interacciones de activación se representan con 1 y las de inhibición con -1. La matriz  $A$  fue dividida en regiones distinguibles que definen conjuntos de nodos que pueden agruparse en cúmulos. Si la región se forma únicamente de 0, se simplifica a 0; si contiene 0 y 1, se simplifica como 1; si tiene 0 y -1, se simplifica como -1 y si tiene 1 y -1, se simplifica como  $\pm 1$ .

$$A = \left( \begin{array}{cccc|cccc|cccc|cccc} +1 & 0 & 0 & +1 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & +1 & +1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & -1 & +1 & +1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ \hline -1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & +1 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & +1 & -1 & 0 \end{array} \right)$$

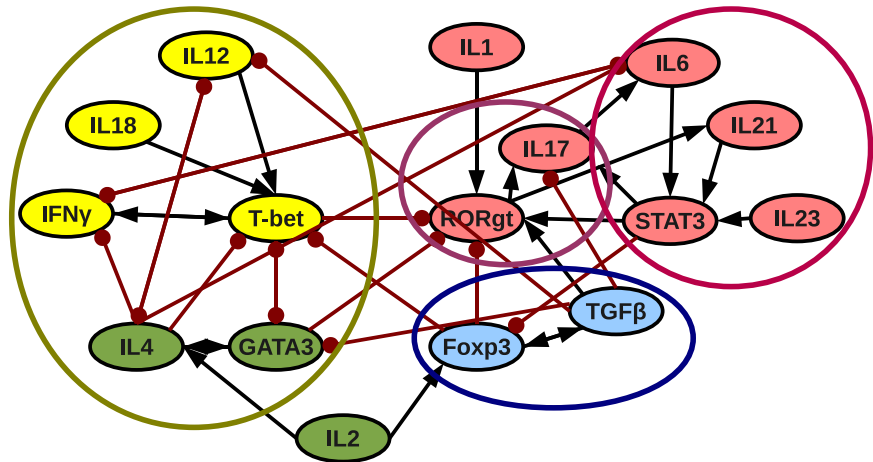
**Cuadro 6.4:** Matriz de adyacencia de la red. Los genes son en orden: TBET, IL-12, IL-18, IFG $\gamma$ , GATA3, 1L-4, IL-2, ROR $\gamma$ t, IL-1, IL-6, IL-17, IL-21, IL23, STAT3, Foxp3, TGF $\beta$ lo y TGF $\beta$ hi

A partir de la matriz simplificada la red puede ser reducida a seis módulos: Th1-Th2, Th17(A) y (B), Treg, IL-2 e IL-1 [Fig: 6.2, Cuadro: 6.5]. Los módulos corresponden parcialmente a los genes de cada linaje. Th1 y Th2 fueron agrupadas en un solo módulo con el signo  $\pm 1$ . Esto parece reflejar la cercana regulación de la diferenciación de ambos linajes. Th17 fue dividida en dos módulos: uno compuesto de ROR $\gamma$ t e IL-6 y otro compuesto de IL-17, IL-21, IL-23 y STAT3, lo cual podría hablar de que hay diferencias en las interacciones de las diversas citocinas de este linaje.



	Th1-Th2	IL2	Th17(A)	IL1	Th17(B)	Treg
Th1-Th2	±1	+1	-1	0	0	-1
IL2	0	+1	0	0	0	0
Th17(A)	-1	0	+1	+1	+1	±1
IL1	0	0	0	+1	0	0
Th17(B)	0	0	+1	0	+1	-1
Treg	0	+1	0	0	-1	±1

**Cuadro 6.5:** Al analizar la red se encontró que se puede agrupar en seis módulos. Estos corresponden a Th1-Th2, Th17(A) y (B), Treg, IL-2 e IL-1.



**Figura 6.2:** Módulos de la red

No fue posible agrupar a IL-1 e IL-2 en ningún módulo, esto podría deberse a su rol en el sistema inmune. Aunque IL-2 puede ser considerada una citocina de Th2, está principalmente relacionada con la activación de las células Th. IL-1 es una citocina relacionada con diversos aspectos de la inflamación y está especialmente relacionada con los macrófagos.

### 6.3. Centralidad

La centralidad de grado refleja que tan conectado está un nodo y puede ser interpretada como la facilidad con la que un nodo es afectado por alteraciones en la red. Los nodos con mayor centralidad de grado son: Tbet, GATA-3, IL-4 y ROR $\gamma$ t.

La centralidad de intermediación habla de cuantas trayectorias pasan por un nodo. Esto puede ser interpretado como en cuantas vías está implicado un gen. Los nodos con mayor centralidad de intermediación son: Tbet, IFN $\gamma$ , ROR $\gamma$ t e IL-6.

Cabe destacar que tanto Tbet como ROR $\gamma$ t tienen altos grados de centralidad. Esto implica que son nodos fundamentales en la transición de la información dentro la red de regulación.

	Centralidad de grado	Centralidad de intermediación
TBET	0.69	0.11
IL4	0.63	0.05
GATA3	0.56	0.03
ROR $\gamma$ t	0.56	0.1
IL12	0.5	0.04
IL6	0.44	0.09
TGFBhi	0.44	0.04
IFN $\gamma$	0.38	0.1
IL21	0.38	0
FOXP3	0.38	0.05
TGFBlo	0.38	0
IL2	0.25	0
IL17	0.25	0.07
STAT3	0.25	0
IL18	0.19	0
IL1	0.19	0
IL23	0.06	0

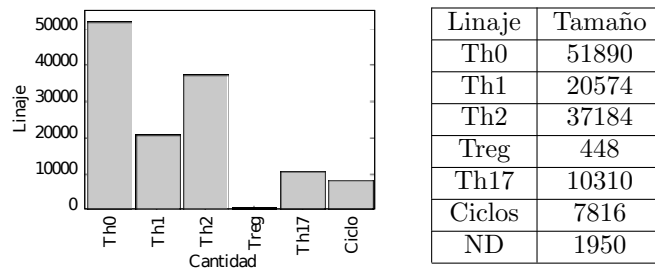
**Cuadro 6.6:** Centralidad de grado y de intermediación de la red

---

# Capítulo 7

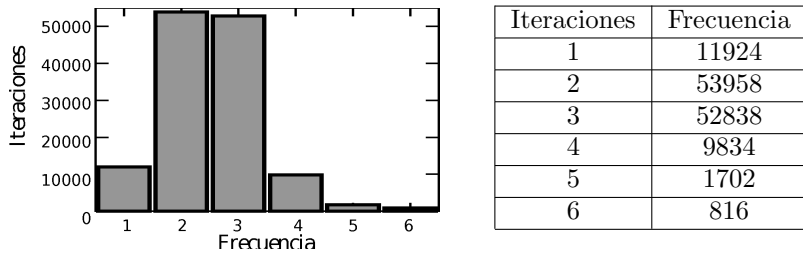
## Atractores

La red de estados consta de 130172 estados que corresponden a las  $2^{17}$  condiciones iniciales. Los atractores de esta red son las 256 soluciones puntuales más los 60 estados correspondientes a soluciones cíclicas. Utilizando las clasificaciones antes descritas se calculó la cuenca de atracción de cada linaje [Fig: 7.1]. Cada cuenca consta de varias soluciones puntuales diferentes. Se consideraron aparte los estados que llevan a ciclos. A partir de esto se pudo determinar que la cuenca de mayor tamaño corresponde a Th0, seguido por Th2, luego Th1, después Th17 y finalmente Treg, lo cual coincide con observaciones de la estabilidad de distintos linajes [15].



**Figura 7.1:** Tamaño de la cuenca de atracción de los distintos linajes.

En promedio, se necesitaron 2.53 iteraciones para comprobar que se ha llegado a un punto fijo o ciclo. El mínimo de iteraciones necesarias fue una y el máximo seis [Fig: 7.2]. Esto implica que la proporción de soluciones es pequeña con respecto al espacio de estados total y que se llega rápidamente a soluciones.



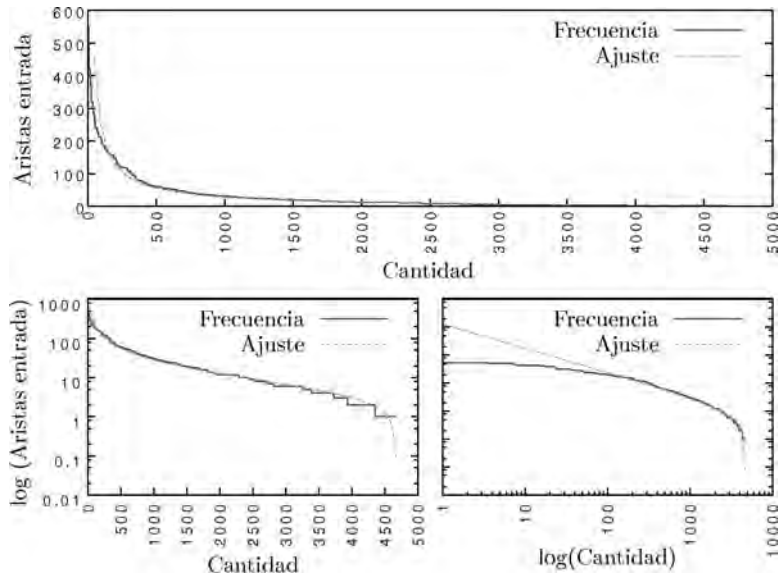
**Figura 7.2:** Iteraciones necesarias para comprobar que se ha llegado a una solución. El promedio se necesitaron 2.53 iteraciones

Cada nodo de la red de estados tienen una sola arista de salida, mientras que puede tener hasta  $2^{17}$  aristas de entrada. Las cúspides no tienen ninguna arista entrante, mientras que en las soluciones puntuales su arista de salida los comunica consigo mismos. Normalmente se considera que si la red está en fase crítica entonces la distribución de las aristas de entrada de la red sigue una ley de potencias [69]. Sin embargo, las leyes de potencias fallan al ajustar las colas de la distribución.

Al graficar la cantidad de aristas de entrada por nodo [Fig: 7.3] se ve que la mayoría de los nodos tienen pocas entradas y unos pocos nodos tienen muchas entradas. Estas pueden ser ajustadas a una función similar a la función  $\beta$  [70]:

$$y = \frac{(R - x + 1)^b}{r^a}$$

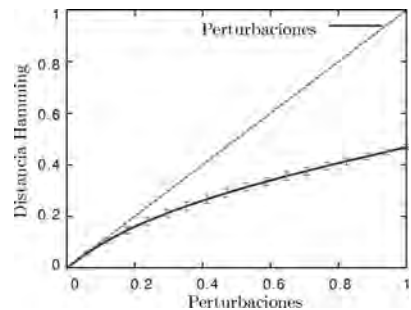
donde  $a$  y  $b$  son ajustados a la gráfica,  $R$  es el número de datos y  $K$  depende de  $a$  y  $b$ . En este caso  $a = 0,87$  y  $b = 0,60$  con  $r = 0,970$ .



**Figura 7.3:** Cantidad de aristas de entrada por nodo. Por claridad no se incluyen las cúspides. Se presentan los mismos datos en A) Sistema decimal, B) Semilog y C) Loglog

## 7.1. Efectos de las perturbaciones

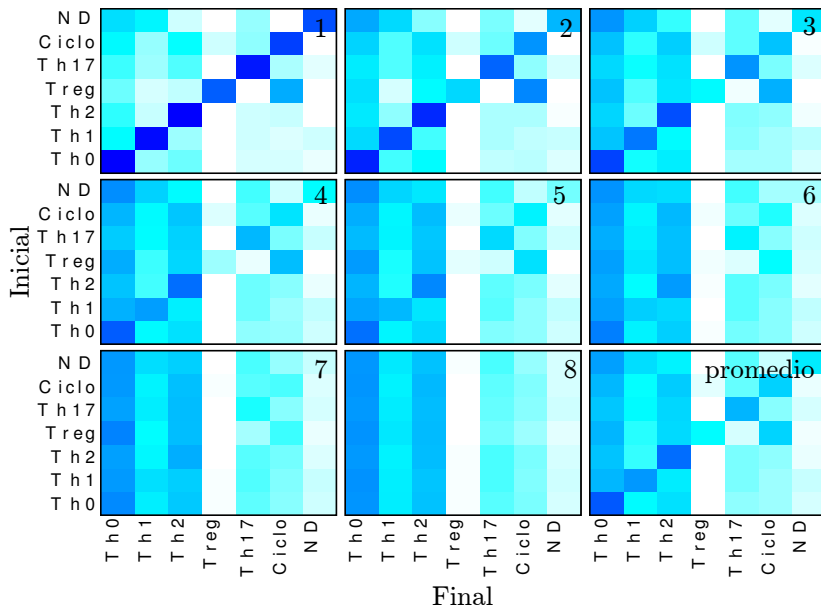
Para analizar la estabilidad de los atractores se les sometió a ruido. Se cambiaron  $n$  nodos al azar y se buscó la distancia de Hamming entre el atractor original y el nuevo [Fig: 7.4]. Al analizar las sensibilidad a las perturbaciones de la red se ve que cerca del origen la pendiente se aproxima a la recta identidad, lo cual es característico de las redes en fase crítica. Si se lleva a cabo una sola perturbación cambia en promedio menos de un nodo, esto demuestra que la red es robusta ante el ruido.



**Figura 7.4:** Sensibilidad a las perturbaciones. Se graficó el porcentaje de nodos perturbados contra la distancia de Hamming. Se analizaron 1700000 estados en total.

La red tiene 17 nodos y 51 interacciones, un parámetro de conectividad promedio  $K = 3$ , una densidad  $\rho = 0,1875$  con  $p = 0,33$  y  $K_c = 2,26$ , de tal forma que está en la fase caótica. Sin embargo, la red original tiene 42 nodos y 84 interacciones con  $K = 1,95$  y una  $\rho = 0,006024096$ , la  $p = 0,43$  con  $K_c = 2,04$ . La diferencia entre ambas es de 4.1%, haciendo que la red esté en la fase ordenada pero muy cerca de la fase crítica. En el caso de las células Th el organismo necesita mantener una respuesta constante ante un mismo patógeno. Pero al mismo tiempo, las células Th necesitan poder adaptarse a distintos retos inmunológicos. El hecho de que la red esté en la fase crítica es un compromiso entre la plasticidad y la robustez necesarias para una adecuada respuesta inmune.

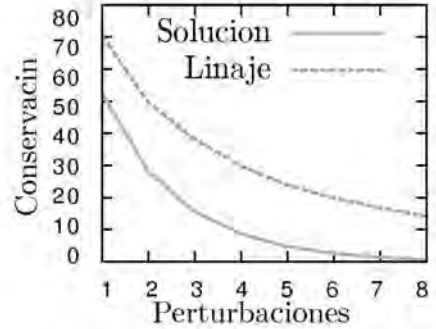
Para ver la estabilidad de los linajes se les sometió a ruido y se vio a que linaje tiende el nuevo estado [Fig: 7.5]. Se perturbaron entre uno y ocho nodos y se comparó el linaje original con el obtenido después de la perturbación utilizando la clasificación de linajes antes expuesta.



**Figura 7.5:** Efectos del ruido en la transición de linajes. La escala va de 0% a 80%. Cada recuadro corresponde al número de nodos perturbados, se analizaron 100000 condiciones.

Se vio que los linajes se conservan a pesar del efecto del ruido, lo cual puede deberse a sus grandes cuencas. Conforme se cambiaron más nodos la red tendió a cambiar más entre un linaje y otro concentrándose en Th0, Th2 y en menor medida en Th1. Ante un solo cambio el 70% de los estados permanecieron en el mismo linaje, sin embargo ante seis solo el 20% permanecieron. Ante un solo cambio el 52% de los estados regresan al mismo punto fijo, mientras que el 70% permanece en el mismo linaje. Esto se debe a que cada atractor incluye varios puntos fijos diferentes.

Treg es el atractor más pequeño y en presencia de ruido tiende a atractores cíclicos. Varios de estos atractores cíclicos expresan Foxp3 de manera oscilatoria, lo cual podría explicar este patrón. Curiosamente, de los estados que se convierten en Treg la mayoría son ciclos.



**Figura 7.6:** Conforme se incrementó el ruido de la red menos estados permanecieron en el mismo punto fijo. La cantidad de estados que permanecen en el mismo linaje disminuye más lentamente.

---

## Capítulo 8

# Conclusiones

El estudio de la diferenciación de células Th es relevante para entender la forma en la que el sistema inmune reacciona ante los diversos retos inmunológicos. El saber como son reguladas las células Th nos da pistas acerca de como tratar problemas de salud como el VIH y la diabetes tipo 2. Al entender las interacciones de las diversas citocinas y factores transcripcionales involucrados podemos elegir mejores blancos terapéuticos.

Utilizando redes es posible modelar sistemas complejos como los biológicos, ya que las redes permiten reflejar las interacciones entre los distintos elementos del sistema. La generación de modelos dinámicos permite entender su funcionamiento y generar nuevas hipótesis que pueden ser verificadas experimentalmente. Durante la elaboración de este trabajo, se publicaron dos artículos que utilizaron redes para modelar la diferenciación de células Th [13, 14].

La red propuesta permite recuperar las soluciones clásicas de Th0, Th1, Th2, Th17 y Treg. La red también mostró patrones de expresión de genes y citocinas que son diferentes a los clásicos pero que han sido reportados en la literatura [34, 38, 39, 68]. Además, el tamaño de las cuencas de atracción predichas coincide con las que se han supuesto a partir de observaciones experimentales [15].



Un detalle interesante es que fue necesario duplicar el nodo de  $TGF\beta$  según su nivel de expresión para ilustrar el comportamiento del sistema. Si se genera un solo nodo de  $TGF\beta$  se pierden los atractores correspondientes a Th17. Esto muestra que es posible recuperar el comportamiento del sistema al suponer que una proteína no se comporta de una manera binaria. Cuando se reporta que una proteína tiene acciones opuestas puede deberse a que diferentes concentraciones activan diferentes vías.

Al ver la modularidad de la red se encontró que puede ser dividida en seis módulos. Estos módulos corresponden aproximadamente a un módulo con los genes de Th1 y Th2, otro correspondiente a Treg, las citocinas y genes de Th17 se encuentran divididas en dos módulos. Curiosamente, fue imposible agrupar IL-1 e IL-2 quedando cada uno como un módulo propio, lo cual podría reflejar su función en el sistema inmune.

La red se encuentra en fase crítica al igual que otras redes biológicas estudiadas [58]. Las células Th necesitan ser capaces de reaccionar de manera distinta ante distintos retos inmunológicos, pero también necesitan ser capaces de montar una respuesta inmune robusta que permita el control y eliminación de los patógenos. Estar en fase crítica representa un compromiso entre plasticidad y robustez.

Esta red permite estudiar la dinámica de la diferenciación de células Th y refleja diversas observaciones experimentales. Sin embargo aún falta mucho por hacer. Es necesario estudiar el efecto de activar constitutivamente o deletar nodos específicos de la red y contrastarlos con los síndromes conocidos. Además se deben agregar nuevos genes, citocinas y factores ambientales para conocer a mayor detalle el comportamiento de la red y encontrar nuevas relaciones.

Entender la diferenciación de las células Th no sólo tiene un interés académico. Conociendo la red es posible buscar nuevos blancos terapéuticos para tratar problemas de salud relacionados con la incorrecta regulación del sistema inmune como el SIDA, la diabetes y las enfermedades autoinmunes. Sin embargo son problemas complejos y aún hace falta mucho trabajo para poder resolver los retos que plantean.

---

# Bibliografía

- [1] World Health Organization. The global burden of disease. *World Health Organization*, 2004.
- [2] M Clericia and GM Shearer. A TH1-TH2 switch is a critical step in the etiology of HIV infection. *Immunology Today*, 14(3):107–111, 1993.
- [3] E Maggi, M Mazzetti, A Ravina, F Annunziato, M de Carli, MP Piccinni, R Manetti, M Carbonari, AM Pesce, and G del Prete. Ability of HIV to promote a Th1 to Th0 shift and to replicate preferentially in Th2 and Th0 cells. *Science*, 265(5169):244–248, 1994.
- [4] J Jian Liu, A Divoux, J Sun, J Zhang, K Clément, JN Glickman, GK Sukhova, PJ Wolters, J Du, CZ Gorgun, A Doria, P Libby, RS Blumberg, BB Kahn, GS Hotamisligil, and GP Shi. Genetic deficiency and pharmacological stabilization of mast cells reduce diet-induced obesity and diabetes in mice. *Nature Medicine*, 15:940–945, 2009.
- [5] S Nishimura, I Manabe, M Nagasaki, K Eto, H Yamashita, M Ohsugi, M Otsu, K Hara, K Ueki, S Sugiura, K Yoshimura, Y Kadowaki, and R Nagai. CD8+ effector T cells contribute to macrophage recruitment and adipose tissue inflammation in obesity. *Nature Medicine*, 15:914–920, 2009.
- [6] S Winer, Y Chan, G Paltser, D Truong, H Tsui, J Bahrami, R Dorfman, Y Wang, J Zielenski, F Mastronardi, Y Maezawa, DJ Drucker, E Engleman, D Winer, and HM Dosch. Normalization of obesity-associated insulin resistance through immunotherapy. *Nature Medicine*, 15:921–929, 2009.

- 
- [7] M Feuerer, L Herrero, D Cipolletta, A Naaz, J Wong, A Nayer, J Lee, AB Goldfine, C Benoist, S Shoelson, and D Mathis. Lean, but not obese, fat is enriched for a unique population of regulatory T cells that affect metabolic parameters. *Nature Medicine*, 15:930–939, 2009.
- [8] K Choi and YB Kim. Molecular mechanism of insulin resistance in obesity and type 2 diabetes. *The Korean Journal of Internal Medicine*, 25(2):119–129, 2010.
- [9] J. Zhu and W. E. Paul. CD4 T cells: fates, functions, and faults. *Blood*, 112(5):1557–1568, 2008.
- [10] G von Dassow, E Meir, and G Odell. The segment polarity network is a robust developmental module. *Nature*, 406:188–193, 2000.
- [11] C. Espinosa-Soto, P Padilla-Longoria, and Alvarez-Buylla E. A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. *Plant Cell*, 16:2923–2939, 2004.
- [12] L Mendoza. A network model for the control of the differentiation process in Th cells. *BioSystems*, 84:101–114, 2006.
- [13] L Mendoza and F Pardo. A robust model to describe the differentiation of t-helper cells. *Theory in Biosciences*, 129(4):183–293, 2010.
- [14] A Naldi, J Carneiro, C Chaouiya, and D Thieffry. Diversity and plasticity of th cell types predicted from regulatory network modelling. *Plos Computational Biology*, 6(9), 2010.
- [15] KM Murpy and B Stockinger. Effector T cell plasticity: flexibility in the face of changing circumstances. *Nature Reviews Immunology*, 11(8):674–680, 2010.
- [16] J Zhu and WE Paul. Peripheral CD4+ T-cell differentiation regulated by networks of cytokines and transcription factors. *Immunological Reviews*, 238: 247–263, 2010.

- 
- [17] B Alberts, A Johnson, J Lewis, M Raff, K Roberts, and P Walter. *Molecular Biology of the Cell, 4th edition*. Garland Science, 2004.
- [18] Wikimedia: Central dogma. <http://commons.wikimedia.org/wiki/File:CentralDogma1970.es.png>, 2010.
- [19] AD Hershey and M Chase. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *The Journal of General Physiology*, 36(1): 39–56, 1952.
- [20] Wikimedia: Estructura del adn. <http://commons.wikimedia.org/wiki/File:DNAChemicalStructure.es.svg>, 2010.
- [21] B Constanze and PN Cockerill. *Chromatin Mechanisms Regulating Gene Expression In Health And Disease*. Landes Bioscience and Springer Science, 2010. Retrived from Madam Curie Bioscience Database.
- [22] K Chen and N Rajewsky. The evolution of gene regulation by transcription factors and microRNAs. *Nature Reviews Genetics*, 8:93–103, 2007.
- [23] W Fiers, R Contreras, F Duerinck, G Haegeman, D Iserentant, J Merregaert, W Min Jou, F Molemans, A Raeymaekers, A Van den Berghe, G Volckaert, and M Ysebaert. Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature*, 260:500–507, 1976.
- [24] SP Gygi, Y Rochon, BR Franza, and R Aebersold. Correlation between protein and mRNA abundance in yeast. *Molecular Cell Biology*, 19(3):1720–1730, 1999.
- [25] P Aloy and RB Rusell. Structural systems biology: modelling protein interactions. *Nature Reviews Molecular Cell Biology*, 7:188–197, 2006.
- [26] CA Janeway, P Travers, M Walport, and MJ Shlomchik. *Immunobiology*. Garland Science, 2001.
- [27] RA Goldsby, TJ Kindt, and BA Osborne. *Kuby Immunology*. W.H. Freeman, 2006.

- [28] AK Abbas and AH Lichtman. *Immunología celular y molecular*. Mc Graw-Hill, 2004.
- [29] T.R. Mosmann, H Cherwinski, M.W. Bond, M.A Giedlin, and R.L. Coffman. Two types of murine helper T cell clone. i. definition according to profiles of lymphokine activities and secreted proteins. *The Journal of Immunology*, 136(7):2348–2357, 1986.
- [30] S Sakaguchi, N Sakaguchi, M Asano, M Itoh, and M Toda. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *Journal of Immunology*, 55(3):1151–1164, 1995.
- [31] M Veldhoen, RJ Hocking, CJ Atkins, RM Locksley, and B Stockinger. TGFB in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity*, 24(2):179–189, 2006.
- [32] MC Anaya, CD Cañas, Cervera RS, and PV Correa. *Auto inmunidad y enfermedad autoinmune*. Corporación para investigaciones biológicas, 2005.
- [33] A Giovannetti, M Pierdominici, F Mazzetta, M Marziali, C Renzi, AN Mileo, M de Felice, B Mora, A Esposito, R Carello, MG Pizzuti, A ans Paggi, R Paganelli, W Malorni, and F Aiuti. Unravelling the complexity of T cell abnormalities in common variable immunodeficiency. 2007, 178(3932-3943), *The Journal of Immunology*.
- [34] T Aarvak, M Chabaud, P Miossec, and JB Natvig. IL-17 is produced by some proinflammatory Th1/Th0 cells but not by Th2 cells. *The Journal of Immunology*, 162:1246–1251, 1999.
- [35] PKK Wong, PJ Egan, BA Croker, K O'Donnell, NA Sims, SK Hiu Kiu, EJ McManus, WS Alexander, AW Roberts, and IP Wicks. SOCS-3 negatively regulates innate and adaptive immune mechanisms in acute IL-1-dependent inflammatory arthritis. *The Journal of Clinical Investigation*, 116(6):1571–1581, 2006.

- [36] K Takada and SC Jameson. Naive T cell homeostasis: from awareness of space to a sense of place. *Nature Reviews Immunology*, 9:283–832, 2009.
- [37] KM Murphy and SL Reiner. The lineage decisions of helper T cells. *Nature Reviews Immunology*, 2:933–944, 2002.
- [38] J Zhu, H Yamane, J Cote-Sierra, L Guo, and WE Paul. GATA-3 promotes Th2 responses through three different mechanisms: induction of Th2 cytokine production, selective growth of Th2 cells and inhibition of Th1 cell-specific factors. *Cell Research*, 16:3–10, 2006.
- [39] KA Mowen and LH Glimcher. Signaling pathways in Th2 development. *Immunological reviews*, 202(1):203–222, 2004.
- [40] SQ Crome, AY Wang, and MK Levings. Translational mini-review series on Th17 cells: Function and regulation of human T helper 17 cells in health and disease. *Clinical and Experimental Immunology*, 159(2):109–119, 2010.
- [41] M Kawaguchi, M Adachi, N Oda, F Kokubu, and S Huang. IL-17 cytokine family. *Journal of Allergy and Clinical Immunology*, 114(6):1265–1273, 2004.
- [42] JE Sims and DE Smith. The IL-1 family: regulators of immunity. *Nature Reviews Immunology*, 2010.
- [43] L Attisano and JL Wrana. Signal transduction by the TGF-beta superfamily. *Science*, 296:1646–1647, 2002.
- [44] ES Hwang, SJ Szabo, PL Schwartzberg, and LH Glimcher. T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science*, 307, 2005.
- [45] B Afzali, P Mitchell, RI Lechler, and G Lombardi. Translational mini-review series on Th17 cells: Induction of interleukin-17 production by regulatory T cells. *Clinical and Experimental Immunology*, 159(2):120–130, 2010.
- [46] E de Jong, T Suddason, and GM Lord. Translational mini-review series on Th17 cells: Development of mouse and human T helper 17 cells. *Clinical and Experimental Immunology*, 159(2):148–158, 2010.

- [47] M Saraiva and A O'Garra. The regulation of IL-10 production by immune cells. *Nature Reviews Immunology*, 10, 2010.
- [48] A Yoshimura, T Naka, and M Kubo. SOCS proteins, cytokine signalling and immune regulation. *Nature Reviews Immunology*, 7:454–470, 2007.
- [49] JJ O'Shea and WE Paul. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science*, 327(5969):1098–1102, 2010.
- [50] KE Wellen and GS Hotamisligil. Inflammation, stress and diabetes. *The Journal of Clinical Investigation*, 115(5), 2005.
- [51] V Cecchinato, CJ Trindade, A Laurence, JM Heraud, JM Brenchley, MG Ferrari, L Zaffiri, E Trynieszewska, WP Tsai, M Vaccari, R Washington Parks, D Venzon, DC Douek, JJ O'Shea, and G Franchini. Altered balance between th17 and th1 cells at mucosal sites predicts aids progression in simian immunodeficiency virus-infected macaques. *Mucosal Immunology*, 1:279–288, 2008.
- [52] G Thorborn, L Pomeroy, H Isohanni, M Perry, B Peters, and A Vyakarnam. Increased sensitivity of CD4+ T-effector cells to CD4+CD25+ Treg suppression compensates for reduced Treg number in asymptomatic HIV-1 infection. *Plos One*, 5, 2010.
- [53] I. D. Vega del Valle. Modelos dinámicos continuos para la red de regulación genética del nicho de células troncales del meristemo apical de raíz de *Arabidopsis thaliana*. Master's thesis, Universidad Nacional Autónoma de México, 2010. Tesis Maestría (Maestría en Ciencias (Física))-UNAM, Facultad de Ciencias.
- [54] D Kaplan and L Glass. *Understanding Nonlinear Dynamics*. Springer, 1995.
- [55] S.A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3):437–467, 1969.
- [56] C. Gershenson. Guiding the self-organization of random boolean networks. *Manuscript*, 2010. manuscript.
- [57] M. Aldana. Boolean dynamics of networks with scale-free topology. *Physica D: Nonlinear Phenomena*, 185(1):45–66, 2003.

- [58] E Balleza, ER Alvarez-Buylla, A Chaos, S Kauffman, I Shmulevich, and M Aldana. Critical dynamics in genetic regulatory networks: Examples from four kingdoms. *PLoS ONE*, 3(6), 2008.
- [59] L Mendoza, D Thieffry, and ER Alvarez-Buylla. Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics*, 15(7):593–606, 1999.
- [60] R Albert and HG Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. *Journal of Theoretical Biology*, 223:1–18, 2002.
- [61] R Albert. Network inference, analysis, and modeling in systems biology. *The Plant Cell*, 19:3327–3338, 2007.
- [62] U. Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8:450–461, 2007.
- [63] SA Kauffman. *Investigations*. Oxford University Press, 2000.
- [64] *The information dynamics of phase transitions in random boolean networks*, Artificial Life XI - Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems, 2008. MIT Press.
- [65] *A fisher information study of phase transitions in random boolean networks*, Artificial Life XII - Proceedings of the Twelfth International Conference on the Simulation and Synthesis of Living Systems, 2010. MIT Press.
- [66] C Ting, MC Olson, KP Barton, and JM Leiden. Transcription factor GATA-3 is required for development of the T-cell lineage. *Nature*, 384:474–478, 1996.
- [67] Networkx, python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. <http://networkx.lanl.gov/index.html>, 2010.
- [68] A O’Garra and N Arai. The molecular basis of T helper 1 and T helper 2 cell differentiation. *Trends in cell biology*, 10(12):542–550, 2000.



- 
- [69] M Aldana, E Balleza, S Kauffman, and O Resendiz. Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical Biology*, 245:433–448, 2007.
- [70] G Naumis and Cocho G. The tails of rank-size distributions due to multiplicative processes: from power laws to stretched exponentials and beta-like functions. *New Journal of Physics*, 9(8), 2007.