



UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

---

---

FACULTAD DE CIENCIAS

Estructura y función en redes y su relación con el  
potencial evolutivo.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Biólogo

PRESENTA:

David Gordillo Alaniz

DIRECTOR DE TESIS:

Pablo Padilla Longoria

Ciudad Universitaria, Cd. Mx.

2018





Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



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

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

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

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

# Contenido

<b>Introducción</b>	<b>3</b>
<b>Sección 1: Marco teórico.</b>	<b>6</b>
Redes dinámicas como modelos de sistemas biológicos. . . . .	6
Estructura. . . . .	7
Dinámica. . . . .	9
Función. . . . .	12
El potencial evolutivo. . . . .	14
La red de Hopfield. . . . .	18
Similitudes entre redes de regulación genética y redes de Hopfield. . . . .	24
<b>Sección 2: Objetivos.</b>	<b>26</b>
<b>Sección 3: Método.</b>	<b>27</b>
El modelo . . . . .	27
Evaluación de la estabilidad de los atractores. . . . .	28
Carga. . . . .	29
Información mutua entre patrones. . . . .	30
Topología. . . . .	32
<b>Sección 4: Resultados y discusión.</b>	<b>34</b>
Simulaciones. . . . .	34
Carga. . . . .	34
Información mutua. . . . .	39
Topología. . . . .	44
Consideraciones biológicas del modelo. . . . .	47
Estructura y potencial evolutivo. . . . .	59
<b>Sección 5: Conclusiones.</b>	<b>62</b>

## Resumen

Las redes dinámicas (RD) son modelos matemáticos y computacionales que han sido ampliamente usados en biología ya que sus propiedades presentan similitudes con los sistemas biológicos. Una de estas propiedades es la capacidad de asociar estados iniciales con un comportamiento particulares, denominados atractores, a partir de información almacenada en su estructura, lo cual puede considerarse como una forma básica de categorización. Por esta razón, se ha hablado de que las redes dinámicas son capaces de crear representaciones internas, ya sea del entorno o de sí mismas. Recientemente, estudios realizados con redes neuronales artificiales (ReNA), un tipo de RD, han mostrado paralelismos entre éstas y redes de regulación genética (RRG) en cuanto a dinámica y evolución se refiere. Siendo que las ReNA son modelos empleados en el estudio de la memoria y el aprendizaje, estos hallazgos sirven como fundamento para la tesis defendida por algunos de que el potencial evolutivo, entendido como la capacidad de un sistema de evolucionar, es una forma restringida de aprendizaje. En este trabajo se exploran algunas estas ideas. Empleando una ReNA denominada red de Hopfield como modelo de una RRG y mediante simulaciones en computadora, se analiza cómo es que cambios en la estructura de la red afectan su función en términos de la estabilidad de sus atractores. Se comparan los resultados obtenidos por las simulaciones con datos biológicos y se analizan las implicaciones biológicas que podrían derivarse del modelo. Finalmente, se discute la relación que existe entre la estructura y el potencial evolutivo. Dadas las similitudes que existen, por un lado entre las redes dinámicas y los sistemas biológicos, y por el otro, entre el aprendizaje y la evolución, los modelos conexionistas del aprendizaje que emplean redes dinámicas como las ReNA pueden darnos información importante acerca de la biología y la evolución.

## Introducción

Un sistema es una entidad formada por un conjunto de componentes interdependientes que se comportan como un todo (von Bertalanffy, 1976). Los seres vivos por ejemplo, pueden ser entendidos como sistemas biológicos formados por moléculas, organelos y células que interactúan entre sí a distintos niveles de organización. Una propiedad de los sistemas es que exhiben comportamientos emergentes, es decir, dinámicas colectivas que surgen de la organización del sistema como unidad y que no están presentes en los componentes individuales (Trewavas, 2006). De hecho, muchas de las propiedades que caracterizan la organización biológica como el metabolismo, la multicelularidad, el desarrollo, el potencial evolutivo e incluso la conciencia, son emergentes, derivadas de procesos más básicos.

La biología de sistemas es el área de la biología que estudia a los seres vivos y sus procesos desde una perspectiva sistémica. Busca entender cómo a partir de las interacciones entre componentes es que emergen todos aquellos comportamientos que caracterizan la organización biológica (Kitano, 2001). Pero entender la complejidad de los sistemas biológicos no es tarea fácil, representa uno de los mayores retos para la biología moderna (Trewavas, 2006) y requiere de la integración de datos de diferentes fuentes de manera colaborativa (Knüpfner y Beckstein, 2013), así como de conceptos y herramientas provenientes de diversas disciplinas. Debido a la complejidad del problema, el papel que juegan los modelos y las simulaciones por computadora es crucial.

Uno de los modelos más usados para el estudio de los sistemas biológicos son las redes dinámicas, que combinan principios de la teoría de grafos y de la teoría de sistemas dinámicos no lineales (Jost, 2007). Una red dinámica está formada por un conjunto de elementos o agentes cuyos estados cambian en el tiempo en función de una regla de activación que toma en cuenta las interacciones con los elementos vecinos (Jost, 2007). Los elementos de una red pueden ser vistos como unidades computacionales que procesan entradas y emiten salidas, mientras que las conexiones determinan el flujo de información entre unidades (Gershenson, 2003). Las redes dinámicas, como sistemas abstractos, manifiestan propiedades emergentes que las hace interesantes por sí mismas, pero que a su vez tienen relevancia biológica (Wuensche, 2004). La propiedad más relevante es la presencia de cuencas de atracción que canalizan la dinámica de la red hacia estados estables de activación llamados atractores (Wuensche, 2004), los cuales dependen de la estructura de la red, es decir de la forma en la que se relacionan sus elementos. Esta capacidad de asociar un patrón de activación con un comportamiento particular a partir de información almacenada en la estructura, puede considerarse como una forma básica de categorización y es relevante para la biología por varias razones. En primer lugar porque el mapeo entre estructuras

y comportamiento se considera como una de las bases para el surgimiento de la función (Knüpfer y Beckstein, 2013), pero además, porque significa que las redes dinámicas son capaces de crear representaciones tanto del entorno como de sí mismas en su estructura. Esto brinda un enfoque para la comprensión de la memoria en distintos tipos de sistemas, desde neuronales hasta regulatorios (Wuensche, 2004).

Estudios con redes neuronales artificiales (ReNA), un tipo de redes dinámicas que han sido empleadas para estudiar la memoria y el aprendizaje, han mostrado similitudes con redes de regulación genética en cuanto a dinámica y evolución se refiere (Chalmers, 1990; Wuensche, 2004; Fernando *et al.*, 2009; Watson *et al.*, 2010, 2011a, 2014). Estas similitudes han hecho que el potencial evolutivo, entendido a groso modo como la capacidad de un sistema de evolucionar, sea considerado por algunos como una forma restringida de aprendizaje en la que el sistema asocia fenotipos que han sido visitados y seleccionados en el pasado (Valiant, 1984, 2009; Watson *et al.*, 2010). Además plantean la posibilidad de usar modelos pertenecientes a áreas como la inteligencia artificial para el estudio de la evolución. Teniendo esto en cuenta, uno de los retos de la biología de sistemas es entender cómo la estructura de los sistemas biológicos se organiza para manifestar comportamientos similares al aprendizaje que les permitan evolucionar sin comprometer los logros pasados y mantener a su vez el acoplamiento con el entorno.

Dadas las similitudes que existen, por un lado entre las redes dinámicas y los sistemas biológicos, y por el otro, entre el aprendizaje y la evolución, los modelos conexionistas del aprendizaje que emplean redes dinámicas como las ReNA pueden darnos información importante acerca de la biología y la evolución. El objetivo de este trabajo es explorar algunas ideas que este tipo de modelos pueden brindar a la biología, mediante simulaciones en computadora y desde una perspectiva teórica.

Empleando un tipo de ReNA denominada red de Hopfield como modelo de un sistema

de regulación genética hipotético, se estudió el efecto que tiene en la función la modificación de la estructura, tomando en cuenta las implicaciones biológicas que podrían derivarse del modelo. La organización general del trabajo es la siguiente: En la Sección 1 se proporciona el marco teórico para el estudio de las redes dinámicas, se definen los conceptos de estructura, dinámica, función y potencial evolutivo y se detalla cómo estos aspectos son abordados mediante los modelos de redes dinámicas. Posteriormente se introduce el tema de las redes de Hopfield y de las redes de regulación genética, haciendo énfasis en las semejanzas entre ambos modelos, así como entre el proceso evolutivo y el aprendizaje. En la sección 2 se especifican los objetivos particulares de este trabajo. En la Sección 3 se habla del método, se explica el diseño de la red así como la manera en la que fueron evaluadas la estructura y la función. En la Sección 4 se discuten los resultados de las simulaciones y se ahonda en las implicaciones biológicas, comparando los resultados obtenidos con datos biológicos reales. Al final de esta sección se discute la relación entre estructura y potencial evolutivo. Por último, en la Sección 5 se dan las conclusiones.

## **Sección 1: Marco teórico.**

### **Redes dinámicas como modelos de sistemas biológicos.**

Las redes dinámicas en biología las podemos encontrar en cualquier lugar, desde el cerebro hasta la ecología (Wuensche, 2004). Como ejemplo de ello tenemos a las redes metabólicas, las redes de señalización bioquímica, las redes de regulación genética, las redes neuronales y las redes tróficas, por mencionar algunas.

Existen varios aspectos de un sistema que pueden ser estudiados mediante estos modelos, pero dos de ellos son de particular importancia: la estructura y la dinámica. La estructura se refiere a los componentes y los patrones de conectividad, mientras que la dinámica se

refiere al comportamiento global de la red en el tiempo. A continuación se detalla cada uno de estos aspectos.

### **Estructura.**

La estructura de un sistema se define como el conjunto de componentes y relaciones que concretamente lo constituyen como unidad (Maturana y Varela, 1984). Una manera útil de representar dicha estructura es mediante una red.

Desde el punto de vista matemático una red es un grafo  $G = (V, E)$  donde  $V$  es un conjunto de nodos o vértices que representan a los elementos del sistema y  $E$  es un conjunto de aristas o conexiones que representan la relación entre elementos. Una conexión  $e = (v_i, v_j)$  se define en función de los dos vértices  $v_i$  y  $v_j$  que determinan la arista; cuando la dirección entre las aristas importa, es decir si  $(v_1, v_2) \neq (v_2, v_1)$ , se dice que la red es dirigida, de lo contrario se dice que es no dirigida y por lo tanto simétrica (Figura 1).

En general se define a  $|V| = N$  como el tamaño de la red y a  $|E| = M$  como el número total de aristas. Cuando se tiene que  $E = V \times V$  entonces se dice que la red es completa (Almaas *et al.*, 2008) ya que tiene el número máximo posible de aristas y por lo tanto está "completamente conectada". La conexión entre dos nodos indica que esos elementos están relacionados de algún modo, pero debido a que en la mayoría de los casos los sistemas naturales poseen una gran heterogeneidad en la intensidad de sus conexiones (Boccaletti *et al.*, 2006), a veces es necesario indicar la intensidad de dicha relación asignando pesos  $w$  entre las conexiones (Figura 2).



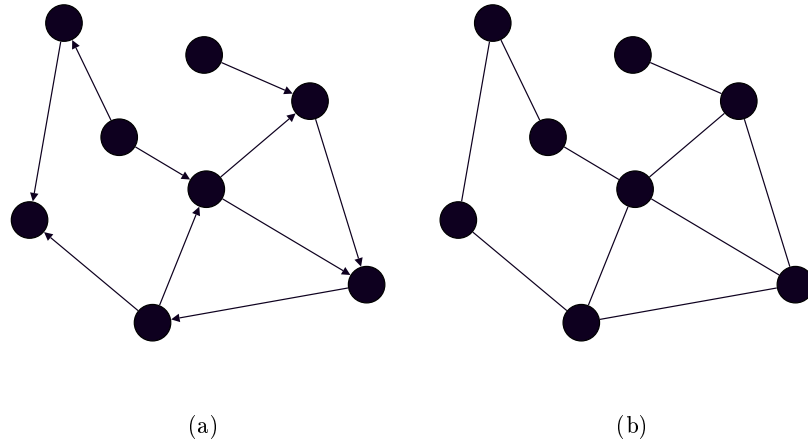


Figura 1: Esquema de una red. En (a) se ilustra una red dirigida y en (b) una no dirigida. En el caso de las redes no dirigidas, las conexiones entre sus vértices se denominan aristas o simplemente conexiones, mientras que en las redes dirigidas se les llaman *arcos*.

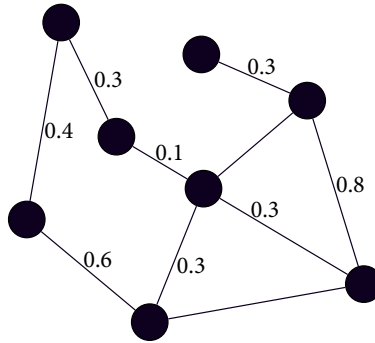


Figura 2: Esquema de una red con pesos en las conexiones.

Toda estructura de una red presenta una manera particular en la que las conexiones se distribuyen entre los nodos, a esto se le conoce como topología de la red. La topología suele representarse mediante la matriz de adyacencia  $A$ , donde  $a_{ij} = 1$  si los nodos  $i, j$  están conectados y  $a_{ij} = 0$  de lo contrario. Una forma de analizar la topología es mediante la

distribución de probabilidad del grado  $P(k)$ , donde  $k$  o “grado” de un nodo se define como el número de conexiones adyacentes a él. Por lo tanto  $P(k)$  se refiere a la probabilidad de que un nodo tomado al azar, tenga  $k$  conexiones.

Existe un gran interés en estudiar el efecto que tienen distintas topologías en la dinámica y el funcionamiento de los sistemas, en particular en las redes libres de escala (RLE). Las RLE se caracterizan porque su distribución del grado sigue una ley de potencias  $P(k) \sim k^{-\gamma}$ , donde  $\gamma$  es un parámetro cuyo valor suele estar entre 2 y 3 (Albert y Barabási, 2002); esto implica que la red posee un número relativamente grande de nodos con pocas conexiones y pocos nodos altamente conectados a los que se les denomina *hubs* (Albert y Barabási, 2002) (Figura 3). Las RLE han llamado la atención ya que muchos sistemas naturales tanto biológicos como sociales e incluso tecnológicos parecieran ser de este tipo (Albert y Barabási, 2002). Entre los sistemas biológicos esto incluye a las redes de interacción proteína-proteína y las redes de regulación genética (Albert y Barabási, 2002), por mencionar algunas.

En el caso de las redes completas, donde todos los nodos están conectados entre sí, la topología no es relevante ya que todos los nodos tienen el mismo número de aristas. En este caso la estructura está determinada únicamente por los pesos de las conexiones.

### **Dinámica.**

La dinámica se refiere al cambio o evolución de un sistema a través del tiempo. En una red el cambio puede observarse a dos niveles: en el estado de activación de las unidades de procesamiento o bien en la propia estructura, modificando la topología y los pesos. Los cambios del primer tipo son abordados por los modelos de redes dinámicas con topologías estáticas, mientras que los del segundo tipo son abordados por los modelos que se conocen como de “redes dinámicas evolutivas” (Belykh *et al.*, 2014). Aunque un enfoque verdaderamente integral para el estudio de los sistemas desde una perspectiva de redes requiere

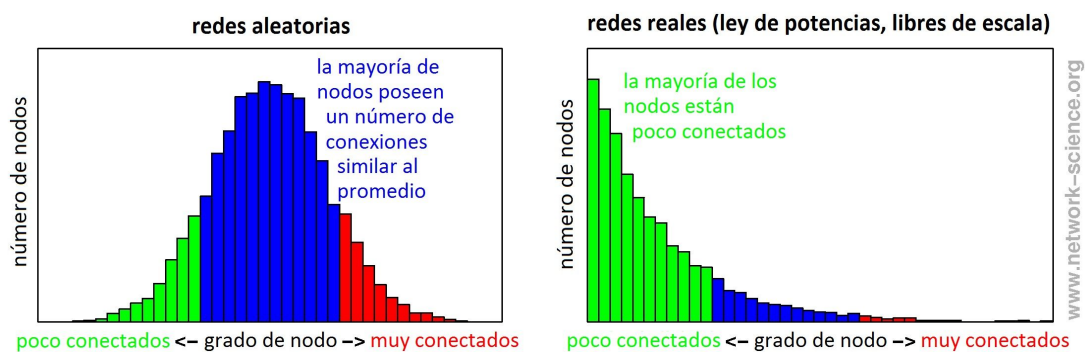


Figura 3: Diagrama de la distribución del grado de una red aleatoria (izquierda) y una red libre de escala (derecha) (modificado de *Node degree distribution. Random networks versus scale-free power-law networks* [Imagen digital]. s.f. Recuperado el 22 de septiembre, 2017, de <http://www.network-science.org/>).

la incorporación de ambos procesos, las redes dinámicas evolutivas quedan fuera de esta discusión. En lo subsecuente se hablará únicamente de las redes dinámicas con topologías estáticas, particularmente de las redes discretas ya que en este trabajo se empleará un modelo de este tipo.

Las redes dinámicas discretas suelen tener topologías y pesos fijos, donde la dinámica no ocurre en la estructura sino en el estado de activación de los elementos del sistema. Cada elemento o “unidad de procesamiento” de la red expresa un valor en un momento determinado, ya sea 1 si se encuentra activa ó 0 si se encuentra inactiva (o bien, 1 y  $-1$  respectivamente). El patrón de ceros y unos de todas las unidades en conjunto se traduce en una cadena de bits que representa el estado de la red en el tiempo  $t$ . Al tiempo  $t + 1$  cada unidad de procesamiento actualiza su estado siguiendo una función recursiva denominada *función de activación*<sup>1</sup>, que puede ser diferente para cada unidad o la misma para todas y que toma en cuenta el estado actual de la red en el tiempo  $t$ . De esta forma el tiempo procede de manera discreta generando una secuencia de estados a la que se le llama *trayectoria* (Wuensche, 2004) (Figura 4). Por lo tanto definiremos a la dinámica de

<sup>1</sup>También llamada *función de cambio* o *regla dinámica de activación*.

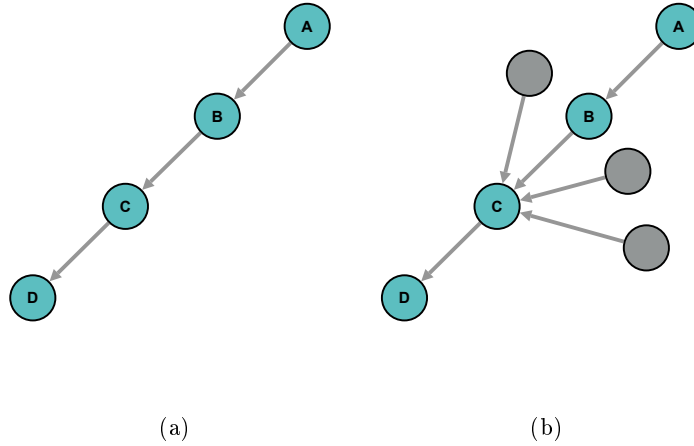


Figura 4: Trayectoria de una red dinámica. A, B y C representan distintos estados del sistema (determinados por cadenas de bits) y las flechas indican la dirección de la trayectoria. (a) Se considera que A conduce a B y B conduce a C, si el cambio en un bit en el estado A produce el estado B y así sucesivamente. (b) En color gris se ilustra el conjunto de estados distintos de B que también conducen a C y a los cuales, junto con B, se les conoce como “preimágenes” de C.

una red como una sucesión de estados que depende del estado anterior del sistema y que es determinada por una función de activación o regla dinámica interna a cada unidad.

Una característica importante de las redes dinámicas es que en muchas de ellas, la trayectoria eventualmente se estabilizará y convergerá a un atractor (Wuensche, 2004; Goudarzi *et al.*, 2012) o sumidero (Chaos, 2010), el cual puede ser periódico o caótico. Si el atractor es periódico de período uno se trata de un estado estacionario al que se le llama punto fijo y si el período es mayor a uno se trata de un ciclo límite (Figura 5); por otro lado, si el atractor es caótico se dice que el sistema presenta un atractor extraño (Wuensche, 2004).

Los atractores de una red dinámica pueden tener múltiples significados dependiendo el área en la que se apliquen. En biología muchas veces se toman como representaciones de perfiles de expresión genética que mapean con fenotipos específicos en estudios de la

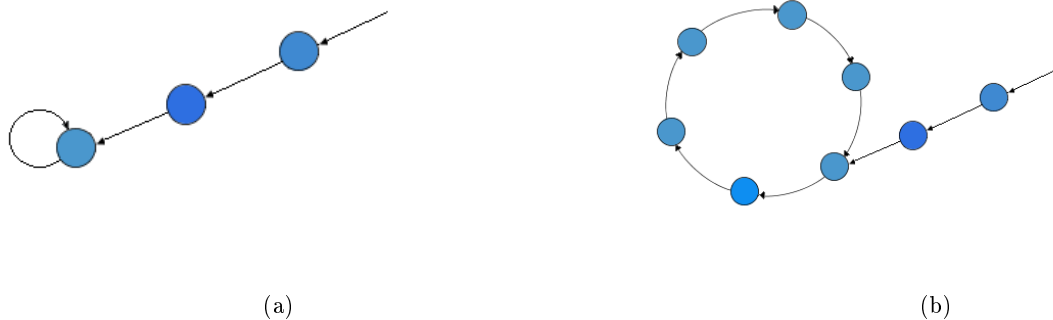


Figura 5: Atractores cíclicos. En (a) se ilustra una trayectoria con un atractor de período 1 (un punto fijo) y en (b) una trayectoria con un atractor de período seis (ciclo límite).

regulación genética, o bien como patrones de activación neuronal en el cerebro cuando se aborda la memoria y el aprendizaje (Hopfield, 1982; Geard y Wiles, 2005; Perotti *et al.*, 2006; Balleza *et al.*, 2008; Karlebach y Shamir, 2008; Dragui y Wagner, 2009; Barrio *et al.*, 2010; Chaos, 2010; Watson *et al.*, 2010; Fernández *et al.*, 2013; Watson *et al.*, 2014; Watson y Szathmáry, 2016).

## **Función.**

Otro aspecto ligado a los sistemas, en particular a los biológicos, es la función. En el contexto de la biología de sistemas es importante conocer la relación entre la estructura y la dinámica de una red con el fin de entender el origen de la función biológica, ya que la función liga la estructura con un comportamiento a nivel de sistema (Knüpfer y Beckstein, 2013).

Hablar de función en biología siempre ha sido motivo de debate, sobre todo en el ámbito de la biología evolutiva, debido a que a la función se le suele vincular con una visión

teleológica de la Naturaleza (Abbot, 1916); sin embargo algunos consideran que hablar en términos de función es inevitable e incluso necesario como parte de un proceso de búsqueda de explicaciones causales a procesos biológicos debido a que establece paralelismos entre fenómenos (Abbot, 1916; Breitenbach, 2009). Más aún, se cree que el hecho de que podamos reconocer funciones habla de nuestra tendencia a reconocer que en la Naturaleza no sólo existen cosas sino también tipos de cosas o clases, cuya existencia como realidades objetivas es a su vez determinante de la capacidad de los individuos de formar conceptos (Abbot, 1916).

Ludwig von Bertalanffy, el fundador de la teoría general de sistemas, consideraba que por tratarse de entidades conformadas por unidades interconectadas, los sistemas debían compartir principios generales en cuanto a organización y control se refiere. A estos principios generales o “paralelismos” a nivel de sistema, von Bertalanffy los llamó isomorfismos (von Bertalanffy, 1976). En este sentido, cuando se dice que la función del corazón es bombear sangre a los pulmones no quiere decir que el corazón haya sido diseñado para esa finalidad por una intencionalidad genuina, sino que reconocemos en su organización las relaciones que deben darse entre las partes para que funcione como una bomba. Ambas estructuras son pues isomórficas.

Consciente del problema que genera hablar en términos de función, von Bertalanffy marcó una distinción entre tipos de finalidad, a los que llamó: *teleología estática*, que se centra en la relación fin-medio, significando que una disposición parece útil para determinado "propósito" y *teleología dinámica*, que se centra en la relación causa-efecto, atribuyendo la función a una directividad de procesos en los que la estructura determina un resultado (von Bertalanffy, 1976). Él consideraba que la interpretación de función que se busca en ciencia es la del tipo causal, no obstante, los isomorfismos en general cuando se establecen cuidadosamente, son útiles para entender fenómenos naturales de los que se sabe poco, ya

que pueden ser expresados en términos de otros procesos más sencillos y mejor conocidos (von Bertalanffy, 1976).

Entendiendo a la función como una teleología dinámica resulta más o menos claro que para que ésta tenga lugar debe establecerse un mapeo entre dos aspectos de un sistema: la estructura y el comportamiento (Knüpfer y Beckstein, 2013). Dicho mapeo se construye a lo largo del tiempo evolutivo del sistema y debe ser en gran medida el origen de sus restricciones (Wagner y Altenberg, 1996). De hecho desde tiempo atrás ya se hablaba de que la función de alguna manera debía determinar a la estructura, justo como a su vez la estructura debía definir límites sobre la función (Abbot, 1916).

Los modelos de redes dinámicas podrían ser útiles para estudiar la relación entre estructura y función, ya que precisamente se tratan de representaciones de la estructura de un sistema que mapea con ciertos comportamientos. En este sentido, los atractores de una red dinámica serían equiparables con la finalidad del sistema según von Bertalanffy ya que es la estructura misma la que determina su resultado.

### **El potencial evolutivo.**

El potencial evolutivo (del inglés "evolvability"), ha sido tema central de múltiples publicaciones en los últimos treinta años (Pigliucci, 2008). El término comienza a aparecer en el ámbito de la biología a principios de 1990, sin embargo el concepto es más viejo. Según Ingo Brigandt éste se remonta a los ochenta, década en la que nace la Evo-Devo como una alternativa al paradigma evolutivo (Brigandt, 2015). Si bien el uso del término potencial evolutivo no era común en aquel entonces ya comenzaba a hablarse de otros términos similares como el cambio evolutivo facilitado y la oportunidad evolutiva, ambos ligados originalmente al concepto de restricciones del desarrollo (Brigandt, 2015).

Actualmente no existe una definición unívoca de lo que es el potencial evolutivo. El término ha sido usado para referirse a distintos fenómenos a varias escalas, por lo cual engloba un amplio espectro de conceptos que se sobrelapan (Pigliucci, 2008). Pero de todos estos conceptos, Massimo Pigliucci identifica tres grupos principales que conforman tres definiciones o formas de concebir el potencial evolutivo, que son: conceptos relacionados con la heredabilidad (potencial evolutivo *sensu* Houle), conceptos relacionados con la variabilidad (potencial evolutivo *sensu* Wagner y Altenberg) y conceptos relacionados con las principales transiciones evolutivas (potencial evolutivo *sensu* Maynard-Smith y Eörs Szathmáry), los cuales se describen a continuación.

- **Potencial evolutivo *sensu* Houle:** Lo conforman conceptos análogos a la heredabilidad, una medida de qué tanto la variación fenotípica de un carácter es causada por la variación genotípica en la población (Wray y Visscher, 2008). Dado que la variación se considera una medida del potencial de la población para responder a la selección natural, muchos han tomado a la heredabilidad como una forma de cuantificar el potencial evolutivo. David Houle (1992) fue uno de los primeros en hablar del potencial evolutivo en términos de la heredabilidad, razón por la cual Pigliucci nombra a este grupo de conceptos como potencial evolutivo *sensu* Houle (Pigliucci, 2008). En realidad Houle considera que el coeficiente de variación es una medida más adecuada para medir el potencial evolutivo que la heredabilidad ya que permite comparar múltiples caracteres y de distintos tipos, ya sean morfológicos o historias de vida (Houle, 1992).
- **Potencial evolutivo *sensu* Wagner y Altenberg:** Este grupo de conceptos parte de la distinción entre variación y variabilidad. De acuerdo con Günter P. Wagner y



Lee Altenberg "*la variación es una propiedad de una colección de objetos que puede ser directamente observada*", mientras que la "*variabilidad describe una propensión [o potencial] de variar*" (Wagner y Altenberg, 1996). La variación puede ser medida a través de la heredabilidad (potencial evolutivo *sensu* Houle) pero la variabilidad, en cambio, requiere de herramientas distintas, que reflejen la estructura genética y del desarrollo, y que tomen en cuenta el mapeo entre el genotipo y el fenotipo (Pigliucci, 2008). En este caso el potencial evolutivo se entiende como cualquier propiedad o conjunto de propiedades de la arquitectura genética y de las restricciones del desarrollo que facilitan la evolución fenotípica.

- **Potencial evolutivo *sensu* Maynard-Smith y Eörs Szathmáry:** El tercer grupo de conceptos está relacionado con la innovación y con lo que John Maynard-Smith y Eörs Szathmáry llaman "las principales transiciones evolutivas" (*The major transitions in evolution*) (Pigliucci, 2008). En este contexto, el potencial evolutivo figura como una capacidad de las especies de superar sus restricciones, pudiendo generar cambios fenotípicos importantes.

Estos tres grupos de conceptos o definiciones muestran la importancia que se le da a la selección y las restricciones a varias escalas en la construcción de entidades de mayor nivel de organización. La razón por la que el potencial evolutivo abarca un amplio espectro de conceptos se debe a que es en realidad el resultado de una composición de propiedades del sistema que en conjunto permiten el acoplamiento y la transición al cambio de una manera organizada, involucrando procesos como la regulación, la compartimentalización y los mecanismos exploratorios (Kirschner y Gerhart, 1998). Algunos han llamado a estas propiedades como "facilitadoras de cambio" (West-Eberhard en Pigliucci, 2008) o "liber-

adoras de restricciones" (Kirschner y Gerhart, 1998).

Pero a los tres grupos de conceptos anteriores podemos sumarles un cuarto grupo: el de los conceptos relacionados con el aprendizaje, el cual conforma lo que llamaremos "definición de potencial evolutivo sensu Valiant".

En 1984 Leslie G. Valiant propuso una metodología para estudiar el aprendizaje desde una perspectiva computacional, en donde el aprendizaje es considerado como un fenómeno de adquisición de conocimiento en ausencia de un algoritmo explícito (Valiant, 1984). Su metodología, denominada aprendizaje PAC (*Probably Approximately Correct*) consiste en una máquina que aprende a formar conceptos, es decir programas para relizar una tarea, únicamente a partir de ejemplos que recibe de su entorno y que no proveen una descripción explícita de dicho programa. Valiant ve a la evolución biológica como una forma restringida de aprendizaje PAC en donde el sistema que aprende lo hace a partir de ejemplos que son obtenidos de una distribución de probabilidad determinada por la disponibilidad de dichos ejemplos en la naturaleza y cuyo "objetivo" es la búsqueda de funciones que sean, en sí mismas, capaces de evolucionar (Valiant, 2009).

De acuerdo con Valiant para que el aprendizaje tenga lugar es necesario que el sistema cuente con una manera de reunir información y una forma de explorar la clase de conceptos que son aprendidos (Valiant, 2009). Pero para que un sistema aprenda es necesario además que su desempeño en alguna tarea mejore con la experiencia y que sea capaz de extrapolar o generalizar a partir de dicha experiencia (Watson *et al.*, 2010). Lo que es realmente interesante es que los sistemas biológicos podrían presentar estas mismas características.

Recientemente se ha demostrado que redes de regulación genética con interacciones recurrentes muy similares a las redes de Hopfield, de las cuales se hablará en el tema siguiente, forman memorias distribuídas de su desarrollo, pudiendo almacenar y recuperar múltiples fenotipos que han sido seleccionados en el pasado a partir de "fenotipos embronarios"

degradados o incompletos (Watson *et al.*, 2010, 2014; Watson y Szathmáry, 2016). Además, bajo ciertas condiciones, estas redes de regulación genética son capaces de generalizar para producir nuevas combinaciones de fenotipos (Watson *et al.*, 2010, 2014).

El hecho de que modelos conexionistas del aprendizaje como las redes neuronales presenten similitudes con otros sistemas biológicos como las redes de regulación genética, no sólo permite el uso en biología de herramientas provenientes de áreas como las ciencias cognitivas, sino que además amplía la forma en la que pensamos el proceso evolutivo y su potencial (Watson y Szathmáry, 2016). En las siguientes subsecciones se hablará de las redes neuronales, en particular el modelo de Hopfield, y de las similitudes que éste presenta con sistemas biológicos como las redes de regulación genética.

## **La red de Hopfield.**

Las ReNA son sistemas dinámicos formados por un conjunto de unidades procesadoras simples denominadas neuronas que se comunican mediante señales a través de conexiones con pesos (Kröse y van der Smagt, 1996). La manera como la información es procesada y almacenada en una ReNA le brinda la capacidad de aprender y generar comportamientos autoorganizados.

Uno de los modelos de red neuronal más conocidos y estudiados fue creado por John Hopfield en 1982 (Hopfield, 1982), en él Hopfield desarrolló varias ideas previas concernientes a este tipo de redes y presentó un análisis matemático detallado de su funcionamiento (Kröse y van der Smagt, 1996), razón por la cual a este tipo de redes se les conoce con su nombre (Beale *et al.*, 1990).

La red de Hopfield consiste en un número  $N$  de neuronas todas interconectadas (Figura

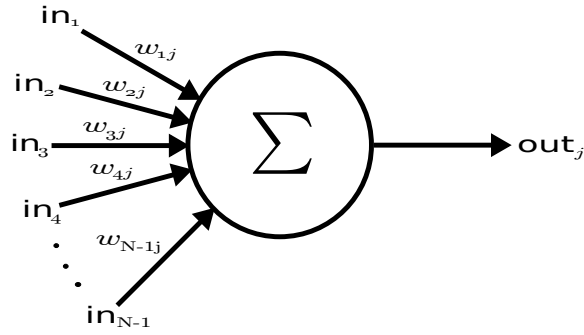


Figura 6: Esquema de una neurona matemática. Cada neurona integra las  $N - 1$  entradas que recibe del resto de las neuronas y así emite una salida. La salida  $out_j$  representa el estado de activación de la neurona  $j$  y corresponde a la suma ponderada de los valores de entrada  $in_i$ , donde  $i = 1, \dots, N - 1$ .

6). La fuerza de la conexión entre una neurona  $i$  y una neurona  $j$  se define mediante el peso  $w_{ij}$  (neuronas no conectadas tienen  $w_{ij} = 0$ ) (Hopfield, 1982). Debido a que todos sus nodos están conectados se trata de una red completa, además de ser simétrica ya que las conexiones de una neurona a otra son iguales en ambas direcciones, es decir que  $w_{ij} = w_{ji}, \forall i \neq j$  (Figura 7).

Como en el caso de otras redes dinámicas, cada neurona  $i$  puede tomar dos estados:  $\sigma_i = 0$  (inactiva) y  $\sigma_i = 1$  (activa), o lo que es equivalente,  $\sigma_i = -1$  y  $\sigma_i = 1$ . El estado instantáneo del sistema está especificado por la lista de los  $N$  valores de  $\sigma_i$ , de forma que se puede representar por un vector binario (o bipolar en caso de tratarse de los valores  $\{-1,1\}$ ) de  $N$  bits (Hopfield, 1982).

La dinámica del sistema cambia en el tiempo de acuerdo a la siguiente función de activación:

$$\sigma_i(t + 1) = \varphi\left[\sum_{j=1}^N w_{ij}\sigma_j(t)\right],$$

donde  $\varphi$  es la función escalón, también llamada de Heaviside

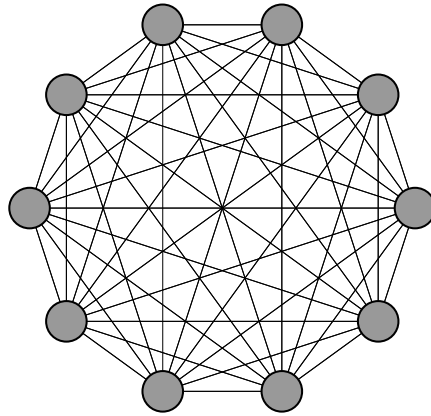


Figura 7: Estructura de una red de Hopfield de diez unidades. Se trata de una red completamente conectada, en donde cada unidad se encuentra conectada con el resto, excepto consigo misma y es simétrica porque no existe una direccionalidad en sus conexiones, representado por medio de aristas sin flecha.

$$\varphi = \begin{cases} 1 & \text{si } x \geq \theta_i \\ 0 & \text{si } x < \theta_i \end{cases}$$

y  $\theta_i$  es el umbral de activación para la neurona  $i$ , el cuál suele tomarse como  $\theta_i = 0$  (Hopfield, 1982).

Cada neurona  $i$  actualiza su estado de manera aleatoria y asíncrona, evaluando si está por arriba o por abajo del umbral y reajusta su estado en correspondencia.

Una de las características que distingue a la red de Hopfield de otros modelos de redes neuronales es que al ser una red completamente conectada, las señales de entrada o *inputs* (que consisten en un vector inicial de valores binarios o bipolares) son aplicadas sobre todos los nodos al mismo tiempo. Las conexiones simétricas permiten que la información circule en ambos sentidos. En una red de Hopfield la primera salida generada o *output* es tomada como un nuevo *input*, el cual produce a su vez un nuevo *output* y así sucesivamente hasta

que la red converge a un atractor (Beale *et al.*, 1990). El *output* final de la red o “solución” se toma como el valor de todos los nodos cuando la red ha alcanzado un estado estacionario.

En una red de Hopfield es posible guardar información de los patrones en la estructura de pesos de la red, haciendo que estos patrones coincidan con los atractores del sistema (Kröse y van der Smagt, 1996). Usando los mismos términos que Hopfield, considérese el vector  $X_a$  y supóngase que es un estado estable del sistema. Si el sistema es iniciado suficientemente cerca de  $X_a$ , digamos  $X = X_a + \Delta$ , éste iterará hasta que  $X \approx X_a$ . Así pues, podemos decir que el sistema posee información almacenada de  $X_a$  que le permite recuperar dicho vector (Hopfield, 1982). Las redes de Hopfield son capaces de recuperar patrones partiendo de información incompleta proveniente del entorno y completándola con información almacenada en su estructura. Por esta razón, la estructura actúa como una memoria asociativa. En este sentido, cuando a la red se le presenta un patrón desconocido como entrada, ésta responderá produciendo el patrón en memoria que más se le parezca (Figuras 8a y 8b). En ocasiones la red se estabiliza en estados distintos a cualquiera de los patrones aprendidos a los cuales se les denomina “estados espurios”.

En el modelo original de Hopfield la memorización de los patrones se logra asociando un peso  $w_{ij}$  a cada conexión de acuerdo a la regla Hebbiana de aprendizaje. En el contexto de las redes neuronales, la regla Hebbiana es a menudo representada bajo la frase "neuronas que disparan juntas se conectan juntas", lo que quiere decir que las conexiones entre neuronas que tienen una correlación en su activación ante un estímulo son reforzadas (Watson *et al.*, 2010). Su fórmula es la siguiente (para una red binaria con estados 0 y 1):

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^P (2\xi_i^\mu - 1)(2\xi_j^\mu - 1),$$

donde  $\xi_i^\mu$ , es el elemento  $i$  del patrón  $\mu$ , con  $i = 1, \dots, N$  y  $\mu = 1, \dots, P$ .

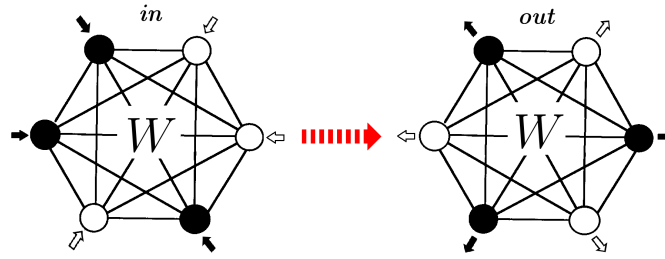
La regla Hebbiana tiene dos características importantes para el aprendizaje en redes neuronales: 1) es local, lo que significa que la modificación de los pesos depende únicamente de las neuronas conectadas y no de toda la red en conjunto y 2) es incremental o creciente, ya que los nuevos patrones pueden ser aprendidos sin tener que usar la información de los patrones previamente utilizados para el aprendizaje. Por lo tanto la actualización de los pesos depende únicamente de los valores anteriores y del nuevo patrón. El carácter local le permite al sistema establecer paralelismos entre patrones mientras que el carácter incremental le permite aprender en un ambiente cambiante (Perumal, 2008).

La red de Hopfield posee un valor escalar  $E$  asociado con cada estado, al que se le conoce como energía, y que permite relacionar a la red con otro tipo de sistemas físicos (Torres *et al.*, 2010). Se le llama energía a  $E$  porque ésta disminuye a medida que la dinámica de la red avanza o bien permanece igual cuando el sistema ha alcanzado un atractor, el cual puede ser visto como un mínimo local de la función de energía (Kröse y van der Smagt, 1996) (Figura 8c).

En resumen el algoritmo general de operación de una red de Hopfield es el siguiente:

1. Asignar los pesos entre conexiones mediante la regla Hebbiana de aprendizaje.
2. Inicializar la red con un patrón desconocido como entrada.
3. Actualizar el estado de las neuronas mediante la función de activación hasta que la red converja a un atractor.

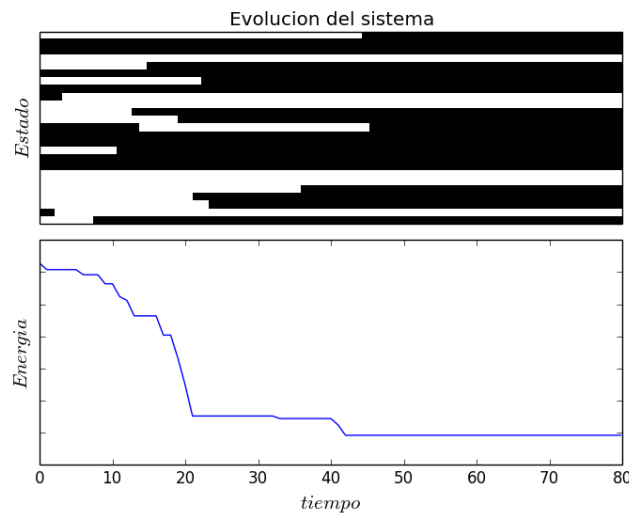
En este tipo de redes los atractores siempre son puntos fijos a menos que se modifiquen aspectos de su estructura como que el hecho de que los pesos no sean simétricos o de la proporción de pesos  $w_{ij} = 0$  (Weaver, 1999).



(a)



(b)



(c)

Figura 8: Funcionamiento de una red de Hopfield. Por tratarse de una red de una sola capa, todas las unidades actúan como entrada y como salida (a). A través de un número de iteraciones, las redes de Hopfield son capaces de recuperar patrones a partir de información incompleta proveniente de su entorno y completándola con información almacenada en su matriz de pesos  $W$  (a) y (b). La trayectoria de la red hacia la estabilidad puede verse como un proceso de disminución de energía a medida que la dinámica de la red avanza (c). (a) modificado de Hopfield Net. [Imagen digital]. Recuperado el 23 de septiembre de 2017, de: <http://www.nnwj.de/hopfield-net.html>; (b) tomado de Beale, R., Jackson, T., & Hilger, A. (1990). Neural Computing : An Introduction. Bristol: Adam Hilger.



Las características que posee el modelo de Hopfield han hecho que sea empleado como modelo del funcionamiento de las redes neuronales biológicas y del aprendizaje en general, sin embargo, como se ha venido mencionando, se cree que sus principios podrían generalizarse a otros procesos biológicos distintos de los neuronales como la regulación genética y la evolución (Wuensche, 2010; Watson y Szathmáry, 2016).

### **Similitudes entre redes de regulación genética y redes de Hopfield.**

Las redes de regulación genética (RRG) regulan la expresión de miles de genes en cada proceso del desarrollo, especificando qué conjunto de genes deben ser expresados en tiempo y espacio específicos (Levine y Davidson, 2005). Una RRG consiste en cientos de secuencias modulares de DNA que reciben e integran múltiples señales activadoras y represoras, a partir de moléculas que reconocen secuencias específicas dentro de éstas. El resultado final es el control preciso de la transcripción de los genes asociados.

Los enfoques para modelar RRG son varios, van desde los modelos matemáticos detallados que se apegan a los datos experimentales, hasta los modelos más abstractos, comunes en campos como los sistemas complejos o la vida artificial (Geard y Wiles, 2005). Mientras que los primeros son empleados a menudo para hacer predicciones cuantitativas, los segundos sirven para investigar aspectos más cualitativos de la regulación (Geard y Wiles, 2005).

Una RRG puede describirse en términos de una red abstracta en donde los nodos representan genes y factores de transcripción y las conexiones representan interacciones regulatorias, ya sean transcripcionales, postranscripcionales o postraduccionales (Aravind *et al.*, 2009). En general una RRG puede modelarse como una red dinámica de modo que los conceptos de atractor, trayectoria y estado, mencionados anteriormente, pueden aplicarse

a este tipo de redes. El estado de la RRG en cualquier momento está determinado por el patrón de activación de los genes individuales (Geard y Wiles, 2005). En los modelos discretos 1 suele ser el valor de un gen que se encuentra activo (o por encima de un umbral de activación) y 0 el valor de un gen inactivo (o por debajo de un umbral de activación).

Las redes de Hopfield han sido empleadas para modelar RRG (Dragui y Wagner, 2009; Watson *et al.*, 2010), en particular en el estudio del cáncer (Maetschke y Ragan, 2014; Szedlak *et al.*, 2014; Fard *et al.*, 2016). En una RRG modelada como una red de Hopfield, cada tipo celular o linaje puede ser visto como uno de los patrones memorizados. Así como una red de Hopfield contiene información parcial en su estructura que le permite recuperar los patrones de activación almacenados en memoria, una célula contiene información parcial para producir distintos tipos celulares a lo largo del desarrollo. De esta manera, la estructura de la RRG actúa como una memoria capaz de organizar los componentes del sistema y de recuperar estados fenotípicos pasados (Wuensche, 2004; Watson *et al.*, 2010, 2011a, 2014).

En cuanto a la naturaleza de los pesos, en el modelo clásico de Hopfield la fuerza de la interacción entre dos unidades  $i, j$  se obtiene mediante la regla Hebbiana de aprendizaje. Establecer una relación directa entre la red de Hopfield y las RRG implicaría que éstas almacenaran información de forma similar. Watson y colaboradores aseguran que las RRG pueden estar sujetas a modificaciones en sus conexiones que en principio son las mismas producidas por el aprendizaje Hebbiano (Watson *et al.*, 2010). Ellos citan el trabajo de Wagner *et al.* (2007) en el que mencionan que la selección puede actuar en genes que afectan la correlación de ciertos caracteres, ya sea incrementando dicha correlación o reduciéndola, dependiendo de si los caracteres se encuentran simultáneamente bajo selección direccional o no, proceso que a la larga resultaría en un reforzamiento de los efectos pleiotrópicos entre caracteres que son seleccionados juntos y en la supresión de los efectos pleiotrópicos de los que no (Wagner *et al.*, 2007 en Watson *et al.* (2010)). Watson y colaboradores señalan que

este comportamiento claramente describe una modificación Hebbiana de las interacciones genéticas (Watson *et al.*, 2010). Dichas modificaciones en las conexiones hacen que el patrón de expresión seleccionado sea más robusto a la variación ambiental y a la larga este proceso causaría que el sistema regulatorio en conjunto se comporte como una memoria asociativa de fenotipos que han sido visitados y seleccionados en el pasado (Watson *et al.*, 2010). La extrapolación de la noción Hebbiana de las neuronas hacia los genes, la cual podríamos reformular como “genes que se activan juntos se conectan juntos”, simplemente sugiere que las causas que dan origen a tales correlaciones entre unidades se mantienen, de manera que estímulos recurrentes generarían estructuras recurrentes.

## Sección 2: Objetivos.

Dadas las similitudes que existen entre el aprendizaje y la evolución, los modelos de aprendizaje como las ReNA pueden brindarnos información acerca de la biología. El objetivo de este trabajo es explorar algunas de las ideas que este tipo de modelos pueden proporcionar. Para ello se empleó una red de Hopfield como modelo de un sistema de regulación genética hipotético. Aunque la red de Hopfield es un modelo muy simplificado de una red de regulación genética (RRG), nos da la posibilidad de estudiar a este tipo de sistemas en términos de su estructura, memoria y mecanismos de adquisición de información.

Los objetivos particulares son:

- Analizar cómo es que cambios en la estructura de una red de Hopfield afectan su función en términos de la estabilidad de los atractores.
- Comparar los resultados obtenidos por las simulaciones con datos biológicos reales y discutir las implicaciones biológicas que podrían derivarse del modelo.

- Discutir la relación que existe entre la estructura y el potencial evolutivo.

### Sección 3: Método.

#### El modelo

Se crearon redes de Hopfield de tamaño  $N = 10, 50, 100, 500$  y  $1,000$  unidades completamente conectadas, con conexiones simétricas y pesos simétricos. Al tratarse de redes completas, simétricas y sin conexiones autodirigidas, el número total de aristas es  $M = (N^2 - N)/2$ . Las redes poseen un estado global  $S$  determinado por el estado local  $\sigma_i$  de cada una de sus unidades, el cual puede ser 0 ó 1 (ó  $\{-1, 1\}$ , ver subsección *Topología* más adelante), es decir  $S = \{\sigma_i \in \{0, 1\}, i = 1, \dots, N\}$ . En el caso de una RRG un estado  $\sigma_i = 0$  representaría un gen inactivo o por debajo de un umbral de activación mientras que un estado  $\sigma_i = 1$  representaría un gen activo o por encima de un umbral de activación.

Las redes actualizan su estado en el tiempo siguiendo la función de activación

$$\sigma_i(t+1) = \varphi\left[\sum_{j=1}^N w_{ij}\sigma_j(t)\right],$$

donde

$$\varphi = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Se entrenó a las redes para que memorizaran  $P$  patrones independientes que representan perfiles de expresión genética donde cada variable  $\xi_i^\mu$  tiene una probabilidad  $p(\xi_i^\mu = \{0, 1\}) = 1/2$ . La memorización de los patrones se hizo mediante la regla Hebiana de aprendizaje  $w_{ij} = \frac{1}{N} \sum_{\mu=1}^P (2\xi_i^\mu - 1)(2\xi_j^\mu - 1)$  (Hopfield, 1982).

## Evaluación de la estabilidad de los atractores.

Para evaluar cómo los cambios en la estructura afectan la estabilidad de los atractores de la red se empleó la función de traslape, definida aquí como:

$$m_S^\mu = \frac{1}{N} \sum_{i=1}^N (2\xi_i^\mu - 1)(2\sigma_i - 1).$$

La función de traslape es frecuentemente usada en el ámbito de las redes neuronales para evaluar el desempeño de la red ya que establece una distancia entre el patrón  $\xi^\mu$  que debería recuperar y el estado  $S$  de la red (Elizalde *et al.*, 1992; Löwe, 1998; McGraw y Menzinger, 2003; Stauffer *et al.*, 2003; Davey *et al.*, 2004; Kim, 2004; Morelli *et al.*, 2004; Torres *et al.*, 2004; Dominguez *et al.*, 2007; Löwe y Vermet, 2015).

Dado que los patrones de entrada consistieron todo el tiempo en cada uno de los patrones almacenados en memoria, sin alguna perturbación, la red debería permanecer estacionaria en el patrón que le fue presentado ya que se trata de un patrón memorizado y por lo tanto un punto estable del sistema. De esta forma el traslape es usado aquí como una medida de la estabilidad de los atractores y por lo tanto de la función de la red. En caso de que la red converja a un estado diferente al que le fue presentado el traslape sería menor a 1, esto querría decir que ese estado ya no es un atractor o es menos estable, a pesar de haber sido establecido mediante la regla Hebbiana de aprendizaje.

En todos los experimentos se reporta el traslape global o promedio de la red, es decir, la suma del traslape de cada uno los patrones en memoria dividido entre el total de patrones memorizados por la red. Si la red tuvo un traslape igual a 1 en cada uno de los patrones memorizados, entonces su traslape global será  $m = 1$ .

## **Modificación de la estructura.**

La estabilidad de los atractores se evaluó en función de tres parámetros: 1) el incremento en la carga de patrones almacenados, 2) la información mutua entre los patrones y 3) el cambio en la topología de la red. Estos tres parámetros son importantes ya que afectan la estructura de la red al incidir directamente sobre la matriz de pesos  $W$ .

Aunque ya se ha observado antes el efecto que tiene el cambio en estos parámetros en el desempeño de la red, la intención de este trabajo es reproducir dichos comportamientos y discutir sus posibles implicaciones biológicas fuera del ámbito de las redes neuronales. La carga, por ejemplo, está relacionada con el número de tipos celulares estables que la red es capaz de “recordar, mientras que la información mutua puede usarse como una medida del grado de correlación entre los perfiles de expresión que corresponden con cada tipo celular. La topología, por otro lado, sería una medida de la geometría de la RRG, establecida por las interacciones físicas entre los componentes. A continuación se explican con mayor detalle cada uno de estos parámetros así como su implementación.

### **Carga.**

La carga  $\alpha$  se define como la fracción que representa el total de patrones memorizados con respecto al tamaño de la red, esto es  $\alpha = P/N$ . Es sabido que conforme la carga aumenta el traslape disminuye (Rojas, 1996; McGraw y Menzinger, 2003; Torres *et al.*, 2010; Dominguez *et al.*, 2007). Para saber cómo es que esto ocurre se obtuvo el traslape de las redes a medida que se incrementaba la carga.

Se comenzó siempre con un solo patrón almacenado en la memoria de modo que  $P_o = 1$  y  $\alpha_o = 1/N$ . En cada iteración se incrementó en +1 el número de patrones, es decir

$\alpha_{t+1} = \frac{1}{N}(P_t + 1)$ , hasta alcanzar un total de  $P = N$  patrones y  $\alpha = 1$ . Este proceso se repitió veinte veces para cada valor de  $N$ , excepto en el caso donde  $N = 1,000$ , que solamente constó de 10 repeticiones.

La estabilidad de los atractores en función de la carga se midió a partir de un conjunto memoria con  $P$  patrones independientes donde cada variable  $\xi_i^\mu = \{0, 1\}$  es aleatoria y con una probabilidad  $p(\xi_i^\mu = \{0, 1\}) = 1/2$ .

Uno de los valores que se desea determinar es el número máximo de patrones  $P_{max}$  que las redes relativamente pequeñas con las que se trabajó pueden almacenar antes de ser inestables, para así calcular la carga crítica  $\alpha_c$  a partir de la cual ocurre la inestabilidad y compararla con los resultados que hayan sido reportados previamente.

Debido a que en todo momento se obtuvo la estabilidad global de redes con varios patrones memorizados, se consideró una cantidad arbitraria de error de recuperación de los patrones del 1%. Esto quiere decir que se definió  $\alpha_c$  como el valor de carga a partir del cual el traslape global de la red es menor a 0.99.

### **Información mutua entre patrones.**

En los experimentos anteriores la estabilidad de la red en función de la carga se evaluó a partir de un conjunto memoria de patrones independientes, pero se sabe que el grado de similitud entre el conjunto de patrones memorizados tienen un efecto en la estabilidad de la red (Rojas, 1996). En este experimento se buscó evaluar la estabilidad en función de la similitud de los patrones usando una medida de correlación. Como medida de correlación se empleó la información mutua entre los patrones memorizados.

La información mutua  $IM$  entre dos variables aleatorias  $X$  e  $Y$ , establece una medida

de dependencia entre ambas variables. En este caso, las variables  $X$  e  $Y$  corresponden a dos patrones cualesquiera  $\xi^a$  y  $\xi^b$  del conjunto memoria. La  $IM$  entre ambos patrones nos indica qué tanto conocer el estado de un patrón reduce nuestra incertidumbre del otro (Latham y Roudi, 2009). Dicho de otro modo, si conocer el estado del patrón  $\xi^a$  reduce nuestra incertidumbre del patrón  $\xi^b$  es porque ambos comparten información y por ende nos permite hacer una predicción de uno a partir del otro.

La definición de  $IM$  se puede expresar en términos de la entropía de Shannon de cada patrón de la siguiente manera:

$$IM(\xi^a; \xi^b) = H(\xi^a) + H(\xi^b) - H(\xi^a, \xi^b),$$

donde  $H(\xi^a)$  y  $H(\xi^b)$  son las entropías marginales de los patrones  $\xi^a$  y  $\xi^b$  respectivamente y  $H(\xi^a, \xi^b)$  es su entropía conjunta (Latham y Roudi, 2009).

Dado que en la mayoría de los casos el conjunto memoria de una red de Hopfield está conformado por más de dos patrones, se empleó una generalización de la  $IM$  para más de dos variables, también llamada correlación total  $C$  (Watanabe, 1960). La correlación total nos habla del grado de correlación que existe entre todos los patrones del conjunto memoria y está dada por

$$C(\xi^1; \dots; \xi^P) = \left[ \sum_{\mu=1}^P H(\xi^\mu) \right] - H(\xi^1, \dots, \xi^P),$$

donde  $P$  es el número total de patrones almacenados.

Aunque estrictamente la correlación total es una ecuación distinta a la de la  $IM$ , nos referiremos a ésta como información mutua, ya que la primera es una generalización de la segunda, de modo que  $C = IM$ , considerando que aquí  $IM$  siempre tomará en cuenta más



de dos patrones.

Con el fin de normalizar la información mutua obtenida para un conjunto memoria y poder compararlo entre distintas redes, se dividió ésta entre la máxima información mutua, es decir:

$$IM_{Norm} = \frac{IM}{IM_{max}},$$

donde  $IM_{max}$  ocurre cuando una de las variables determina al resto (Rothstein, 1952). En este caso un patrón por sí mismo brindaría toda la información acerca de los valores del resto de los patrones y por lo tanto éstos se encontrarían relacionados de forma máxima.

Para esta parte se generaron redes de tamaño  $N = 100$  con un conjunto memoria conformado por  $P$  patrones donde la variable  $\xi_i^\mu$  de cada nuevo patrón tiene una probabilidad  $p$  de tomar el valor negativo<sup>2</sup> al de la variable  $\xi_i^1$  del primer patrón creado ( $\xi_i^\mu = 1$  si  $\xi_i^1 = 0$ ,  $\xi_i^\mu = 0$  si  $\xi_i^1 = 1$ ). En cada caso se comenzó con un patrón inicial  $\xi^1$  almacenado en la memoria a partir del cual se crearon el resto de los patrones. Para los valores de  $p$  se utilizó el intervalo  $[0, 1]$ , incrementando en 0.1 cada nuevo valor de  $p$ .

El experimento constó de 100 ensayos para cada valor de probabilidad, en los cuales se obtuvo la información mutua del conjunto memoria y el traslape.

## Topología.

Como se ha comentado anteriormente, la topología puede influir en el comportamiento global del sistema ya que determina la manera cómo las conexiones se distribuyen entre

---

<sup>2</sup>En lógica binaria, 1 y 0 son representaciones de los valores de verdad True (Verdadero) y False (Falso), respectivamente. En este caso el negativo de un valor es su negación, lo que nos da:  $\neg 1 = 0$  y  $\neg 0 = 1$ .

los nodos. Por esta razón se evaluó también el papel que tiene la topología en la función de la red. Tomando en cuenta la importancia que han adquirido las redes libres de escala recientemente (Stauffer *et al.*, 2003; Torres *et al.*, 2004; Perotti *et al.*, 2006), se comparó la estabilidad de los atractores en redes libres de escala (RLE) y en redes aleatorias (RA).

Las RLE se crearon a partir de secuencias de números enteros que siguieran una distribución de tipo ley de potencias con un exponente  $\gamma = 2.3$ . Se eligió este valor para  $\gamma$  debido a que las redes biológicas que exhiben una topología libre de escala parecen presentar exponentes dentro del intervalo  $2 < \gamma < 3$  (Albert y Barabási, 2002).

Para generar las redes con la topología deseada se siguió la misma metodología reportada anteriormente en estudios similares (McGraw y Menzinger, 2003; Kim, 2004; Dominguez *et al.*, 2007). Ésta consiste en multiplicar la matriz de pesos original creada mediante el aprendizaje Hebbiano por la matriz de adyacencia que representa la topología. De esta manera los pesos de la red resultante se reescriben como:

$$w_{ij} = a_{ij} \left[ \frac{1}{N} \sum_{\mu=1}^P (2\xi_i^\mu - 1)(2\xi_j^\mu - 1) \right],$$

donde  $a_{ij}$  es el elemento de la matriz de adyacencia que representa la conectividad entre las unidades  $i, j$ , por lo tanto si  $a_{ij} = 0$  entonces el peso resultante es cero sin importar el valor  $w_{ij}$  de la matriz de pesos original.

Para el caso de las RA se procedió de manera similar. Éstas se generaron a partir de las RLE con las que fueron comparadas, de modo que tuvieran los mismos valores  $N, M$  (tamaño de la red y número total de aristas). Cabe mencionar que el valor de  $M$  no fue el mismo en todos los ensayos debido a que cada vez que se creó una RLE se partió de distribuciones de grado distintas, sin embargo, el valor de  $M$  sí fue el mismo en cada par de redes cuya topología fue comparada.

En todo momento se compararon redes del mismo tamaño ( $N = 3,000$ ) y el mismo número de conexiones totales  $M$ .

Únicamente en esta parte del trabajo se evaluó también la estabilidad de redes bipolares en las que  $\sigma_i \in \{-1, 1\}$ . Para el caso de las redes bipolares se emplearon las mismas ecuaciones arriba descritas con sus respectivas modificaciones.

## **Sección 4: Resultados y discusión.**

Esta sección se divide en tres partes. En primer lugar se discuten los resultados de las simulaciones, posteriormente se discuten las implicaciones biológicas que podrían derivarse del modelo a partir de los resultados obtenidos y finalmente la relación entre estructura y potencial evolutivo.

### **Simulaciones.**

#### **Carga.**

En el análisis de la estabilidad de los atractores en función de la carga se observó una disminución de la estabilidad conforme la carga incrementaba. El que un patrón sea estable se refiere a que el atractor conserve su integridad, de forma que se tenga un traslape  $m = 1$  siempre que a la red se le presente como estado inicial uno de los patrones memorizados sin ninguna perturbación. Debido a que el número de repeticiones fue pequeño (20 para  $N = 10, 50, 100, 500$  y 10 para  $N = 1,000$ ), se reportan los valores de las medianas de  $m$  en vez del promedio.

Al principio de cada experimento las redes presentaron una estabilidad óptima con un traslape  $m = 1$  para todos sus patrones (por lo tanto el traslape global es igual a 1), pero

a partir de cierto umbral de carga crítica el traslape decae aceleradamente pero de manera continua (Figura 9). Esto contrasta con los resultados obtenidos por Dominguez *et al.* (2007) en los que se reporta una transición abrupta de  $m \rightarrow 0$  para valores de carga por encima de  $\alpha_c$ . Muy probablemente esto se deba a que aquí se reporta el traslape global de todos los patrones memorizados y no sólo de un patrón, ya que si se toma únicamente el traslape de uno de los patrones en medida que se incrementa la carga, sí se observa dicha transición.

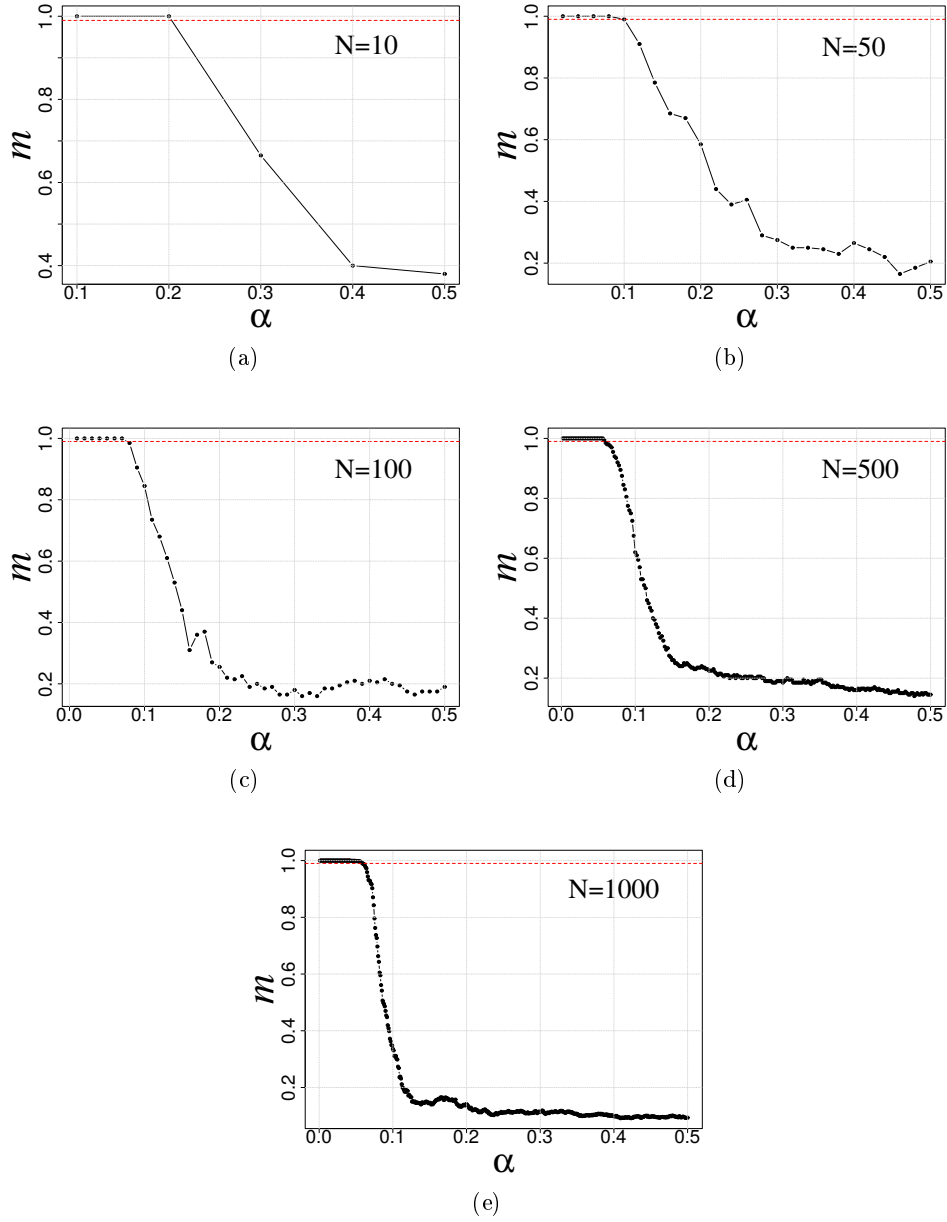


Figura 9: Traslape en función de la carga. Al principio la red se comporta de forma óptima hasta que a partir de un umbral de carga crítica se vuelve inestable. Cada punto corresponde a la mediana del traslape global de una muestra de  $n$  redes con carga  $\alpha$ . Para el caso de las redes con  $N = 10, 50, 100$  y  $500$ ,  $n = 20$  y para el caso de las redes con  $N = 1000$ ,  $n = 10$ .

Tabla 1: Comparación entre distintos tamaños de red ( $N$ )

Comparación $N$	p-val
10-50	8.823e-06
10-100	1.034e-07
10-500	3.498e-08
10-1000	4.185e-06
50-100	0.0155
50-500	3.813e-06
50-1000	9.097e-06
100-500	0.003366
100-1000	0.0002805
500-1000	0.0006301

Los valores de carga crítica para redes de tamaño  $N = 10, 50, 100, 500$  y  $1,000$  fueron  $\alpha_c = 0.2, 0.09, 0.08, 0.06$  y  $0.05$  respectivamente, lo cual pareciera indicar que los valores de  $\alpha_c$  disminuyen conforme el tamaño de la red aumenta (Figuras 9 y 10). Para comprobar lo anterior, dado que los valores de carga crítica no tienen una distribución normal y por tratarse de muestras independientes, se realizó primero una prueba no paramétrica de Kruskal-Wallis para comparar sus distribuciones. Ésta reveló que al menos dos tamaños de red poseen diferencias significativas en la distribución de su valores de  $\alpha_c$  [ $X^2 = 65.8769$ ,  $gl = 4$ ,  $p - val = 1.682e - 13$ ]. Para determinar cuáles fueron los tamaños de red cuyas  $\alpha_c$  varían, se llevó a cabo una prueba parizada de Wilcoxon-Matt-Whitney (también conocida como prueba  $U$  de Mann-Whitney). Éste último análisis mostró que los valores de las medianas de  $\alpha_c$  varían significativamente entre tamaños distintos de red al compararse entre sí (Tabla 1).

Si se grafican los valores de las medianas de  $\alpha_c$  obtenidos en los experimentos en función de  $N$ , vemos que éstas decaen en proporción  $1/\delta \ln N$ , lo cual coincide con lo reportado anteriormente por McEliece *et al.* (1987) para un conjunto memoria con  $P$  patrones in-

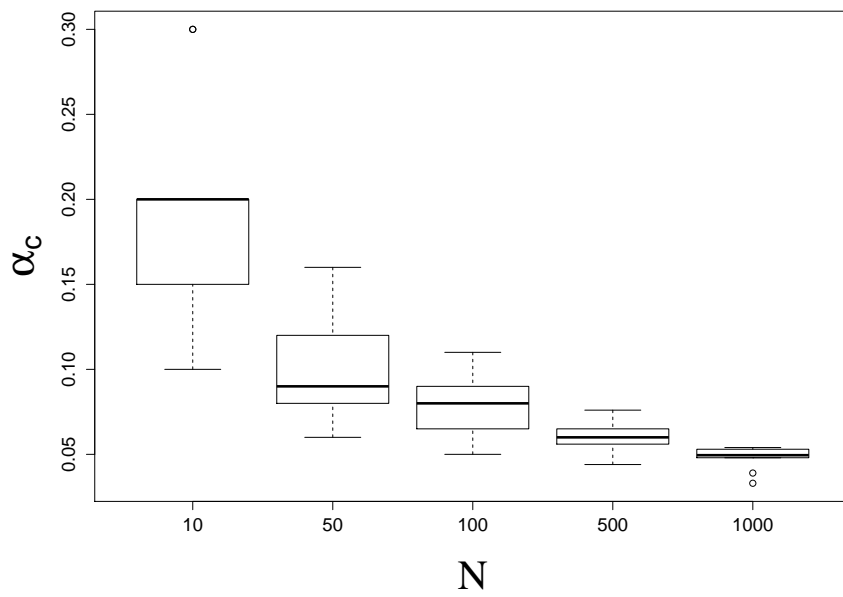


Figura 10: Distribución de los valores de carga crítica  $\alpha_c$  en función del tamaño de la red ( $N$ ). Los valores de  $\alpha_c$  así como su variación, disminuyen conforme el tamaño de la red aumenta.

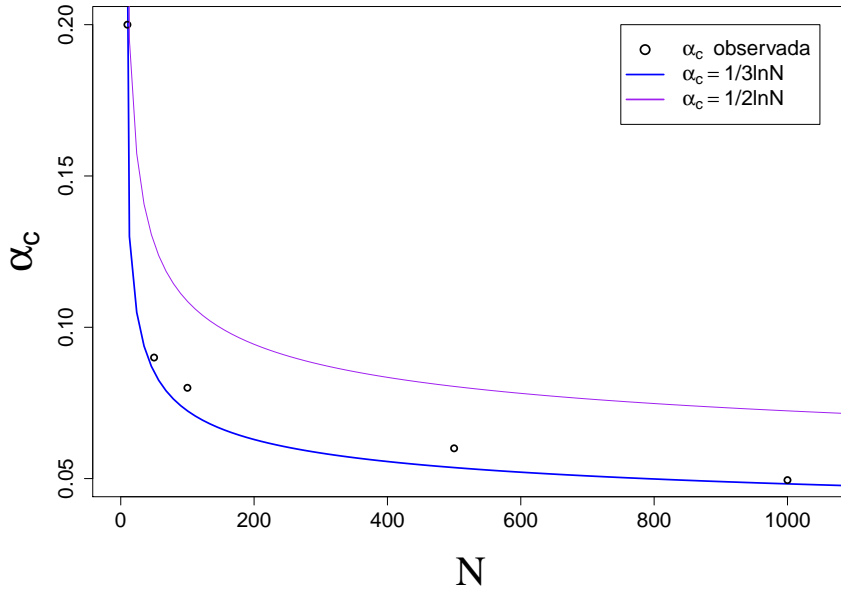


Figura 11: Valores de las medianas de las cargas críticas en función del tamaño de la red ( $N$ ). Se observó un decaimiento de  $\alpha_c$  en proporción  $1/\delta \ln N$ , con  $\delta \approx 3$ .

dependientes con  $p(\xi_i^\mu = \{0, 1\}) = 1/2$ . McEliece *et al.* colocan a  $P_{max}$  en el orden de  $N/\delta \ln N$ , con  $\delta$  no mayor a 4, si se desea que la red recupere los  $P$  patrones de manera exacta. En este caso  $\delta \approx 3$  como se muestra en la Figura 11, considerando un error permitido de 0.01 en el traslape global de la red.

### Información mutua.

En este experimento se utilizaron redes de tamaño  $N = 100$  y  $P = 5$ , ya que el valor de carga crítica para redes de este tamaño, sin aceptar ningún error, fue de 0.05. Esto se hizo con la finalidad de evitar el ruido que pudiese ser atribuido a la inestabilidad causada por la carga y no por la correlación entre los patrones.



Nuevamente para comprobar si existen diferencias significativas dado el efecto de  $p$ , se efectuó una prueba de Kruskal-Wallis seguida de una prueba de Wilcoxon-Matt-Whitney entre los distintos valores de  $p$ , tanto para la información mutua como para el traslape. Como es de esperar la información mutua entre los patrones memorizados presentó variación en función de  $p$  (con un p-val  $< 2.2e-16 < 0.05$ , siendo este último el nivel de significancia empleado para esta prueba). En la Tabla 2 se muestran los resultados de las pruebas de Wilcoxon-Matt-Whitney para cada comparación (los p-val no significativos se indican con \*).

En la Figura 12 (arriba) se observa que en medida que  $p$  aumenta su información mutua disminuye hasta alcanzar un valor mínimo ( $IM \approx 0.1$ ) que corresponde a un valor de  $p = 0.5$ . A partir de este punto, conforme  $p$  continúa incrementandose, también lo hace la información mutua. Este comportamiento es lo que se esperaría ya que en el intervalo  $0 < p < 0.5$ , conforme  $p$  incrementa, los patrones están menos correlacionados y por ende su información mutua es menor. Un valor de  $p = 0.5$  significa que los patrones son independientes entre sí y por lo tanto, su información mutua es mínima ( $IM \approx 0.1$ ). Por otro lado, en el intervalo  $0.5 < p < 1$ , conforme  $p$  incrementa, los patrones comienzan a parecerse otra vez más entre sí, lo que se ve reflejado en un incremento en la información mutua del conjunto memoria. Considérese que en los extremos del intervalo de variación de  $p$ , cuando  $p = 0$ , todos los patrones son iguales al patrón inicial y la información mutua es máxima ( $IM = 1$ ); asimismo, cuando  $p = 1$ , todos los demás patrones son el negativo del patrón inicial y por lo tanto iguales entre sí excepto el patrón inicial, sin embargo, el grado de correlación entre el conjunto memoria también es máximo, lo que resulta nuevamente en un valor de  $IM = 1$ . Los resultados de la prueba de Wilcoxon-Matt-Whitney (Tabla 2) indican que el comportamiento de la  $IM$  en función de  $p$  es en efecto simétrico entre los valores del intervalo  $0 < p < 0.5$ , con respecto a los del intervalo  $0.5 < p < 1$ .

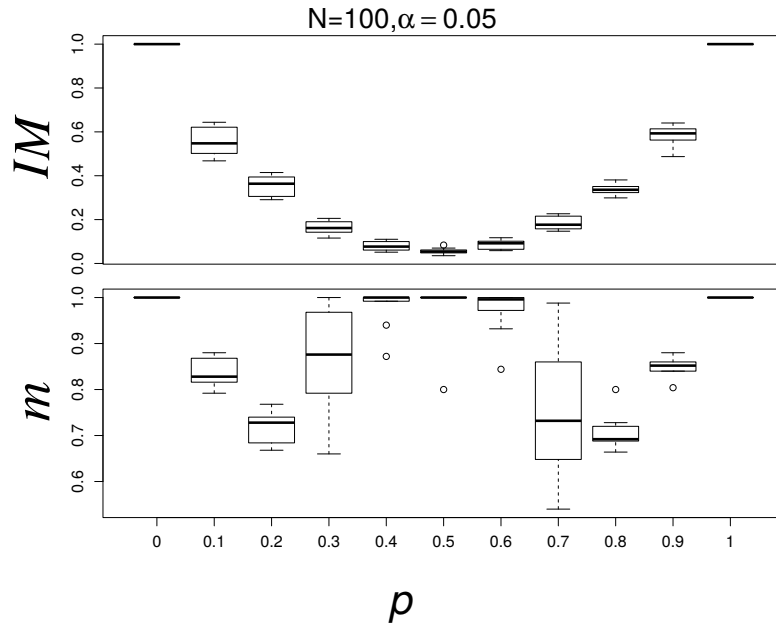


Figura 12: . Información mutua y traslape en función de  $p$ . Se hicieron 10 ensayos para cada valor de  $p$ .

Así como sucede con la información mutua, el cambio en  $p$  también provoca variaciones en la estabilidad de la red (con un  $p\text{-val} < 2.2e-16 < 0.05$ ). Los  $p\text{-val}$  de la prueba de Wilcoxon-Matt-Whitney se muestran en la Tabla 2 (nuevamente los  $p\text{-val}$  no significativos se indican con \*). En la Figura 12 (abajo) se observa que cuando  $IM = 1$  (esto es en los valores de  $p = 0$  y  $p = 1$ ), la estabilidad de la red es máxima ( $m = 1$  en todos los ensayos), pero fuera de estos dos valores (para  $0.1 < p < 0.9$ ) en la medida que disminuye la  $IM$  entre los patrones, la estabilidad de las redes también disminuye hasta alcanzar valores mínimos de traslape. En cambio cuando los patrones del conjunto memoria son prácticamente independientes entre sí (esto es en  $p = 0.4, 0.5$  y  $0.6$ ), y por lo tanto la  $IM$  entre ellos es mínima ( $IM \approx 0.1$ ), las redes vuelven a presentar una estabilidad casi óptima, con un traslape  $m \approx 1$ . Este comportamiento es más sencillo de comprender si se analiza desde el punto de vista del tamaño y el tipo de cuenca de atracción de cada

Tabla 2: Comparación entre distintos valores de  $p$ . Información mutua ( $IM$ ) y traslape ( $m$ ).

Comparación p	p-val ( $IM$ )	p-val ( $m$ )
0.1-0.2	2.835e-06	7.005e-05
0.1-0.3	2.835e-06	0.0002637
0.1-0.4	2.835e-06	6.154e-05
0.1-0.5	2.835e-06	4.692e-05
0.1-0.6	2.835e-06	0.003376
0.1-0.7	2.835e-06	0.0003898
0.1-0.8	2.835e-06	7.039e-05
0.1-0.9	0.7477 *	0.6693 *
0.2-0.3	2.835e-06	0.01951
0.2-0.4	2.835e-06	6.065e-05
0.2-0.5	2.835e-06	4.62e-05
0.2-0.6	2.835e-06	6.648e-05
0.2-0.7	2.835e-06	0.2782 *
0.2-0.8	0.4009 *	0.6449 *
0.2-0.9	2.835e-06	7.005e-05
0.3-0.4	2.835e-06	6.035e-05
0.3-0.5	2.835e-06	4.596e-05
0.3-0.6	2.835e-06	8.719e-05
0.3-0.7	0.1513 *	0.818 *
0.3-0.8	2.835e-06	0.01133
0.3-0.9	2.835e-06	9.173e-05
0.4-0.5	0.001401	0.3346 *
0.4-0.6	0.4009 *	0.2866 *
0.4-0.7	2.835e-06	6.154e-05
0.4-0.8	2.835e-06	6.095e-05
0.4-0.9	2.835e-06	6.154e-05
0.5-0.6	0.00104	0.07853 *
0.5-0.7	2.835e-06	4.692e-05
0.5-0.8	2.835e-06	4.644e-05
0.5-0.9	2.835e-06	4.692e-05
0.6-0.7	2.835e-06	0.0001165
0.6-0.8	2.835e-06	6.68e-05
0.6-0.9	2.835e-06	0.003027
0.7-0.8	2.835e-06	0.2783 *
0.7-0.9	2.835e-06	1.134e-05
0.8-0.9	2.835e-06	7.039e-05

uno de los atractores del sistema. Cuando  $p = 0$ , todos los patrones son iguales y por lo tanto el sistema posee un único atractor. Ya que la cuenca de atracción es una sola y de gran tamaño, todos los estados iniciales conducen a ese único estado estable y por ende el traslape es máximo. Conforme  $p$  incrementa y los patrones del conjunto memoria varían más entre sí, comienzan a aparecer nuevos atractores (atractores espurios). En este caso, entre mayor sea la información mutua de los atractores, y por lo tanto su correlación, podría ocurrir que las cuencas de atracción se estén traslapando, lo que provocaría la inestabilidad del sistema. Por otro lado, cuando los patrones del conjunto memoria son independientes y su información mutua es mínima ( $p = 0.4, 0.5$  y  $0.6$ ) podríamos decir, con fines meramente explicativos, que cada atractor posee la “menor información del resto de los atractores”, pero la “mayor información de sí mismo”, generando un espacio de estados donde cada atractor está bien delimitado y la distancia de sus cuencas de atracción es la mayor posible. Finalmente, cuando  $p = 1$ , sólo se tienen dos atractores, pero éstos son complementarios ya que uno de los patrones es el negativo del otro, por lo tanto están altamente correlacionados y presentan una  $IM$  máxima. Por la misma razón de ser ortogonales, sus cuencas de atracción se encuentran bien delimitadas y separadas entre sí, lo que permite que el sistema sea estable.

Por lo tanto el grado de correlación entre los patrones almacenados en memoria tiene un efecto en la estabilidad de sus atractores y en la función del sistema. Cuando los patrones están menos correlacionados y su información mutua es baja, la estabilidad es mayor y el sistema se desempeña de manera más eficiente.

## Topología.

En el análisis de la topología se observó que en general tanto las RLE como las RA muestran un desempeño más bajo que redes de Hopfield completamente conectadas.

Cuando se comparó la topología entre RLE y RA binarias no se observaron diferencias significativas en la estabilidad de los atractores para valores de  $\alpha < 0.012$ . Pasado dicho valor de carga, las RA se desempeñaron mejor que las RLE, sin embargo la diferencia en el traslape no fue muy grande, aunado a que el desempeño de ambas topologías fue más bien pobre para estos niveles de carga. También se observó un mayor traslape por parte de las RA cuando se tenía únicamente un patrón en memoria ( $P = 1$ ,  $m \approx 0.89$ ), en comparación con el de las RLE ( $P = 1$ ,  $m \approx 0.69$ ) (Figura 13).

En las redes bipolares en cambio, sí se observaron diferencias significativas en la estabilidad de los atractores. Para casi todos los valores de carga las RA presentaron un mejor desempeño que las RLE. El único caso en el que las RLE presentaron mayor estabilidad fue cuando se tenía un solo patrón almacenado, en cuyo caso el traslape para RLE y RA fue de  $m = 1$  y  $m = 0.998$  respectivamente, sin embargo dicha diferencia no fue significativa. Se observó también que para el caso de las RLE, la diferencia entre el traslape cuando se tiene un solo patrón ( $P = 1$ ) y cuando se añade un patrón más ( $P = 2$ ), fue de 0.26, en comparación con el de las RA que fue sólo de 0.05. Esto significa un desempeño cinco veces mayor por parte de las RA con respecto a las RLE en la transición de  $P = 1$  a  $P = 2$  (Figura 14).

En general las RA construídas mediante la metodología aquí descrita, parecieran desempeñarse mejor que las RLE del mismo tamaño y con el mismo número de conexiones ( $N$ ,  $M$ ). Estos resultados concuerdan con otros estudios similares como los de McGraw y

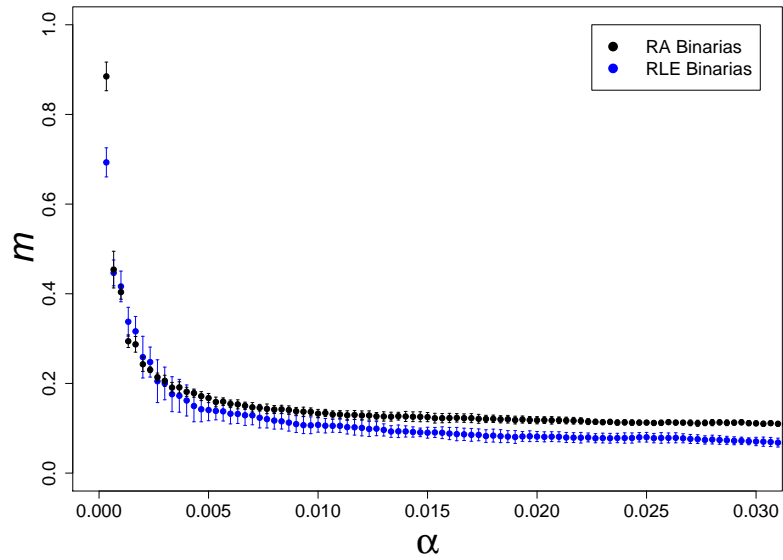


Figura 13: Comparación del desempeño de topologías en redes binarias. RA (Redes Aleatorias), RLE (Redes Libres de Escala).

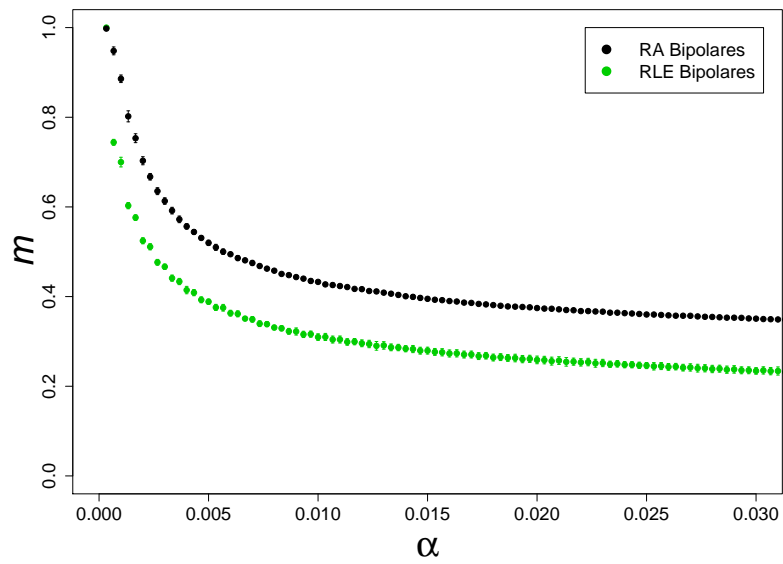


Figura 14: Comparación del desempeño de topologías en redes bipolares. RA (Redes Aleatorias), RLE (Redes Libres de Escala).

Menzinger (2003), también con redes bipolares, en los que se reporta un mejor desempeño por parte de las RA, sin embargo en otros casos se reporta lo contrario. Por ejemplo, Torres *et al.* (2004) estudiaron la influencia que tiene la topología en la capacidad de una red bipolar de recuperar un solo patrón con distintos grados de perturbación. Los autores reportan un mejor desempeño para las RLE en comparación con las RA. A pesar de que en los resultados aquí obtenidos también se observó cierta ventaja por parte de las RLE bipolares cuando  $P = 1$ , las diferencias no fueron significativas como para asegurar este hecho, además faltaría comprobar que el desempeño mencionado por Torres *et al.* se mantiene para  $P > 1$ .

La mayoría de estudios revisados para este trabajo en torno a la topología y el desempeño de las redes apuntan a que la capacidad de la red incrementa en función del grado de los nodos (Labieuse *et al.*, 2002; McGraw y Menzinger, 2003; Kim, 2004; Morelli *et al.*, 2004; Löwe y Vermet, 2015). Ésta pareciera ser la causa de que las RA hayan presentado un mejor desempeño que las RLE, dado que la distribución de las conexiones en las RA es más uniforme que en las RLE. De hecho, estudios como el de Kim (2004) sugieren que la eficiencia de las redes de Hopfield depende mucho más del coeficiente de agrupamiento o “clustering” que de la propia topología. En general, entre menor es el coeficiente de agrupamiento mejor es el desempeño, lo cual concuerda con los resultados obtenidos. En este caso las RA presentaron un coeficiente de agrupamiento menor que las RLE (Tabla 3). Esto es de esperar si se considera que las RLE suelen ser redes de mundo pequeño, siendo uno de los requisitos para que una red sea de mundo pequeño el que su coeficiente de agrupamiento sea mayor al de una red aleatoria con el mismo número de nodos.

Tabla 3: Valores promedio de los parámetros de las redes con las que se experimentó. Tamaño de la red ( $N$ ), número de aristas ( $M$ ), grado promedio ( $K$ ), grado máximo ( $k_{max}$ ) y coeficiente de agrupamiento ( $C$ ).

<i>Parámetro</i>	RLE	RA
$N$	3000	3000
$M$	5311.333	5311.333
$K$	3.541	3.541
$k_{max}$	621.333	11.333
$C$	0.781	0.001

### Consideraciones biológicas del modelo.

Como se ha venido mencionando, de manera general una RRG presenta similitudes con una red de Hopfield. Durante el desarrollo embrionario, por ejemplo, los estímulos externos que inciden directa o indirectamente sobre la RRG llevan al sistema a adquirir un estado a partir del cual la estructura toma control del proceso. Ésta es básicamente la idea detrás del concepto de canalización de Waddington, siendo además que la canalización no requiere de mecanismos especializados sino que surge como una propiedad emergente de las RRG (Bergman y Siegal, 2003).

Es cierto que el modelo de Hopfield implica situaciones que parecieran ser poco plausibles en términos biológicos, como suponer sistemas totalmente conectados y relaciones simétricas entre sus partes, pero partir del hecho de que los sistemas naturales poseen topologías diluidas puede resultar igual de arbitrario, todo depende de qué aspectos del sistema se busque representar y la escala a la que se establecen las partes y las relaciones. Por ejemplo, si un sistema está conformado por pocos elementos es de suponer que éstos estén altamente correlacionados. Además, las relaciones entre los componentes no necesariamente representan interacciones físicas. En el caso de las redes de Hopfield la matriz de pesos  $W$  no representa las interacciones que podrían darse entre los componentes sino su grado de



correlación (Rojas, 1996). De esta forma en la matriz de pesos de una RRG modelada como una red de Hopfield estarían implícitos una serie de procesos tanto genéticos como epigenéticos que en conjunto provocan que la expresión de un gen específico influya en la expresión de otro.

Suponiendo que las RRG puedan comportarse como redes de Hopfield y presentar una evolución en principio similar al aprendizaje Hebbiano como afirman Watson y colaboradores (Watson *et al.*, 2010; Watson y Szathmáry, 2016; Watson *et al.*, 2014, 2011b), los resultados de los experimentos en los que se probó el efecto de la carga y de la información mutua, dan pie a preguntarse dos cosas: 1) ¿En qué medida el número de genes determina el número de tipos celulares de un organismo, y 2) ¿Cómo es la correlación entre los perfiles de expresión que caracterizan los tipos celulares?

En relación a la primera pregunta ciertamente sería un error decir que los valores de  $P_{max}$  obtenidos en las simulaciones corresponden con el número de tipos celulares en organismos con genomas de ese mismo tamaño, ya que los valores que se esperarían a partir de la relación  $N/\delta \ln N$  (con  $\delta$  entre 2 y 6) es muy superior al número de tipos celulares que supuestamente se observa en la mayoría de estos organismos. Lo que sí valdría la pena preguntarse en este caso es qué tan válido es incluir genomas completos en este tipo de análisis, ya que es sabido que no todos sino sólo una fracción de los genes participan en el establecimiento del tipo celular (Szathmáry *et al.*, 2001; Xu y Su, 2015). Factores como la duplicación de genes así como el efecto que puedan tener distintas “estrategias” biológicas en la expresión genética podrían estar introduciendo ruido en las mediciones. Por ejemplo, mientras los vertebrados poseen un gran número de genes relacionados con la formación y funcionamiento de los sistemas inmune y nervioso, las plantas, posiblemente a causa de su vida sésil, dependen de sensor de manera precisa su entorno para responder acorde al cambio, dedicando gran parte de su genoma a su metabolismo secundario (Szathmáry

*et al.*, 2001). Algo similar podría estar ocurriendo en la determinación del tipo celular, en la que distintos subconjuntos del genoma parecen contribuir de forma diferencial, o donde básicamente sea un grupo relativamente pequeño de genes el que determine las principales líneas celulares.

Tomando en cuenta lo anterior y a modo meramente heurístico, se hizo una regresión lineal para conocer la relación entre el número de tipos celulares reportados para distintas especies eucariontes y el número de genes, considerando 1) el número total de genes en el genoma, y 2) únicamente el número de factores de transcripción (FT). Las especies que fueron consideradas se listan en la Tabla 4.

En este pequeño experimento se observó un mejor ajuste entre el número de tipos celulares y los FT que con respecto al número total de genes ( $R^2 = 0.128$  y  $R^2 = 0.053$  respectivamente; Figuras 15b y 15a). Más interesante aún, si no se incluye a las plantas en el análisis, el coeficiente de determinación incrementa notablemente en ambos casos, aunque sigue siendo mayor si se toma en cuenta únicamente el número de FT en vez del número total de genes ( $R^2 = 0.348$  tomando todos los genes y  $R^2 = 0.698$  tomando únicamente el número de FT; Figuras 15c y 15d). Estos resultados coinciden con los reportados por Schad *et al.* (2011), quienes encontraron una correlación entre el número de genes (tomado como el tamaño del proteoma reducido) y el número de tipos celulares de los mismos 52 organismos eucariontes considerados aquí. Aunque haría falta hacer un análisis más detallado que incluya un número mayor de organismos, estos resultados sugieren que si bien el tamaño del genoma no determina el número de tipos celulares, sí podrían existir grupos de genes o módulos que contribuyan más que otros, siendo además que dichos módulos varíen entre especies.

Por otro lado, determinar cuántos tipos celulares hay en un tejido o en un organismo es complicado, la mayoría de las veces los tipos celulares se definen ya sea por su morfología

o por la actividad de un número pequeño de genes. En el caso del ser humano, por ejemplo, estimaciones aproximadas reportan entre 150 y 200 tipos celulares diferentes (169 de acuerdo con Schad *et al.* (2011)), mientras que los listados más completos reportan que en el individuo adulto existen unos 400 tipos distintos (Vickaryous y Hall, 2006). Sin embargo muchos de los tipos celulares que un óvulo fecundado es capaz de producir no necesariamente están presentes en el individuo, algunos de ellos conforman la placenta y otros tejidos que controlan el desarrollo del embrión y el feto (Galvão *et al.*, 2010). Tomando todo esto en cuenta el número de tipos celulares reportados en la literatura es de alrededor de 800 (Galvão *et al.*, 2010), el cual, curiosamente, se acerca mucho al número de tipos celulares esperados para un genoma de 25,000 genes como en el caso del ser humano, si se toma  $P_{max} = N/\delta \ln N$  con  $\delta = 3$ , que es alrededor de 822. Considerando lo anterior, probablemente la aplicación de técnicas de secuenciación de nueva generación y los nuevos métodos bioinformáticos para la identificación de tipos celulares, con un enfoque más genómico, nos permitan “*elaborar definiciones más sistémicas de lo que es un tipo celular*” (Schwartz, 1999).

Con respecto a la segunda pregunta, ¿cómo es la correlación entre los perfiles de expresión que caracterizan los tipos celulares?, en los experimentos en los que se usó la información mutua como medida de correlación entre patrones, se observó que la *IM* afecta la estabilidad de la red. Por lo tanto, si una RRG fuera similar a una red de Hopfield en este aspecto, significaría que el grado de correlación entre los patrones, en este caso los perfiles de expresión que determinan un tipo celular, afectaría la diferenciación. Para indagar un poco en ello se tomaron perfiles de expresión reales y se midió su *IM*. Los perfiles que se consideraron corresponden a los patrones de expresión genética de 21,399 genes pertenecientes a 10 tipos celulares distintos, provenientes de tejidos sanos en el ser humano y obtenidos

Tabla 4: Número de tipos celulares, tamaño de genoma y número de FT de distintas especies eucariontes. El número de tipos celulares es el promedio reportado para esa especie. Datos obtenidos de Schadt *et al.* (2011) y completados a partir de las bases de datos REGULATOR (Wang y Nishida, 2015), PlantTFDB (Jin *et al.*, 2017) y DBD (Wilson *et al.*, 2008).

	Clasificación	Nombre de Sp.	Tipos Cel.	Tam. Genoma	Num. FT
1	algas	Ostreococcus lucimarinus	12.50	7651	111
2	algas	Ostreococcus tauri	12.50	7725	99
3	algas	Micromonas pusilla	12.50	NA	150
4	algas	Volvox carteri	14.50	15544	125
5	algas	Chlamydomonas reinhardtii	12.50	15256	212
6	deuterostomados	Ciona intestinalis	74.00	14180	394
7	deuterostomados	Tupaia belangeri	159.00	15462	798
8	deuterostomados	Anolis carolinensis	140.00	28102	1211
9	deuterostomados	Gallus gallus	154.00	16736	758
10	deuterostomados	Xenopus tropicalis	129.50	18023	1030
11	deuterostomados	Takifugu rubripes	119.50	18523	1233
12	deuterostomados	Canis familiaris	159.00	19304	1028
13	deuterostomados	Tetraodon nigroviridis	119.50	27991	1362
14	deuterostomados	Pan troglodytes	169.00	20543	1366
15	deuterostomados	Bos taurus	159.00	21746	1217
16	deuterostomados	Danio rerio	119.50	21322	1291
17	deuterostomados	Homo sapiens	169.00	22997	1508
18	deuterostomados	Rattus norvegicus	159.00	22503	1202
19	deuterostomados	Mus musculus	159.00	23873	1426
20	deuterostomados	Branchiostoma floridae	100.00	50817	2498
21	hongos	Encephalitozoon cuniculi	3.35	1996	35
22	hongos	Schizosaccharomyces pombe	4.35	5001	100
23	hongos	Kluyveromyces lactis	3.05	5331	155
24	hongos	Yarrowia lipolytica	3.05	6659	156
25	hongos	Saccharomyces cerevisiae	3.05	6698	117
26	hongos	Ustilago maydis	4.35	6522	192
27	hongos	Neurospora crassa	5.55	10620	213
28	hongos	Emericella nidulans (Aspergillus n.)	5.55	NA	357
29	hongos	Phanerochaete chrysosporium	4.35	10048	146
30	protostomados	Amphimedon queenslandica	16.00	NA	227
31	protostomados	Anopheles gambiae	64.00	12457	564
32	protostomados	Drosophila melanogaster	64.00	14141	601
33	protostomados	Hydra magnipapillata	22.00	17398	441
34	protostomados	Caenorhabditis elegans	28.50	20054	695
35	protostomados	Nematostella vectensis	22.00	27273	701
36	protostomados	Daphnia pulex	50.00	30940	563
37	protozoos	Theileria annulata	7.85	3795	16
38	protozoos	Plasmodium falciparum	7.85	5405	18
39	protozoos	Leishmania major	7.85	8280	27
40	protozoos	Trypanosoma brucei	7.85	8758	30
41	protozoos	Entamoeba histolytica	4.65	9772	69
42	protozoos	Dictyostelium discoideum	4.65	13681	116
43	protozoos	Phytophthora ramorum	7.85	16066	160
44	plantas	Physcomitrella patens	22.00	35938	1284
45	plantas	Selaginella moellendorffii	25.00	NA	665
46	plantas	Brachypodium distachyon	27.25	NA	1751
47	plantas	Zea mays	27.25	64180	3316
48	plantas	Oryza sativa	27.25	66710	1891
49	plantas	Sorghum bicolor	27.25	34496	2198
50	plantas	Arabidopsis thaliana	27.25	27235	1356
51	plantas	Vitis vinifera	27.25	30434	1500
52	plantas	Populus balsamifera	28.50	NA	4288

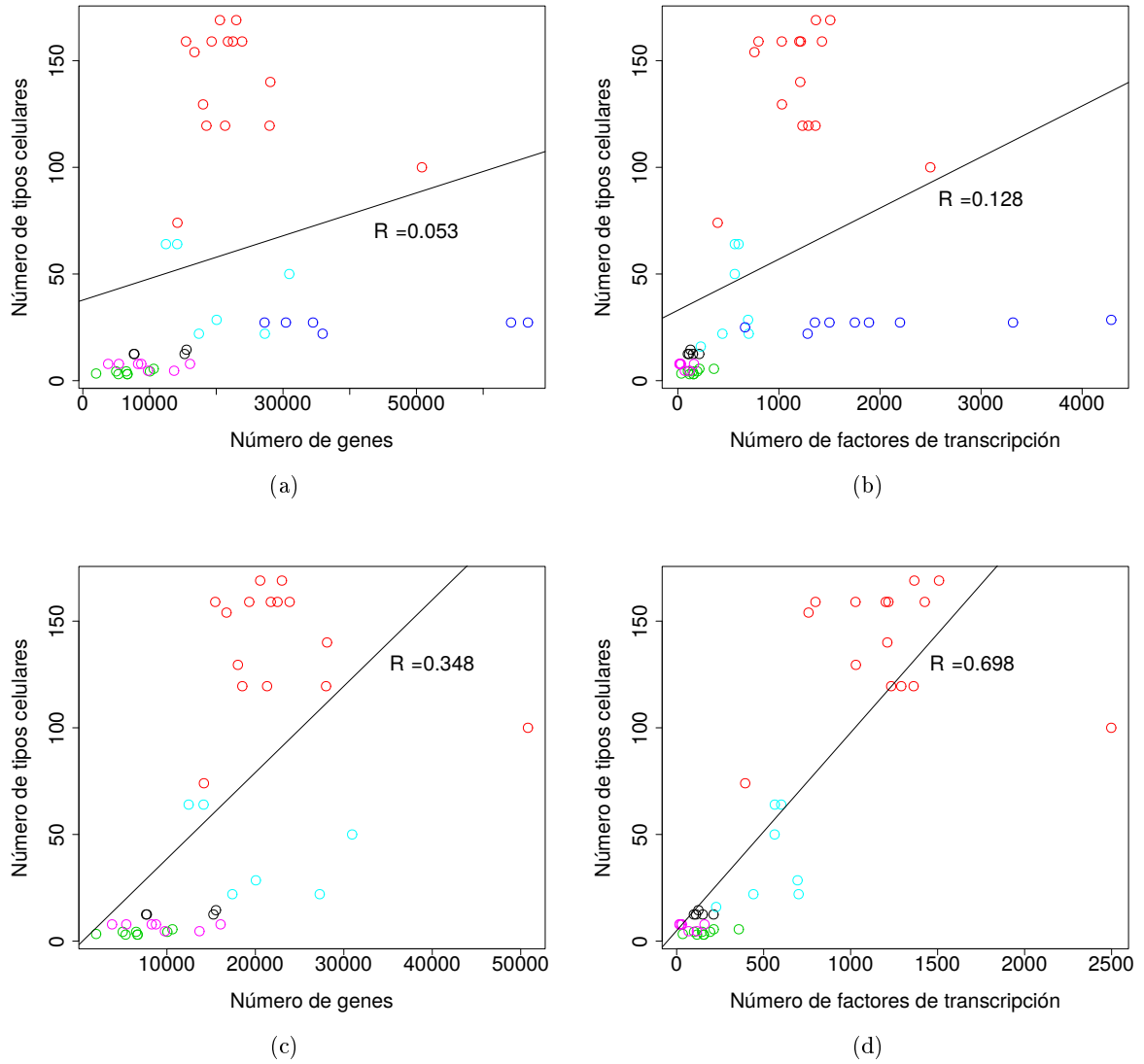


Figura 15: Correlación entre el número de tipos celulares y el número de genes. En (a) y (b) se muestra la correlación tomando en cuenta todas las especies de la Tabla 3 y en (c) y (d) la correlación sin considerar a las plantas (en azul oscuro). En (a) y (c) el número de genes corresponde al total de genes en el genoma, mientras que en (b) y (d) corresponde únicamente a los factores de transcripción (FT). En ambos casos se observó un mayor coeficiente de determinación al compararse el número de tipos celulares con el número de FT, pero cuando no se toman en cuenta a las plantas, el coeficiente incrementa notablemente.

de la base de datos ReNA Seq Atlas (Krupp *et al.*, 2012). Los valores de expresión de cada gen proporcionados por ReNA Seq Atlas son valores estandarizados (Z-score), de modo que para discretizar los perfiles de expresión y crear un vector con valores 0 y 1, se tomaron como 0 todos los valores de expresión con signo negativo y como 1 todos los valores con signo positivo.

El valor de información mutua que se obtuvo de los perfiles de expresión considerados fue  $IM = 0.32$ . Si bien los tipos celulares no están altamente correlacionados, estos tampoco son aleatorios (en cuyo caso  $IM = 0$ ). Lo anterior puede comprobarse también, aunque de forma cualitativa, en las figuras 16 y 17. En la Figura 16 se muestra la matriz de pesos de los diez tipos celulares considerados, tomando únicamente los primeros 4,000 genes (Figura 16c), contrastada con las matrices de pesos de dos redes del mismo tamaño, una formada con patrones aleatorios (Figura 16a) y otra con patrones correlacionados (Figura 16b). En la Figura 17 se muestran las distribuciones de los pesos de cada una de las redes arriba mencionadas. La matriz de pesos resultante de los perfiles de expresión es cualitativamente distinta a la que se obtendría a partir de patrones ya sean aleatorios o correlacionados. Como es de esperar, en la matriz de patrones aleatorios predominan los pesos con bajo valor absoluto (0 y 2, Figura 17a) debido a que la probabilidad de que cualesquiera dos nodos mantengan un patrón de activación a lo largo de todos los perfiles es bajo. Por el contrario, en la matriz de patrones correlacionados los pesos se refuerzan, predominando aquellos de mayor valor absoluto (8 y 10, Figura 17b). En cambio, en la matriz de pesos de los perfiles de expresión humanos se observa una distribución de los pesos un tanto más homogénea, con valores absolutos entre 0 y 10 (Figura 17c), tomando en cuenta que  $|w_{ij}| = 10$  significa que los genes  $i, j$  conservaron el mismo patrón de activación en todos los tipos celulares. Muy probablemente lo que aquí se observa se deba a que los seres vivos (y por supuesto no sólo los seres vivos) tienden a variar al mismo tiempo que conservan estados

anteriores. Esto tiene sentido desde un punto de vista evolutivo si se considera que todos los tipos celulares de un organismo debieron originarse a partir de un tipo celular primitivo; además, como sabemos, no todos los genes están involucrados en la determinación del tipo celular, o bien lo hacen de forma modular, lo que podría estar generando regiones de genes en los perfiles de expresión que tienden a variar más que otras.

Si se toma el valor de información mutua observado en los perfiles de expresión humanos ( $IM = 0.32$ ) y se compara con una red de Hopfield con un valor de  $IM$  similar, vemos que el valor del traslape es  $m < 0.7$ , la cual podría considerarse relativamente bajo si pensamos en lo robusto que es el proceso de desarrollo. En este sentido, para que un sistema genético con estas características fuera estable y pudiera hacerse más complejo, debería ser capaz de escapar de las restricciones de la correlación simultánea de todos sus genes. Esto lo hace mediante la regulación. La regulación genética es en realidad un tipo de modularidad espacio-temporal en el control genético o fisiológico en el que las relaciones entre componentes cambian en el tiempo. De acuerdo con Wagner *et al.* (2007), la modularidad es un concepto abstracto que se refleja a varios niveles, formando distintos tipos de heterogeneidad en los organismos y es considerado un aspecto fundamental de la organización biológica. La modularidad podría generar una disminución de los costos de conectividad y proveer al organismo de algunas ventajas como la variación a un nivel mayor de organización (Watson *et al.*, 2010), por lo tanto, la modularidad podría ser una de las formas en las que el sistema gane grados de libertad.

En un estudio realizado por Galvão *et al.* (2010) se encontró que durante el desarrollo embrionario humano sólo unos cuantos tipos celulares se originan en un intervalo de tiempo dado, de modo que la actividad de la red se centra en un número pequeño de tipos celulares a la vez. Paralelamente es probable que sólo un grupo de genes se encuentren altamente

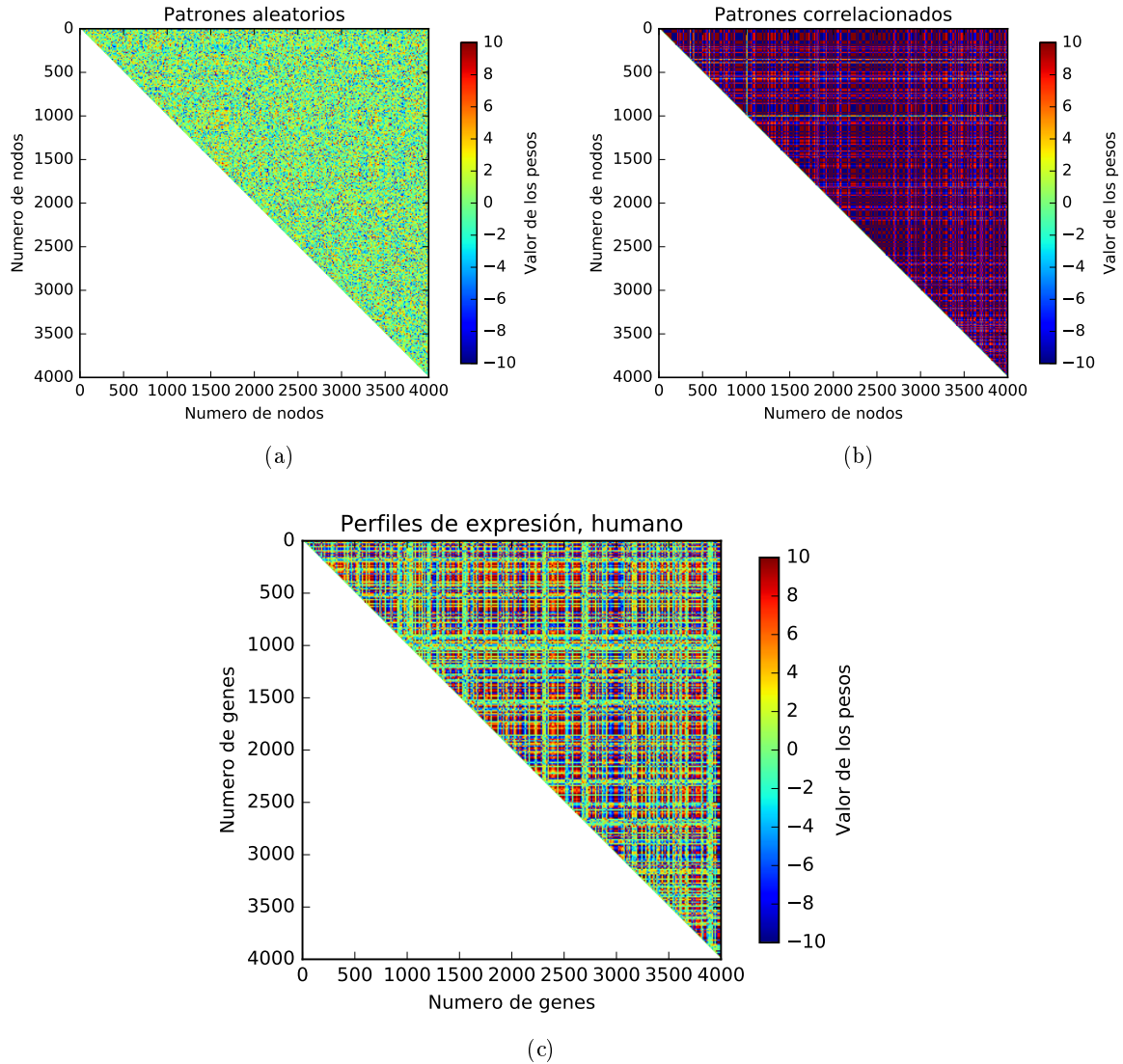


Figura 16: Comparación de tres matrices de pesos  $W$  construídas a partir de: (a) patrones con estados de activación aleatorios, (b) patrones estados de activación altamente correlacionados y (c) perfiles de expresión reales en el ser humano. En los tres casos se usaron redes de  $N = 4,000$  nodos y  $P = 10$ . En los perfiles de expresión humanos estos valores corresponden con 4,000 genes expresados en 10 tipos celulares distintos.



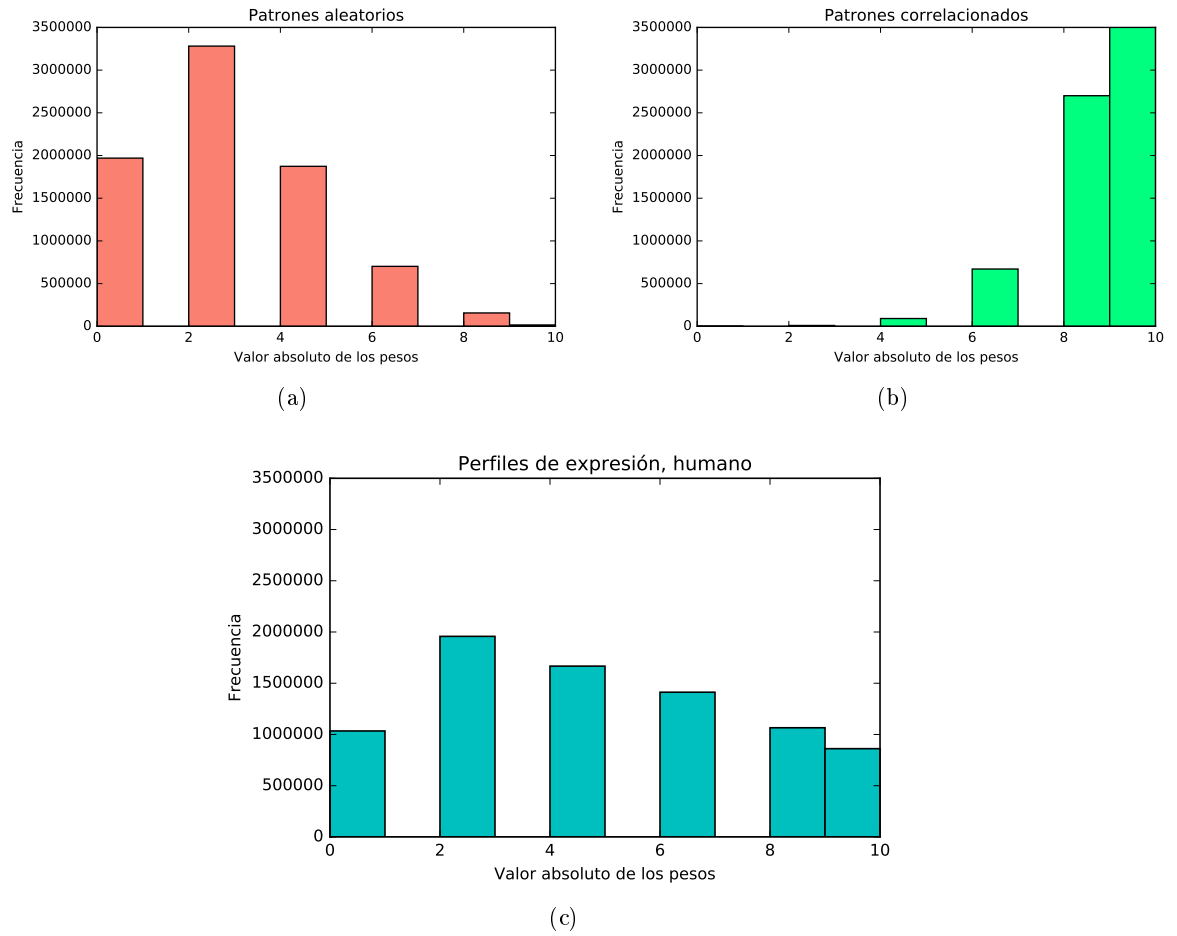


Figura 17: Distribución de los pesos de las matrices de la figura 16. En (a) se muestra la distribución de los pesos de los patrones aleatorios, en (b) la de los pesos de los patrones correlacionados y en (c) la de los pesos de los perfiles de expresión.

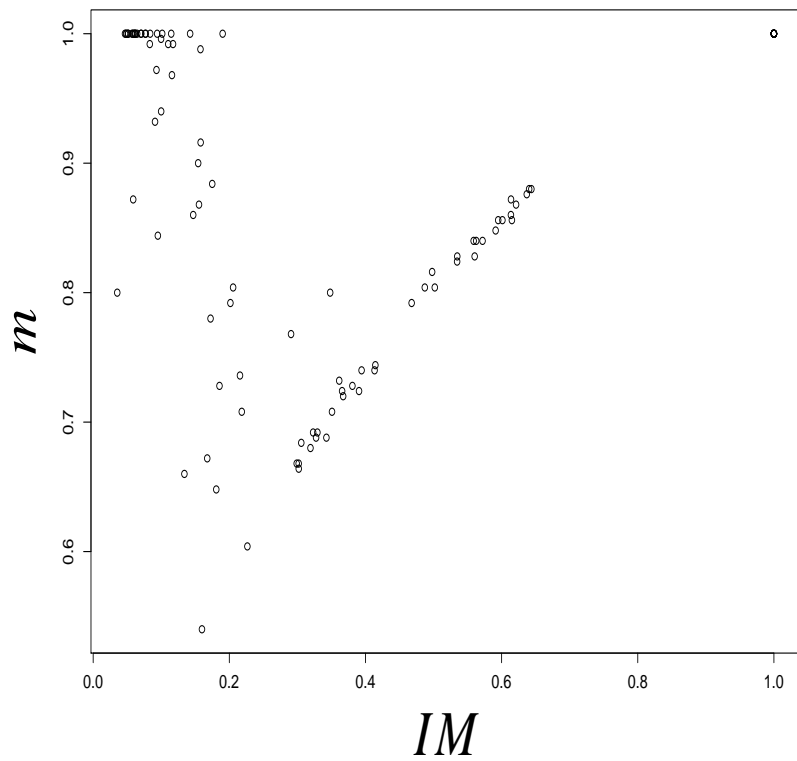


Figura 18: Traslape ( $m$ ) en función de la información mutua ( $IM$ ).

correlacionados en cada intervalo de tiempo. En otros estudios se menciona además que el origen de la especialización funcional entre caracteres homólogos es acompañado por una disminución de la correlación entre dichos caracteres. Wagner *et al.* (2007), por ejemplo, mencionan los trabajos de Berg (1960) y Hansen *et al.* (2007) en plantas, y de Young y Hallgrímsson (2005) en animales, en los que muestran que las partes con funciones más especializadas del organismo están menos correlacionadas entre sí que las partes menos especializadas. En este sentido la regulación genética y su efecto modular, evita la "diafonía" causada por la correlación de todos los genes. Por lo tanto, si bien es poco probable que todos los genes de un genoma se encuentren altamente correlacionados, sí podría ser éste el caso para ciertos grupos de genes que actúan como módulos, los cuales llevan a cabo funciones fisiológicas o regulatorias que son semiautónomas en relación a las de otros módulos. Adicionalmente, la correlación de un conjunto completo de genes, podría haberse dado en etapas muy tempranas de la vida, cuando se tenían relativamente pocos genes y donde la modularidad y la regulación era aún muy baja o poco probable.

Por último, con respecto al efecto de la topología, no es posible afirmar que las redes aleatorias sean más eficientes que las libres de escala, ya que en primer lugar depende de si se toma como modelo una red binaria o una red bipolar. Aunque las redes aleatorias presentaron en general un mejor desempeño, los resultados obtenidos de la *IM* en perfiles de expresión reales en el ser humano indican que los pesos en estas redes de Hopfield no son aleatorias. La razón por la que las RA presentaron en este caso mejor desempeño es porque afectan menos conexiones por nodo que las RLE. Aunque la mayoría de los estudios de este tipo emplean la matriz de adyacencia  $A$  para generar la topología deseada en redes de Hopfield, esta metodología no captura el verdadero efecto que la topología pudiera tener en el sistema ya que prácticamente no toma en cuenta la información de los pesos. En el modelo tradicional (completamente conectado) de Hopfield cada peso entre dos unidades

cualesquiera lleva implícita información del sistema, obtenida a partir de la regla Hebbiana de aprendizaje, pero cuando se multiplica  $W \times A$  muchos de los pesos son igualados a cero ya que  $a_{ij}w_{ij} = 0$  si  $a_{ij} = 0$ . En este caso la topología es libre de escala pero los pesos de la red no necesariamente reflejan un sistema que ha evolucionado manteniendo dicha topología. Para estudiar el efecto de la topología en redes asociativas como la de Hopfield, sería conveniente generar modelos que tomen en cuenta la memoria de la estructura de pesos. Al final de cuentas, como establecen Boccaletti *et al.* (2006) “*La estructura de una red real es el resultado de la continua evolución de las fuerzas que la formaron, y ciertamente afecta la función del sistema*”.

### **Estructura y potencial evolutivo.**

En un sistema dinámico las restricciones impuestas por la estructura determinan en gran medida su trayectoria. En una red de Hopfield estas restricciones están dadas por la naturaleza de la función de activación y por la matriz de pesos, ambas determinan el número y tipo de patrones que la red es capaz de recuperar. La estructura, por lo tanto, es fuente de información que evita que el sistema se comporte de manera aleatoria. Pero para que un sistema pueda evolucionar la estructura no puede actuar únicamente como canalizadora sino también como facilitadora de cambio. Precisamente de esto se trata el potencial evolutivo.

Maynard Smith y Eörs Szathmáry consideran que el incremento en complejidad que se observa en los seres vivos está ligado a un cambio en la manera en la que procesan la información y la transmiten de una generación a otra (citado en Collier, 2008). Por lo tanto para que un sistema pueda evolucionar la estructura debe ser capaz de establecer nuevas formas de procesar la información, posibilitando al sistema ganar grados de libertad.

La manera en que las estructuras se organizan para permitir nuevas formas de procesar la

información está lejos de entenderse, sin embargo existen algunas teorías que podrían servir como base para comprender este fenómeno. Una de ellas es la teoría del flujo de información (también conocida como teoría de canales), formulada en 1997 por Jon Barwise y Jerry Selligman (Barwise y Selligman, 1997). De acuerdo con Barwise y Selligman la información presupone un sistema de clasificación que no depende de un vocabulario particular y el cual consiste en un conjunto de objetos, un conjunto de tipos de objetos y una relación binaria entre ambos como la de pertenencia (Barwise y Selligman, 1997). Los diferentes objetos de una clasificación pueden verse como el mismo sistema físico a lo largo de distintos puntos en el tiempo y los tipos como las descripciones del estado del sistema (Ostalé, 2010).

La teoría del flujo de información parte del concepto de infomorfismo. Un infomorfismo es una relación entre dos clasificaciones que permite el flujo de información entre ambas (Barwise y Selligman, 1997). Aunque la definición formal de infomorfismo está fuera de esta discusión, lo que nos dice es que cuando existe un infomorfismo entre dos estructuras (en este caso estructuras relacionales), las características informacionales que posee un objeto en una estructura se preservan en la otra. Desde el punto de vista biológico, esto sugiere que dos estructuras que son físicamente distintas, podrían ser informacionalmente análogas.

Cuando von Bertalanffy planteó su teoría general de sistemas, él consideraba que éstos debían presentar características comunes que obedecían principios generales a nivel de sistema, a las que denominó isomorfismos (von Bertalanffy, 1976). En biología, por ejemplo, los principios de la aerodinámica hacen que los organismos que han evolucionado para volar presenten similitudes estructurales; estos organismos son por lo tanto isomorfos. Los isomorfismos, también llamados analogías u homoplasias, no son considerados por la sistemática moderna ya que no reflejen relaciones de ancestría y descendencia, sin embargo esto no quiere decir que las agrupaciones que inducen no sean "naturales", simplemente nos hablan de una relación distinta a la genalógica. De la misma manera, un infomorfismo nos

hablaría de restricciones informacionales que servirían para agrupar a los sistemas (vivos y no vivos) de acuerdo a la lógica interna de cómo sus partes actúan en conjunto para facilitar el flujo de información.

Además, los infomorfismos pueden agruparse para formar canales de información que vinculen una regularidad con otra, ya que es debido a estas regularidades que la información de un componente puede llevar información de otros componentes (Collier, 2008), tal como la posición de un *switch* lleva información al observador del estado de activación del bulbo al que está conectado, o el estado de activación de un gen lleva información acerca de un fenotipo y viceversa. Por esta razón a estos sistemas se les llama "distribuidos", es decir, sistemas cuyos componentes se informan unos a los otros mediante canales de información. En términos biológicos esto sugiere que deben existir otras fuentes de información que no son genéticas ni tampoco ambientales, sino que son propias de la organización interna del individuo (Collier, 2008). En la medida que la información fluye por estos canales, sería completada hasta construir el fenotipo. Es probable que dichos canales de información existan. Hoy día se sabe que en el proceso de desarrollo se dan un gran número de interacciones, desde interacciones químicas a nivel genético, hasta interacciones físicas y geométricas a nivel celular, que son integradas a nivel de sistema (Álvarez-buylla *et al.*, 2016).

Finalmente, la razón por la que la evolución y el aprendizaje se asemejan, radica en que ambos procesos presentan características que son informacionalmente similares. Esto naturalmente no quiere decir que la evolución biológica y el aprendizaje que solemos reconocer, ligado a la cognición humana, sean equivalentes, sino que ambos procesos comparten principios que están más extendidos en la naturaleza de lo que se cree. La particularidad de la cognición humana se debe a la estructura del sistema nervioso y su operar, que vincula aspectos de la fenomenología "orgánica" como la percepción que ocurre en un dominio físico, con aspectos de una fenomenología más bien abstracta como el razonamiento y la compren-

sión, que ocurren en un dominio al que podríamos llamar “mental”. Cabe señalar que para este dominio mental, si bien tiene su propia estructura y organización, su existencia es una proyección del sistema físico que le da origen y en cuya realidad material influye (y sin el cual dicha proyección no tiene sentido).

## Sección 5: Conclusiones.

En este trabajo se mostró que existen similitudes entre el aprendizaje y la evolución. Dadas estas similitudes, se exploró el tipo de pistas o conjeturas que modelos bien conocidos de redes neuronales como la red de Hopfield, pueden brindar al campo de la biología y la evolución.

De las implicaciones biológicas del modelo consideradas a partir de los resultados de las simulaciones, así como de las pruebas realizadas con datos reales, se concluye que:

- El número de tipos celulares sí podría tener una relación con el tamaño del genoma. La manera en la que suelen hacerse las correlaciones entre genoma y complejidad (medida como el número de tipos celulares) podría no ser la más indicada ya que suelen tomar el genoma completo, siendo que no todos los genes están involucrados en la determinación del tipo celular. Además, no todos los organismos parecieran tener las mismas restricciones, como se observa cuando se quita a las plantas de la comparación. Por otro lado, el número de tipos celulares para ciertos organismos podría ser mucho mayor al que se cree como revelan Galvão *et al.* (2010). Por lo tanto, todavía se desconoce mucho acerca de la relación que podría existir entre genoma y complejidad. Modelos como el de Hopfield podrían ser útiles para revelar nuevos

tipos celulares aún no identificados, al mismo tiempo que estudios con un enfoque más genómico, permitirán elaborar nuevas subclasificaciones de los tipos celulares a partir de perfiles completos de expresión genética.

- El grado de correlación de los perfiles de expresión podría tener algún efecto en su estabilidad. El valor de IM que existe entre los perfiles de expresión genética de los distintos tipos celulares aquí estudiados muestra que, si bien éstos no están altamente correlacionados, tampoco son aleatorios. Estos resultados son de esperarse si se piensa que los estados convergen en un ancestro común pero divergen durante el proceso de desarrollo. Siendo que se reconoce a la modularidad como un proceso crucial en la evolución, sería interesante estudiar el papel que ésta tiene en la estabilidad del sistema al disminuir las restricciones impuestas por la corregulación simultánea de un gran número de genes.
- Claramente el modelo de Hopfield es una sobresimplificación de una RRG, sin embargo captura aspectos importantes que vale la pena considerar. Uno de ellos es que las conexiones en una red no representan interacciones físicas entre los componentes. En la matriz de pesos de una RRG modelada como una red de Hopfield estarían implícitos una serie de procesos tanto genéticos como epigenéticos que en conjunto provocan que la expresión de un gen específico influya en la de otro. Aunque la mayoría de los estudios de este tipo, emplean la matriz de adyacencia  $A$ , para generar la topología deseada en redes de Hopfield, esta metodología no captura el verdadero efecto que la topología pudiera tener en el sistema, ya que prácticamente no toma en cuenta la información de los pesos.
- Con respecto a la estructura y el potencial evolutivo, los modelos de redes dinámicas podrían ser útiles ya que éstos son representaciones de la estructura de un sistema que



mapean con ciertos comportamientos. La estructura determina el comportamiento del sistema gracias a que la estructura es fuente de información que restringe su trayectoria. Así pues, el potencial evolutivo debe abordarse como una capacidad de los sistemas biológicos de establecer nuevas formas de procesar y almacenar información que le permitan ganar grados de libertad.

- Es necesario crear un puente entre la teoría de la información y la teoría de sistemas, ya que entender la relación entre estructura e información permitirá a su vez entender por qué la evolución, vista como un proceso en el que las estructuras almacenan información de su pasado, presenta similitudes con el aprendizaje. Esto naturalmente no quiere decir que los seres vivos (y sobre todo los sistemas de regulación genética) y las redes neuronales aprendan exactamente de la misma forma, sino que deben existir principios generales para el aprendizaje que cualquier sistema que exhiba este comportamiento debe seguir.

Resultados como los que aquí se muestran van acorde con la idea de que existen restricciones estructurales que pueden ser estudiadas mediante modelos sencillos que tomen en cuenta aspectos como la memoria, la exploración y la adaptación.

## Referencias

- Abbot, E. S. (1916). The Causal Relations between Structure and Function in Biology. *The American Journal of Psychology*, 27(2):245–250.
- Albert, R. y Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97.
- Almaas, E., Holder, A. G., y Livingstone, K. D. (2008). An Introduction to Systems Biology

- for Mathematical Programmers. En Lim, G. J. y Lee, E. K., editors, *Optimization in Medicine and Biology*, chapter 1, pages 331–354. Auerbach Publications, Nueva York.
- Álvarez-buylla, E. R., Dávila-velderrain, J., y Martínez-garcía, J. C. (2016). Systems Biology Approaches to Development beyond Bioinformatics : Nonlinear Mechanistic Models Using Plant Systems. *BioScience*, 66(5):371–383.
- Aravind, L., Anantharaman, V., y Venancio, T. M. (2009). Apprehending multicellularity: regulatory networks, genomics, and evolution. *Birth defects research. Part C, Embryo today : reviews*, 87(2):143–64.
- Balleza, E., Álvarez-Buylla, E. R., Chaos, A., Kauffman, S., Shmulevich, I., y Aldana, M. (2008). Critical Dynamics in Genetic Regulatory Networks: Examples from Four Kingdoms. *PLoS ONE*, 3(6):e2456.
- Barrio, R. Á., Hernandez-Machado, A., Varea, C., Romero-Arias, J. R., y Álvarez-Buylla, E. R. (2010). Flower development as an interplay between dynamical physical fields and genetic networks. *PloS ONE*, 5(10):e13523.
- Barwise, J. y Selligman, J. (1997). *Information Flow: The Logic of Distributed System*. Cambridge University Press, Cambridge.
- Beale, R., Jackson, T., y Hilger, A. (1990). *Neural Computing : An Introduction*. Adam Hilger, Bristol.
- Belykh, I., di Bernardo, M., Kurths, J., y Porfiri, M. (2014). Evolving dynamical networks. *Physica D*, 267:1–6.
- Bergman, A. y Siegal, M. L. (2003). Evolutionary Capacitance As a General Feature of Complex Gene Networks. *Nature*, 424:549–552.
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., y Hwang, D. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424:175–308.
- Breitenbach, A. (2009). Teleology in Biology: A Kantian Perspective. *Kant Yearbook*, 1:31–56.
- Brigandt, I. (2015). From Developmental Constraint to Evolvability: How Concepts Figure in Explanation and Disciplinary Identity. En Love, A. C., editor, *Conceptual Change in Biology: Scientific and Philosophical Perspectives on Evolution and Development*, chapter IV, pages 305–325. Springer Netherlands.
- Chalmers, D. J. (1990). The evolution of learning: An experiment in genetic connectionism. En Touretzky, D. S., Elman, J. L., Sejnowski, T. J., y Hinton, G. E., editors, *Proceedings of the 1990 Connectionist Models Summer School*, pages 1–20. Morgan Kaufmann, San Mateo, California.

- Chaos, A. (2010). *Importancia biológica y evolutiva de la estructura y de la dinámica de redes genéticas pequeñas: modelos de simulación inspirados en redes reales*. Tesis doctoral, Universidad Nacional Autónoma de México.
- Collier, J. (2008). Information in Biological Systems. *Philosophy of Information*, (November 2013):763–787.
- Davey, N., Christianson, B., y Adams, R. (2004). High Capacity Associative Memories and Small World Networks. *2004 IEEE International Joint Conference on Neural Networks*, 4:117–182.
- Dominguez, D., Koroutchev, K., Serrano, E., y Rodriguez, F. B. (2007). Information and topology in attractor neural networks. *Neural Computation*, 19(4):956–973.
- Dragui, J. y Wagner, G. P. (2009). The evolutionary dynamics of evolvability in a gene network model. *Journal of Evolutionary Biology*, 22:599–611.
- Elizalde, E., Gómez, S., y Romeo, A. (1992). Maximum overlap neural networks for associative memory. *Physics Letters A*, 170:95–98.
- Fard, A. T., Srihari, S., Mar, J. C., y Ragan, M. A. (2016). Not just a colourful metaphor: modelling the landscape of cellular development using Hopfield networks. *npj Systems Biology and Applications*, 2(1):16001.
- Fernández, N., Maldonado, C., y Gershenson, C. (2013). Information Measures of Complexity, Emergence, Self-organization, Homeostasis, and Autopoiesis. *Complexity*, pages 1–35.
- Fernando, C. T., Liekens, A. M., Bingle, L. E. H., Beck, C., Lenser, T., Stekel, D. J., y Rowe, J. E. (2009). Molecular circuits for associative learning in single-celled organisms. *J. R. Soc. Interface*, 6:463–469.
- Galvão, V., Miranda, J. G. V., Andrade, R. F. S., Andrade, J. S., Gallos, L. K., y Makse, H. a. (2010). Modularity map of the network of human cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 107(13):5750–5.
- Geard, N. y Wiles, J. (2005). A Gene Network Model for Developing Cell Lineages. *Artif Life*, 11:249–267.
- Gershenson, C. (2003). Artificial neural networks for beginners. *arXiv preprint cs/0308031*.
- Goudarzi, A., Teuscher, C., Gulbahce, N., y Rohlf, T. (2012). Emergent criticality through adaptive information processing in boolean networks. *Physical Review Letters*, 108(12):128702.

- Hopfield, J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558.
- Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130:195–204.
- Jin, J. P., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J. C., y Gao, C. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research*, 45(D1):D1040–D1045.
- Jost, J. (2007). Dynamical networks. En Feng, J., Jost, J., y Qian, M., editors, *Networks: from biology to theory*. Springer.
- Karlebach, G. y Shamir, R. (2008). Modelling and analysis of gene regulatory networks. *Nat Rev Mol Cell Biol*, 9(10):770–780.
- Kim, B. J. (2004). Performance of networks of artificial neurons: the role of clustering. *Physical Review E*, 69(4):045101.
- Kirschner, M. y Gerhart, J. (1998). Evolvability. *Proceedings of the National Academy of Sciences of the United States of America*, 95(18):10712–10717.
- Kitano, H. (2001). Systems Biology: Towards System-level Understanding of Biological Systems. En Kitano, H., editor, *Foundations of Systems Biology*, pages 1–36. The MIT Press, Cambridge, Massachusetts.
- Knüpfner, C. y Beckstein, C. (2013). Function of dynamic models in systems biology : linking structure to behaviour. pages 1–9.
- Kröse, B. y van der Smagt, P. (1996). *An introduction to Neural Networks*. The University of Amsterdam, eighth edi edition.
- Krupp, M., Marquardt, J. U., Sahin, U., Galle, P. R., Castle, J., y Teufel, A. (2012). RNA-Seq Atlas—a reference database for gene expression profiling in normal tissue by next-generation sequencing. *Bioinformatics*, 28(8):1184–1185.
- Labiose, C. L., Salah, A. A., y Starikova, I. (2002). The impact of connectivity on the memory capacity and the retrieval dynamics of Hopfield-type networks. En *Proc. of the Santa Fe Complex Systems Summer School*. Santa Fe.
- Latham, P. E. y Roudi, Y. (2009). Mutual information. *Scholarpedia*, 4(1):1658.
- Levine, M. y Davidson, E. H. (2005). Gene regulatory networks for development. *Proceedings of the National Academy of Sciences*, 102(14):4936–4942.

- Löwe, M. (1998). On the storage capacity of Hopfield models with correlated patterns. *The Annals of Applied Probability*, 8(4):1216–1250.
- Löwe, M. y Vermet, F. (2015). Capacity of an associative memory model on random graph architectures. *Bernoulli*, 21(3):1884–1910.
- Maetschke, S. R. y Ragan, M. A. (2014). Characterizing cancer subtypes as attractors of Hopfield networks. *Bioinformatics*, 30(9):1273–1279.
- Maturana, H. y Varela, F. (1984). *El árbol del conocimiento*. Editorial Universitaria, Santiago de Chile.
- McEliece, R., Posner, E., Rodemich, E., y Venkatesh, S. (1987). The Capacity of the Hopfield Associative Memory. *IEEE Transactions on Information Theory*, 33:461–482.
- McGraw, P. N. y Menzinger, M. (2003). Topology and Computational Performance of Attractor Neural Networks. *Physical Review E*, 68(4):047102.
- Morelli, L. G., Abramson, G., y Kuperman, M. N. (2004). Associative memory on a small-world neural network. *The European Physical Journal B*, 38(3):495–500.
- Ostalé, J. (2010). Channel Theory.
- Perotti, J. I., Tamarit, F. A., y Cannas, S. A. (2006). A scale free neural network for modeling neurogenesis.
- Perumal, S. (2008). Stability and Switchability in Recurrent Neural Networks. *College of Engineering*.
- Pigliucci, M. (2008). Is evolvability evolvable?
- Rojas, R. (1996). *Neural Networks - A Systematic Introduction*. Springer-Verlag Berlin Heidelberg.
- Rothstein, J. (1952). Organization and entropy. *Journal of Applied Physics*, 23:1281–1282.
- Schad, E., Tompa, P., y Hegyi, H. (2011). The relationship between proteome size, structural disorder and organism complexity. *Genome Biology*, 12(12):R120.
- Schwartz, S. (1999). The Definition of Cell Type. *Circulation Research*, 84:1234–1235.
- Stauffer, D., Aharony, A., da Fontoura Costa, L., y Adler, J. (2003). Efficient Hopfield pattern recognition on scale free network. *Eur. Phys. J. B.*, 32(3):395–399.
- Szathmáry, E., Jordán, F., y Pál, C. (2001). Molecular Biology and Evolution: Can Genes Explain Biological Complexity? *Science*, 292(5520):1315–1316.

- Szedlak, A., Paternostro, G., y Piermarocchi, C. (2014). Control of asymmetric hopfield networks and application to cancer attractors. *PLoS ONE*, 9(8).
- Torres, J. J., de Franciscis, S., Johnson, S., y Marro, J. (2010). Excitable Networks: Nonequilibrium Criticality and Optimum Topology. *International Journal of Bifurcation and Chaos*, 20(3):869–975.
- Torres, J. J., Muñoz, M. A., Marro, J., y Garrido, P. L. (2004). Influence of topology on the performance of a neural network. *Neurocomputing*, 58(60):229–234.
- Trewavas, A. (2006). A Brief History of Systems Biology. *The Plant Cell*, 18(10):2420–2430.
- Valiant, L. G. (1984). A Theory of the Learnable.
- Valiant, L. G. (2009). Evolvability. *J. Assoc. Computing Machinery*, 56(1):1–21.
- Vickaryous, M. K. y Hall, B. K. (2006). Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biological reviews of the Cambridge Philosophical Society*, 81(3):425–55.
- von Bertalanffy, L. (1976). *Teoría general de los sistemas: fundamentos, desarrollo, aplicaciones*. Fondo de Cultura Económica, México.
- Wagner, G. P. y Altenberg, L. (1996). Complex Adaptations and The Evolution of Evolvability. *Evolution*, 50:1967–1976.
- Wagner, G. P., Pavlicev, M., y Cheverud, J. M. (2007). The road to modularity. *Nature Reviews Genetics*, 8:921–931.
- Wang, K. y Nishida, H. (2015). REGULATOR: a database of metazoan transcription factors and maternal factors for developmental studies. *BMC Bioinformatics*, Abril 10(16):114.
- Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development*, 4(January):66–82.
- Watson, R. A., Buckley, C., y Mills, R. (2011a). Optimization in "Self-modeling" Complex Adaptive Systems. *Complexity*, 16(5):17–26.
- Watson, R. A., Buckley, C., Mills, R., y Davies, A. (2010). Associative memory in gene regulation networks. En *Artificial Life XII: Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems*, pages 194–202. MIT Press.
- Watson, R. A., Mills, R., y Buckley, C. (2011b). Global Adaptation in Networks of Selfish Components: Emergent Associative Memory at the System Scale. *Artificial Life*, 17(3):147–166.

- Watson, R. A. y Szathmary, E. (2016). How Can Evolution Learn? *Trends in Ecology & Evolution*, 31(2):147–157.
- Watson, R. A., Wagner, P., G., Pavlicev, M., Weinreich, D. M., y Mills, R. (2014). The evolution of phenotypic correlations and 'developmental'. *Evolution*, 68(4):1124–1138.
- Weaver, D. C. (1999). Pacific Symposium on Biocomputing 4:112-123 (1999). 123:112–123.
- Wilson, D., Charoensawan, V., Kummerfeld, S. K., y Teichmann, S. A. (2008). DBD - Taxonomically broad transcription factor predictions: New content and functionality. *Nucleic Acids Research*, 36(D1):D88–92.
- Wray, N. y Visscher, P. (2008). Estimating Trait Heritability. *Nature Education*, 1(29).
- Wuensche, A. (2004). Basins of attraction in network dynamics: A conceptual framework for biomolecular networks. En Schlosser, G. y Wagner, G., editors, *Modularity in developmental and evolution*, pages 288–311. Chicago University Press.
- Wuensche, A. (2010). Complex and Chaotic Dynamics, Basins of Attraction, and Memory in Discrete Networks. *Acta Physica Polonica B*, 3(2):463–478.
- Xu, C. y Su, Z. (2015). Identification of cell types from single-cell transcriptomes using a novel clustering method. *Bioinformatics*, 31(12):1974–1980.