



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

EXPLORACIÓN DE LAS PROPIEDADES QUE DELIMITAN EL  
PAISAJE ORGANIZACIONAL DE REDES DE REGULACIÓN EN  
BACTERIAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

BIÓLOGO

P R E S E N T A :

CARLOS ROBERTO CRUZ MALDONADO

DIRECTOR DE TESIS:

DR. JULIO AUGUSTO FREYRE-GONZÁLEZ



CIUDAD UNIVERSITARIA, CD. MX. 2017



Universidad Nacional  
Autónoma de México



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

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

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

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

**Exploración de las propiedades que delimitan el paisaje  
organizacional de redes de regulación en bacterias**

por

Carlos Roberto Cruz Maldonado

Tesis presentada para obtener el grado de

Biólogo

en la

FACULTAD DE CIENCIAS

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

CIUDAD UNIVERSITARIA, CD. MX. 2017

Dedico esta obra a mis padres,  
A mis padres académicos y  
A mis cachorras, Nova y Maya

# Agradecimientos

Antes que nada, quiero agradecer a mi tutor Julio Freyre, por haber estado conmigo en cada paso de la realización de este trabajo. Sin sus consejos e invaluable guía tanto en lo académico como en lo personal, no hubiera concluido la escritura de mi tesis. Todavía sigo aprendiendo y trabajando en varios aspectos, sin embargo, han tomado un gran papel en mi formación como persona y en lo profesional, como futuro investigador. Gracias Julio por apoyarme tanto, creer en mí desde el principio, así como de ser un gran guía y amigo.

A mi familia le debo desde el principio el haberme inculcado esas ganas de ser universitario y de llegar a ser puma, como mi papá y mi abuelo. Papá y Mamá constantemente me han dado todo el apoyo que un hijo podría pedir. Yo sé que es difícil mantener la comunicación y pasar tanto tiempo a la distancia; sin su ejemplo yo no me hubiera mantenido firme fuera de nuestra tierra ni hubiera obtenido tantos logros, como precisamente éste y los que faltan por venir. Los amo, y cada día que paso fuera de la península trato de dar lo mejor de mí, porque como un día dijo mi abuelo “Yo no mido mi éxito por los logros personales, sino por los logros de mis hijos”.

Yo realmente pienso que el disfrutar del éxito, es ver triunfar y realizándose a las personas a quienes uno brindo tiempo y formación.

A mis padres académicos principalmente Pablo Padilla y Claudia Segal, les debo todo mi avance e inspiración durante la carrera, me apoyaron muchísimo con su amistad y su tutoría. Me adentraron al campo de la biología que me terminaría atrapando y me enseñaron más que cosas académicas. Pablo, gracias por siempre apoyarme, creer en mí, escucharme y compartirme tanto de música y de vida. Claudia, gracias por ser mi tutora los últimos años de carrera y de tanta paciencia y regaños, verás que si rindieron fruto.

A todos mis amigos Kevin, Gabriel, Hani, Luis, Juan, Silvia, Wilo, Arturo, Ivan, Alissa, Fani, Ilse, Lorenia, Carolina, han sido y son mi mayor apoyo moral en todo este recorrido. Aunque en algunos casos estemos a kilómetros de distancia los llevo conmigo siempre. Gracias por estar siempre para mí y compartir tantas cosas. Leslie, gracias, creo que sin tu apoyo los últimos años de carrera, la motivación y todo lo que compartimos no hubiera llegado hasta acá. Nunca me hubiera animado a venir por mi cuenta a Cuernavaca, por lo tanto, no hubiera conocido lo que ahora realmente amo de la biología. Te debo muchísimo, no solo lo que aprendí de ti y de Dora del amor a la ciudad y a la universidad, del amor perruno y muchas cosas más. Dora, gracias por ser mi segunda mamá y apoyarnos tanto.

Y por último a mis hermosas cachorras perricolas, Nova y Maya, sin su compañía, besitos mientras escribía, la alegría de cada día y en cada situación, calentando mis pies mientras leía, caminatas interminables y el apoyo moral y tanto amor que me brindaron y me siguen brindando aun en los momentos más difíciles que solo ustedes me vieron vivir, no lo hubiera logrado sin la calma que ustedes me brindan. Las amo bonitas.

En lo económico:

Sin el enorme esfuerzo económico que han estado realizando mis padres durante estos ya siete años fuera de casa, no podría haberme mantenido aquí realmente enfocado en estar realizando mis sueños.

Este trabajo fue financiado por los donativos del PAPIIT-UNAM IA200615 y IA200616 otorgados al Dr. Julio Augusto Freyre González.

Y, por último, como la frase que escribió Newton a Hooke en 1676, “*Si he visto más lejos es porque estoy sentado sobre los hombros de gigantes.*” Mis padres biológicos y mis padres académicos son mi inspiración, mi ejemplo y los gigantes que me han traído hasta donde estoy parado hoy en día. Gracias.

# Índice general

Agradecimientos .....	ii
Resumen.....	0
Abstract .....	1
Introducción .....	2
Teoría de grafos.....	3
Algunos conceptos básicos de teoría de grafos:.....	5
Propiedades de las redes .....	8
Modelos teóricos de redes.....	10
Modelo de redes aleatorias .....	11
Modelo de redes libres de escala.....	12
Modelo de redes jerárquicas .....	13
Redes biológicas.....	14
Paisaje organizacional de redes de regulación .....	14
Hacia el estudio de la organización y propiedades de redes de regulación .....	17
Enfoque de descomposición natural.....	18
Auto-similitud.....	21
¿Qué es un fractal? .....	21
Dimensión fractal .....	22
Método de box-covering.....	24
Resultados y discusión .....	25
Obteniendo las dimensiones fractales de redes de regulación .....	25
Implementación del método de box-covering.....	25
Optimización del algoritmo de box-covering.....	26
Dimensión fractal en modelos de redes biológicas .....	28
Obtención de redes de regulación bacterianas .....	28
Cálculo de dimensión fractal.....	29
Análisis de la dimensión fractal a través del tiempo.....	30
Dimensión fractal en modelos de redes teóricas.....	31
Generación de redes aleatorias .....	32
Generación de redes libres de escala.....	32
Cálculo de dimensión fractal de redes teóricas .....	32

Análisis de las distribuciones de dimensión fractal.....	32
Paisaje organizacional .....	35
Conclusiones y perspectivas.....	37
Conclusiones .....	37
Perspectivas .....	40
Referencias:.....	41

# Índice de figuras

FIGURA 1: ILUSTRACIÓN DE LA CIUDAD DE KÖNIGSBERG (IZQUIERDA) EN LA CUAL SE OBSERVAN LOS SIETE PUENTES, SOBREPUESTA SOBRE LA ILUSTRACIÓN SE ENCUENTRAN LOS POSIBLES RECORRIDOS A TRAVÉS DE LOS PUENTES. REPRESENTACIÓN SIMPLIFICADA COMO GRAFO DE LOS PUENTES DE KÖNIGSBERG (DERECHA). .....	4
FIGURA 2: REPRESENTACIÓN DE LAS INTERACCIONES REGULATORIAS. EN EL MODELO (IZQUIERDA) OBSERVAMOS CÓMO LA SECUENCIA DE UN GEN SE TRADUCE EN UN PRODUCTO GÉNICO Y A SU VEZ ESTE PRODUCTO GÉNICO ACTÚA SOBRE OTRO GEN. FINALMENTE, EN LA RED DE REGULACIÓN (DERECHA) ÚNICAMENTE REPRESENTAMOS LA CAUSALIDAD DE LAS INTERACCIONES. ....	9
FIGURA 3 REPRESENTACIÓN DE UNA RED DE REGULACIÓN EN LA CUAL LOS GENES SON REPRESENTADOS POR NODOS Y LAS INTERACCIONES REGULATORIAS POR ARISTAS. ADEMÁS DE REPRESENTAR LA INTERACCIÓN, EN UNA RED DE REGULACIÓN SE PUEDE REPRESENTAR LA DIRECCIÓN DE LA INTERACCIÓN VOLVIENDO FLECHAS LAS ARISTAS.....	9
FIGURA 4: MODELOS TEÓRICOS DE REDES Y SUS PROPIEDADES DISTINTIVAS. (A) REDES ALEATORIAS. (B) REDES LIBRES DE ESCALA (SCALE-FREE). (C) REDES JERÁRQUICO-MODULARES. SUS PROPIEDADES CARACTERÍSTICAS ESTÁN REPRESENTADAS DEBAJO DE LOS MODELOS (B) DISTRIBUCIÓN DE CONECTIVIDAD, $P(k)$ , Y (C) DE COEFICIENTE DE AGRUPAMIENTO, $C(k)$ . IMAGEN TOMADA DE (BARABASI AND OLTVAI 2004). .....	11
FIGURA 5: PAISAJE ORGANIZACIONAL. EN ESTA REPRESENTACIÓN SIMPLIFICADA DEL PAISAJE ORGANIZACIONAL SOLO TENEMOS DOS DIMENSIONES, ES DECIR DOS PROPIEDADES CONTRASTANDO SUS RANGOS. EL ESPACIO VERDE QUE SE PUEDE OBSERVAR ES EL ÁREA DE POSIBLES VALORES PARA ESAS DOS PROPIEDADES QUE PUEDEN OCUPAR LAS REDES.....	15
FIGURA 6: LA PROPIEDAD DE DENSIDAD EN REDES DE REGULACIÓN. (IZQUIERDA) REPRESENTACIÓN DE UN GRAFO CON DIFERENTES DENSIDADES, DE ARRIBA HACIA ABAJO SE MUESTRA CÓMO VA DECRECIENDO LA DENSIDAD. (DERECHA) DENSIDAD DE DIFERENTES REDES DE REGULACIÓN GRAFICADAS CON SU RESPECTIVA COBERTURA GENÓMICA.....	16
FIGURA 7: ENFOQUE DE DESCOMPOSICIÓN NATURAL. DESCRIPCIÓN DE LA SERIE DE PASOS QUE CONFORMAN LA DESCOMPOSICIÓN NATURAL DE REDES DE REGULACIÓN Y CON LA CUAL SE OBTIENE LA CLASIFICACIÓN DE LOS NODOS DE LA RED. IMAGEN ADAPTADA DE (IBARRA-ARELLANO, CAMPOS-GONZALEZ ET AL. 2016). .....	20
FIGURA 8: REDES DE REGULACIÓN Y FRACTALES, ESTRUCTURAS ANIDADAS. (A) RED DE REGULACIÓN DE BACILLUS SUBTILIS. (B) RED DE REGULACIÓN DE ESCHERICHIA COLI. EN AMBAS REDES DE REGULACIÓN LOS MÓDULOS FUERON REDUCIDOS A UN SOLO NODO, LAS FLECHAS DE COLORES REPRESENTAN LAS INTERACCIONES REGULATORIAS DE LOS REGULADORES GLOBALES MOSTRANDO EL EFECTO DE SU INTERACCIÓN. IMAGEN TOMADA DE (FREYRE-GONZALEZ, TREVINO-QUINTANILLA ET AL. 2012). (C) TRIANGULO DE SIERPINSKI. EN LAS TRES FIGURAS SE OBSERVAN ESTRUCTURAS ANIDADAS LO CUAL AL COMPARAR LAS INTERACCIONES ANIDADAS CON UNA FIGURA FRACTAL, SURGE LA NOCIÓN DE LA AUTO-SIMILITUD EN REDES DE REGULACIÓN. ....	20
FIGURA 9 CAMBIO DE LA DIMENSIÓN FRACTAL DE LA RED DE E. COLI CURADA A TRAVÉS DEL TIEMPO .....	30
FIGURA 10 GRAFICAS DE CAJA Y BIGOTE QUE MUESTRAN LA DISTRIBUCION DE DIMENSION FRACTAL PARA LAS DIFERENTES REDES ANALIZADAS. ERDOS-RENYI (ER), REDES BIOLÓGICAS (RN), Y BARABASI-ALBERT (BA), PRUEBA U DE MANN-WHITNEY (MWU) Y KOLMOGOROV-SMIRNOV (KS) .....	33
FIGURA 11: HEATMAP Y DENDOGRAMA CON CLUSTERING DE PROPIEDADES. EN EL EJE Y TENEMOS LAS PROPIEDADES CALCULADAS PARA TODAS LAS REDES DE REGULACION. EN EL EJE X ESTÁN LOS IDENTIFICADORES DE LAS REDES DE REGULACION POR SU TAXID DE NCBI ORDENADOS POR LA COBERTURA GENOMICA DE LAS REDES. ....	36

# Resumen

Las redes de regulación son altamente plásticas y evolucionan incluso más rápido que los genes estructurales. Sin embargo, ha sido poco explorada la manera en que la evolución moldea la arquitectura funcional de las redes de regulación y sus sistemas. Para adentrarse en la solución de esta problemática se requiere una gran cantidad de redes regulatorias reconstruidas para diferentes organismos. Abasy Atlas, una base de datos desarrollada por nuestro grupo de trabajo, provee una extensa colección de redes de regulación de una diversidad de bacterias. Un análisis previo de estos datos muestra que no solo las distribuciones de conectividades y de coeficiente de agrupamiento de las redes siguen un comportamiento de leyes de potencias, sino que también la distribución de tamaño de sus módulos, tal como los identifica el método de descomposición natural (NDA, por sus siglas en inglés), presenta el mismo comportamiento. La aparición recurrente de un comportamiento de leyes de potencia sugiere que la auto-similitud es un principio que gobierna la arquitectura de redes de regulación. La auto-similitud, o invariancia hacia cambios de escala, es una propiedad de muchos sistemas naturales que exhiben características fractales, además, la fractalidad nunca ha sido estudiada en redes de regulación de bacterias e incluso ha habido controversia acerca de su posible existencia. Por otro lado, hemos encontrado que la densidad, una propiedad de las redes, se encuentra fuertemente conservada a lo largo de las redes de regulación, llevando a preguntarnos qué otras propiedades o parámetros topológicos se encuentran conservados y cómo su limitado rango de posibles valores delimitan el paisaje organizacional de las redes de regulación. En este trabajo se explora la posible naturaleza fractal de 50 redes de regulación bacterianas, reconstruidas y disponibles en Abasy Atlas. Nosotros implementamos una versión optimizada del algoritmo de *box-covering* para obtener los exponentes de dimensión, es decir la dimensión fractal. En las redes de regulación analizadas, a pesar de sus diferentes niveles de cobertura genómica, encontramos que su dimensión fractal se encuentra conservada y limitada a un corto rango. Nuestros resultados sugieren que la fractalidad, así como la densidad son propiedades organizacionales de las redes de regulación bacterianas que pueden imponer limitaciones a su dinámica evolutiva y a su vez pueden ser explotadas para mejorar los modelos evolutivos que existen actualmente para redes redes.

# Abstract

It is known that regulatory networks (RNs) are highly plastic, evolving faster than structural genes. Nevertheless, the way evolution shapes RNs functional architecture and their systems is poorly explored. In order to address this issue, a wide number of reconstructed RNs for different organisms is required. Abasy Atlas, developed by our group, provides an extensive collection of RNs across bacteria. An analysis of this data shows that not only RNs degree and clustering coefficient distributions follows a power law, but also the module sizes distribution, as identified by the natural decomposition approach, has the same behaviour. The recurrent emergence of a power law behavior could suggest that self-similarity is a principle governing RNs architecture. Self-similarity, invariance against changes in scale, is a property of many natural systems exhibiting fractal features. Although, fractality has never been explored nor observed in RNs and even its existence has been controversial. On the other hand, we have found that density is strongly conserved across RNs, raising the question of which other topological parameters are conserved and how their limited range of possible values delimit the RNs organizational landscape. In this study, we explored the possible fractal nature of 51 bacterial RN reconstructed and available in Abasy Atlas. We implemented an optimized version of the box-covering algorithm to obtain dimension exponents. We found that fractal dimension is conserved and constrained into a small range across all analyzed RNs despite their genomic completeness. Our results suggest that fractality, as well as density, is an organizational property of bacterial RNs that could impose constraints to their evolutionary dynamics and could be exploited to improve our current network evolutionary models.

# Introducción

En el mundo real podemos encontrar infinidad de sistemas complejos que se caracterizan por la confluencia de múltiples procesos cuyas interrelaciones constituyen la estructura de un sistema que funciona como una totalidad (García 2006). La biología de sistemas, la cual estudia diversos sistemas complejos, es una disciplina que tiene como objetivo comprender la complejidad en la biología tratando de reunir sus elementos para entenderlos en su contexto de sistemas.

Debido a que cada sistema presenta una historia evolutiva distinta, es por medio de la aplicación de métodos de ingeniería inversa que podemos develar el proceso evolutivo o serie de pasos que confieren las características que observamos actualmente en un sistema particular. Además, tampoco es sencillo realizar inferencias sobre un sistema a partir de las propiedades de sus partes y las leyes de sus interacciones. Por mucho tiempo se han estado identificando todos los genes y proteínas de diversos organismos, esto es equivalente a estar haciendo una lista de las partes que conforman una máquina. Si bien, esta lista es como tener un catálogo de componentes individuales, por si misma no es suficiente para entender la complejidad que subyace a la máquina (Kitano 2002). Para llegar a entender esta complejidad resulta necesario el conocimiento de cómo estas partes se relacionan para dar su forma y estructura.

El entendimiento a nivel de sistemas de un sistema biológico implica el estudio de su estructura, es decir, estudiar la relación entre sus partes y las características de estas interacciones. Para llevar a cabo un análisis de esta magnitud recurrimos al análisis de redes generadas a partir de estas interacciones (Barabasi and Oltvai 2004). De esta manera podemos decir que, para entender la complejidad en la biología, por ejemplo, tenemos no solo que entender cuál es la función de un gen, sino cómo interactúa uno con otro y con el ambiente.

En un principio el objetivo de este estudio era hacer un análisis sobre las diferentes propiedades que pueden conformar el paisaje organizacional de redes de regulación bacterianas y cómo éstas podrían imponer restricciones sobre su evolución. Sin embargo, al revisar las propiedades que analizaríamos encontramos que la propiedad fractal, que se puede analizar por el método de *box-covering*, no era un cálculo sencillo y aparte no se había implementado un algoritmo óptimo para resolverlo. Por ende, abordamos el problema de la fractalidad en redes de regulación y la implementación de un algoritmo para su análisis.

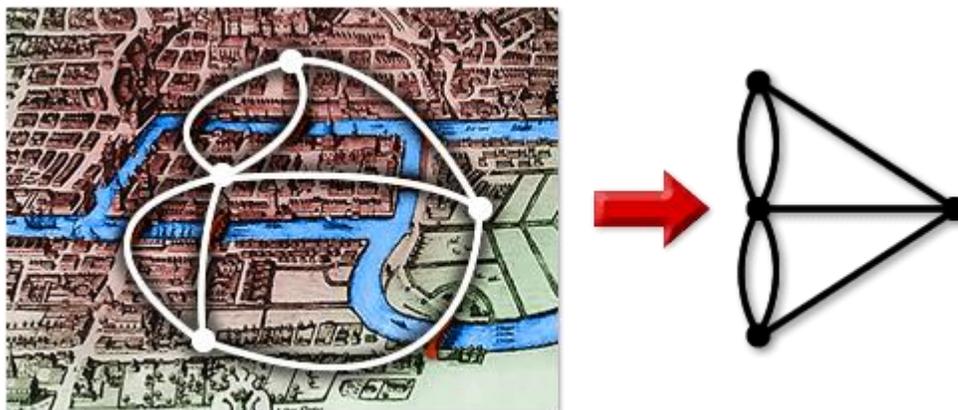
Trabajamos sobre una variedad de redes de regulación bacterianas, las cuales según sus propiedades presentan una característica inherente a los fractales que puede ser analizada, la auto-similitud. La auto-similitud y la fractalidad son propiedades fundamentales de las redes complejas. Al hablar de auto-similitud en una red nos referimos a que hay patrones que se repiten a diferentes escalas, es decir, la estructura global de la red se ve muy parecida a una subsección de sí misma y la presencia de esta propiedad nos lleva a considerar a las estructuras que la presentan como fractales. La auto-similitud y la fractalidad de una red están directamente relacionadas con su estructura interna y organización. Es por esto que es importante un análisis y caracterización fractal de redes para tener una mejor comprensión de los mecanismos generativos del sistema. Este trabajo presenta los primeros pasos hacia ese objetivo.

## **Teoría de grafos**

Para comprender sobre las redes o grafos, situémonos en el siglo XVIII al este de Prusia, en una ciudad llamada Königsberg. A esta ciudad la atraviesa un río llamado Pregel el cual rodea a su vez una isla, Kneiphof, quedando de esta manera dividida la ciudad en cuatro regiones diferentes, las cuales son unidas por siete puentes. Los habitantes de Königsberg como parte del atractivo de su ciudad mantenían vivo un

problema a manera de juego mental “¿Puede uno caminar a través de los siete puentes sin cruzar el mismo dos veces?”.

En este mismo siglo y para ser más exactos en 1736, Leonhard Euler, un matemático suizo, acababa de publicar un artículo respondiendo y probando matemáticamente que no había una solución posible al famoso problema de la ciudad de Königsberg (Euler 1953). En ese momento, Euler no solo estaba dando una respuesta formal al problema de Königsberg, sino que a su vez por la manera de abordar el problema estaba dando nacimiento a un área de las matemáticas ahora conocida como teoría de grafos. Euler visualizó el problema como un grafo (un grupo de nodos conectados por aristas) en el cual las regiones de Königsberg fueron representados como nodos y los puentes las aristas que los unen. Es sobre esta representación simplificada que Euler notó que los nodos que tienen un número impar de aristas deben ser aquellos donde se comience o termine el recorrido de tal manera que se pueda cumplir la condición de que en el recorrido debe de haber sólo un comienzo y un final. Dado que el grafo representando a Königsberg tiene más de dos nodos con esas características, no es posible encontrar un camino que recorra todos los puntos sin pasar dos veces por el mismo puente y esto es una propiedad inherente al grafo .



*Figura 1: Ilustración de la ciudad de Königsberg (izquierda) en la cual se observan los siete puentes, sobrepuesta sobre la ilustración se encuentran los posibles recorridos a través de los puentes. Representación simplificada como grafo de los puentes de Königsberg (derecha).*

En la biología, así como en las matemáticas, la física o las ciencias computacionales y sociales, se pueden tener sistemas como objeto de estudio. La arquitectura o patrón de conexiones entre componentes de un sistema es crucial para entender su comportamiento y esta a su vez se representa como una red. Las redes se encuentran en muchos campos de la ciencia y en la vida diaria, el principal interés en ellas radica en que las podemos ver como una representación simplificada un sistema a una estructura abstracta. Esta abstracción presenta solo los patrones de conexión básicos entre elementos del sistema, así como también información adicional que nos dé más detalles sobre él. Estos elementos son representados por nodos y las conexiones entre ellos por aristas, además, las aristas pueden ser dirigidas o no dirigidas, dependiendo del tipo de relación que representen.

La caracterización de la topología de una red es muy importante para un amplio rango de propiedades estadísticas y dinámicas, un ejemplo inmediato sería la topología de una red social, la cual tiene una gran influencia en la dispersión de información y enfermedad.

Es importante mencionar que en la literatura en general y en este trabajo la mayoría de las veces la utilización de la palabra “red” y “grafo” son conceptos intercambiables. Sin embargo, en la mayoría de los casos se usa “red” cuando queremos enfatizar que estamos hablando de una red del mundo real y “grafo” cuando estamos tomando en cuenta sus propiedades matemáticas (Molontay 2015).

## **Algunos conceptos básicos de teoría de grafos:**

En esta sección presentaré algunas de las definiciones más importantes de la teoría de grafos que se estarán utilizando y tienen relación con este trabajo. Cabe mencionar que entendemos por red un sistema que puede ser modelado por un grafo, y por lo tanto la teoría de redes surge de la teoría de grafos. Algunas definiciones fueron tomadas y adaptadas de Newman (2011), Freyre-Gonzalez (2008) y Molontay

(2015) por lo tanto, si se desea profundizar en este tema pueden referirse a los trabajos anteriormente mencionados, en ellos podrán encontrar introducciones más detalladas al campo, ejemplos y definiciones básicas.

**Grafo.** Es un par ordenado  $G = (N, E)$ , donde  $N$  es el conjunto de nodos junto con un conjunto de aristas o conexiones  $E$ , estas aristas al estar conectando dos nodos cada una pasaría a ser un subconjunto de  $N$ .

**Camino.** Una arista conecta de forma directa dos nodos; de forma similar, una sucesión de aristas, donde el objetivo de la arista previa es el comienzo de la siguiente arista, nos permite trazar un camino que conecta de forma indirecta dos nodos. La longitud de un camino es el número de aristas necesarias para llegar desde un nodo a otro.

**Camino mínimo o geodésico.** Es el camino más corto entre un cierto par de nodos. La longitud geodésica entre dos nodos  $p$  y  $q$  es la distancia  $d(p, q)$  de estos nodos.

**Diámetro de un grafo.** Si calculamos los caminos mínimos entre todos los pares de nodos de un grafo, el diámetro del grafo será igual al camino mínimo de mayor longitud. El diámetro de un grafo  $D(G)$  cumplirá siempre con  $D(G) \leq |N| - 1$ . Se le llama *longitud del camino característico*  $L$  al promedio de las distancias geodésicas del grafo, o sub-grafo.

**Vecindario de un nodo.** El *vecindario- $l$* , representado como  $N_p^l$  de un nodo  $p$  es el conjunto de nodos  $q$  tales que su distancia de  $p$  no sea mayor a  $l$ .

**Adyacencia.** Se dice que dos nodos  $p$  y  $q$  son adyacentes o vecinos si existe una arista que los una.

**Matriz de adyacencia.** Una manera matemática de representar una red o grafo para su análisis es con una matriz. Se puede construir una matriz de adyacencia simplemente con el número de nodos de una red y una lista de sus aristas. En esta matriz cuadrada de  $n \times n$  las filas y las columnas representaran los nodos y a sus intersecciones, si es que hay una arista que una esos dos nodos, se le da el valor de 1, en el caso contrario se le da el valor de 0.

**Grado de un nodo.** En grafos no dirigidos, el grado o conectividad ( $k$ ) de un nodo es el número de aristas que convergen en él. En grafos dirigidos, existen dos grados posibles, el grado de entrada ( $k_{in}$ ) y el grado de salida ( $k_{out}$ ); siendo cada uno, respectivamente, el número de aristas que llegan al nodo y el número de aristas que salen del nodo.

**Red de mundo pequeño (*small-world*).** Una red es llamada así si el tamaño del camino característico  $l$  crece proporcionalmente al logaritmo del número de nodos en la red.

$$L \propto \log |N|$$

**Distribución de grado de conectividad.** Una de las primeras herramientas para analizar la estructura topológica global de una red es la distribución de conectividad,  $P(k)$ , la cual mide la proporción de nodos con grado  $k$  en una red. Es decir, es la probabilidad de que si tomamos un nodo al azar su grado sea igual a  $k$ .

**Red libre de escala (*scale-free*).** Una red es libre de escala si su distribución de grados de conectividad sigue una ley de potencias (*power-law*) es decir,  $P(k)$  es proporcional a una potencia de  $k$ , para algún número  $\gamma \geq 1$ :

$$P(k) \propto k^{-\gamma}$$

Para las redes complejas y en este caso las redes biológicas, el exponente de grado  $\gamma$  generalmente se encuentra entre 2 y 3, es decir  $2 < \gamma < 3$ .

**Coefficiente de agrupamiento.** Está definida por la probabilidad de que dos nodos, vecinos de cierto nodo, sean también vecinos entre ellos. También, es la medida del grado en que los nodos de un grafo tienden a agruparse, y la relación entre elementos que se encuentran conectados por medio de un intermediario, por lo tanto, se encuentran conectados, mas no directamente.

## Propiedades de las redes

A pesar de la gran diversidad de redes que existe, la mayoría de las redes que observamos en el mundo real comparten propiedades específicas que difieren en distintas maneras de las redes aleatorias. Estas propiedades, observables al analizar su estructura (topología), a su vez imponen limitaciones a su dinámica. En biología se hace uso de muchos tipos diferentes de redes, yendo de redes visibles físicamente tales como redes neuronales hasta redes que son representaciones más abstractas como las cadenas alimenticias o redes epidemiológicas en las que se ve la transmisión de una enfermedad. También existen las redes bioquímicas, dentro de las cuales se encuentran las redes metabólicas, de interacción proteína-proteína y de regulación genética.

En las redes de regulación genética, como es el caso de las redes analizadas en este trabajo, los productos génicos son representados por nodos y las interacciones regulatorias entre ellos por aristas (figura 2 y figura 3). Como hemos mencionado anteriormente, analizar la estructura de estas redes es crucial para llegar a la comprensión de la complejidad de la regulación celular. Al analizar la estructura de una red resultaría tal vez lógico, para alguien alejado del campo de teoría de redes, el tomar una imagen de ella y observarla en búsqueda de patrones o características que no podríamos obtener de los datos crudos con los que generamos la red. Sin embargo, esto solo se mantendría, dentro de lo posible, para redes con un pequeño número de nodos y aristas. Las redes de regulación llegan a tener miles de nodos y aristas lo cual

complica su análisis visual, es por esto que se emplean otras técnicas de análisis como las proporcionadas por la teoría de grafos (Bornholdt 2005, Newman 2011).

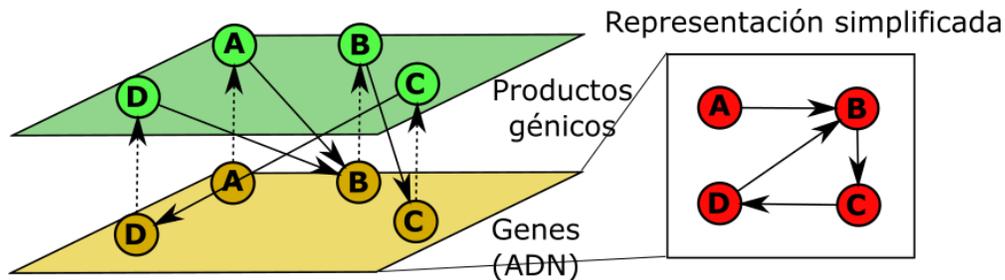


Figura 2: Representación de las interacciones regulatorias. En el modelo (izquierda) observamos cómo la secuencia de un gen se traduce en un producto génico y a su vez este producto génico actúa sobre otro gen. Finalmente, en la red de regulación (derecha) únicamente representamos la causalidad de las interacciones.

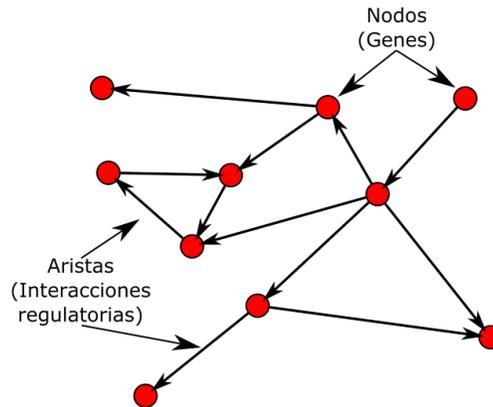


Figura 3 Representación de una red de regulación en la cual los genes son representados por nodos y las interacciones regulatorias por aristas. Además de representar la interacción, en una red de regulación se puede representar la dirección de la interacción volviendo flechas las aristas.

Entre las principales propiedades para analizar y caracterizar redes están el grado de nodo y su distribución de probabilidad o frecuencia. Una de las características fundamentales de las redes complejas es que son libres de escala (*scale-free networks*) es decir la distribución de probabilidad de sus conectividades sigue una ley de potencias (*power-law*), las redes de regulación se clasifican como redes libres de escala. La presencia de esta ley de potencias muestra que existe una relación funcional entre dos magnitudes, donde el cambio relativo en una magnitud resulta en un cambio proporcional relativo en otra magnitud, es decir una magnitud varía como potencia

de la otra. Además, como las leyes de potencia están relacionadas con la invariancia en escala, las leyes de potencia con un exponente escalar particular son equivalentes dado que cada una es una versión escalada de las demás (Newman 2011).

Estas características si bien son mediciones directas de la red, al analizarlas en conjunto revelan una propiedad interesante ya que, desde el grado de conectividad, indican que hay un tipo particular de nodo. Este tipo de nodo que se encuentra en pequeña cantidad es llamado “*hub*” y posee altos grados de conectividad mientras que el resto de nodos, la mayoría, presentan bajos grados de conectividad. La presencia de una pequeña cantidad de nodos “*hub*” a su vez es conocida como la propiedad o característica del mundo pequeño (*small-world*) la cual también es una característica de las redes biológicas. En una red de mundo pequeño los nodos “*hub*” reducen la distancia promedio entre nodos menos conectados lo cual si lo trasladamos a una red de regulación éstos le confieren la ventaja de poder distribuir una señal mucho más rápido (Freyre-Gonzalez 2008).

## **Modelos teóricos de redes**

Si queremos analizar alguna propiedad en particular, una de las mejores maneras de entender y observar los efectos de los diferentes tipos de redes sobre sus propiedades es por medio de la utilización de modelos teóricos de redes. Debido a la gran cantidad de tipos de redes, una manera de comprender su estructura y comportamiento es compararlos con redes artificiales que tengan características tanto similares como distintas, de esta manera establecemos características controladas y podemos explorar sus diferencias.

Otra ventaja que presenta la comparación de redes teóricas con redes reales es que mediante el crecimiento de las primeras con reglas definidas podemos tener un punto de comparación de los diferentes procesos generativos, así también sobre qué

tipo de modelos se aproximan a un modelo nulo de la estructura que estamos observando.

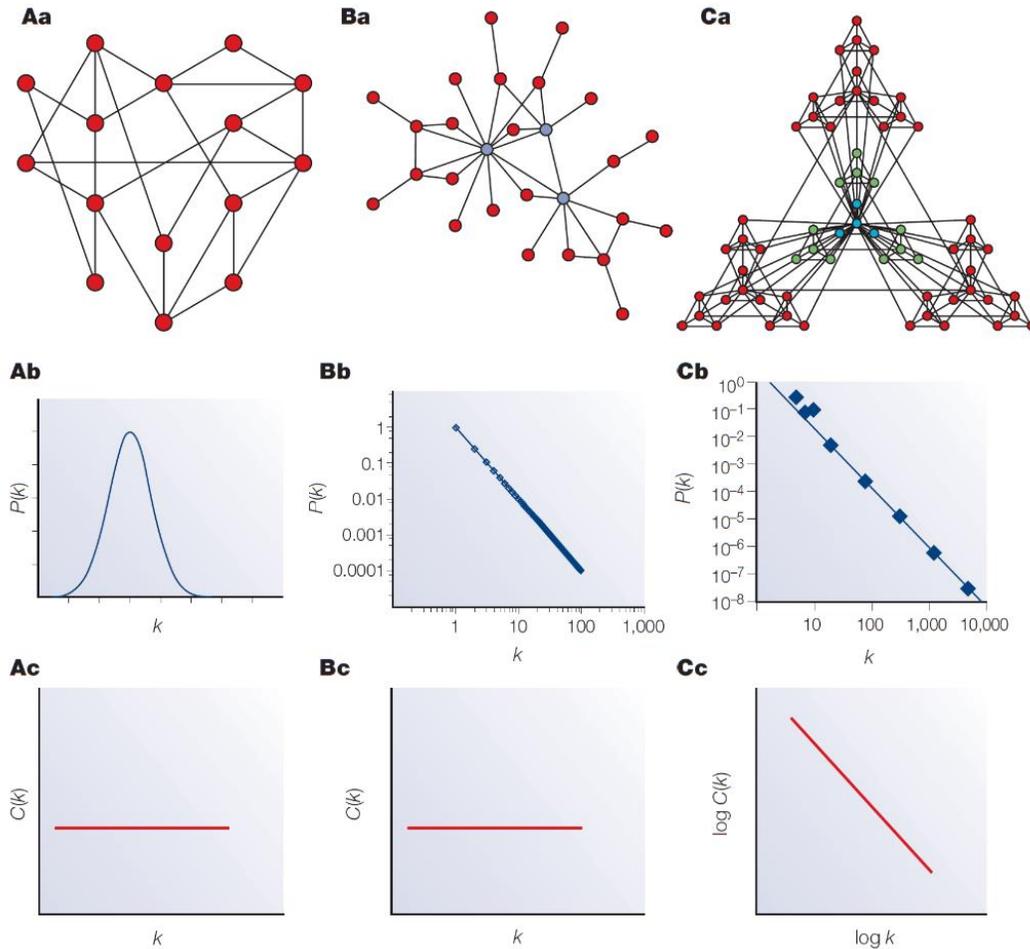


Figura 4: Modelos teóricos de redes y sus propiedades distintivas. (A) Redes aleatorias. (B) Redes libres de escala (scale-free). (C) Redes jerárquico-modulares. Sus propiedades características están representadas debajo de los modelos (b) distribución de conectividad,  $P(k)$ , y (c) de coeficiente de agrupamiento,  $C(k)$ . Imagen tomada de (Barabasi and Oltvai 2004).

## Modelo de redes aleatorias

El principal modelo de redes aleatorias es el de comunidades, como características principales, presenta modularidad mayor a cero, no presenta característica libre de escala ni una jerarquía evidente. Su distribución de

conectividades tiene una distribución poisson y pueden o no presentar coeficiente de agrupamiento.

El modelo Erdős-Rényi (Figura 4, A) (Erdos and Rényi 1960) para redes aleatorias genera redes mediante un mecanismo de aleatorización para usarlas en comparaciones como modelo nulo. En este modelo fijamos el número de nodos y la probabilidad de que pueda haber una arista entre un par de nodos. Es decir, el número de aristas no es fijo, sino que colocamos una arista entre cada par de nodos con una probabilidad independiente. El que cada arista sea incluida con una probabilidad diferente puede llevarnos a obtener redes que tengan elementos desconectados de la red principal. Por esta razón cuando se utiliza este modelo de red se busca el componente gigante, que es un componente conectado con la mayor cantidad de aristas del grafo entero.

### **Modelo de redes libres de escala**

El modelo de redes aleatorias libres de escala nos presenta estructuras tipo árbol, como su principal característica es la ausencia de modularidad, sin embargo, presenta un comportamiento libre de escala, así como una jerarquía entre sus elementos. Su distribución de conectividades se comporta como una ley de potencias y su coeficiente de agrupamiento es constante.

En el modelo Barabási-Albert (Figura 4, B) (Barabási and Albert 1999) basado en conexión preferencial se fija el número de nodos y una constante llamada constante de ramificación, la cual indica el número de conexiones que debe de tener cada uno de esos nodos. Los nodos van siendo añadidos uno por uno a la red y van siendo conectados a un grupo de nodos previamente añadidos. Las conexiones son realizadas a los nodos existentes con una probabilidad equivalente a su grado previo (conexión preferencial), debido a esto, siempre se obtienen árboles de este modelo y no hay componentes desconectados.

La conexión preferencial (*preferential attachment*), en la cual los nodos que están más conectados o poseen una antigüedad mayor tienen una mayor capacidad de atraer más conexiones mientras la red va creciendo, a este comportamiento se le llama “el rico se vuelve más rico” (“*rich-get-richer*”) y da como resultado redes con una estructura libre de escala. A diferencia de las redes aleatorias, las redes libres de escala son robustas a la remoción aleatoria de nodos. Esto es debido a la presencia de “*hubs*” los cuales se encuentran en bajas cantidades y son los encargados de mantener unida la red.

### **Modelo de redes jerárquicas**

El modelo de redes jerárquico modular, presenta como características principales modularidad mayor a cero, un comportamiento libre de escala, y una jerarquía entre sus elementos. Su distribución de conectividades presenta un comportamiento de ley de potencias al igual que su coeficiente de agrupamiento.

El modelo jerárquico de redes (Figura 4, C) integra una topología libre de escala con una estructura modular, describiendo una red con una distribución de conectividades que presenta una ley de potencias. La característica principal de las redes jerárquico-modulares es el coeficiente de agrupamiento escalado y que se ve como una pendiente a una escala logarítmica. Una arquitectura jerárquica nos dice que los nodos con más bajas conectividades se encuentran agrupados en zonas y estas zonas a su vez se encuentran unidas o comunicadas por nodos “*hub*”(Barabasi and Oltvai 2004).

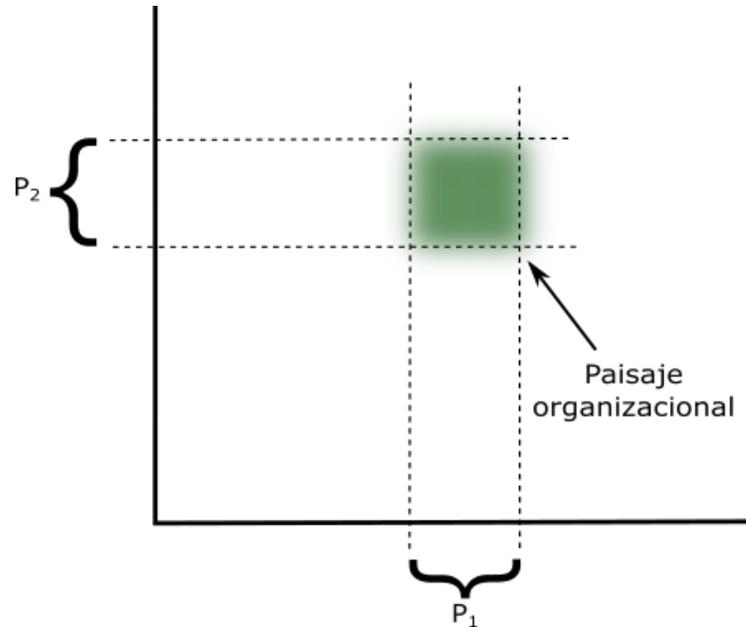
## **Redes biológicas**

Se ha explorado poco cómo la evolución moldea los sistemas y la arquitectura funcional de las redes de regulación. En bacterias se ha identificado que las redes de regulación poseen una estructura jerárquico-modular (Ma, Buer et al. 2004, Resendis-Antonio, Freyre-Gonzalez et al. 2005). Esta organización les confiere dos propiedades al parecer contradictorias, robustez y plasticidad, las cuales contribuyen en la presencia de características biológicas deseables. De hecho, estudios han demostrado que en bacterias las redes de regulación no solo son plásticas sino son altamente plásticas en su arquitectura (Price, Dehal et al. 2007). Esta plasticidad en su regulación le permite al sistema biológico adaptarse exitosamente a nuevos nichos, incrementando así su capacidad de evolucionar por medio del desarrollo de circuitos regulatorios alternativos en respuesta a diferentes ambientes. Por otra parte, la robustez permite a estas redes perder interacciones y seguir manteniendo el sistema en funcionamiento.

## **Paisaje organizacional de redes de regulación**

A pesar de que las redes de regulación poseen una alta plasticidad, se ha descrito la existencia de principios organizacionales conservados en bacterias distantes filogenéticamente (Freyre-Gonzalez, Trevino-Quintanilla et al. 2012), mediante convergencia evolutiva, e incluso su aparente universalidad (Ibarra-Arellano, Campos-Gonzalez et al. 2016). Estos principios organizacionales deben establecer un conjunto de restricciones sobre la evolución de redes regulatorias, delimitándolas a moverse dentro de un espacio restringido de alternativas al cual nosotros llamaremos “paisaje organizacional” de las redes de regulación. Esto implica que algunos parámetros o propiedades de las redes puedan ser invariantes o encontrarse acotados a un rango operativo, los cuales su vez pueden estar asociados a propiedades biológicas o genómicas. Este espacio restringido de alternativas, el paisaje

organizacional, nos muestra las conformaciones o posibles conformaciones que la evolución ha explorado y que, por lo tanto, por sus características tienen una relevancia biológica.



*Figura 5: Paisaje organizacional. En esta representación simplificada del paisaje organizacional solo tenemos dos dimensiones, es decir dos propiedades contrastando sus rangos. El espacio verde que se puede observar es el área de posibles valores para esas dos propiedades que pueden ocupar las redes.*

El paisaje organizacional puede tener  $n$  dimensiones, dependiendo de cuántas propiedades se consideren; en la Figura 5 por simplificación únicamente representamos dos. En esta representación observamos los rangos de valores de dos propiedades contrastados, el área verde sería el paisaje organizacional de los posibles valores que pueden ocupar estas dos propiedades para un tipo de red dada. De esta forma podemos caracterizar los tipos de redes y analizar las implicaciones de sus propiedades. Por ejemplo, obteniendo el paisaje organizacional conformado por las propiedades más relevantes de redes de regulación podría ser utilizado como un buen criterio de selección de redes generadas por predicciones si es que al analizar sus propiedades éstas no concuerdan con lo observado en redes biológicas de regulación.

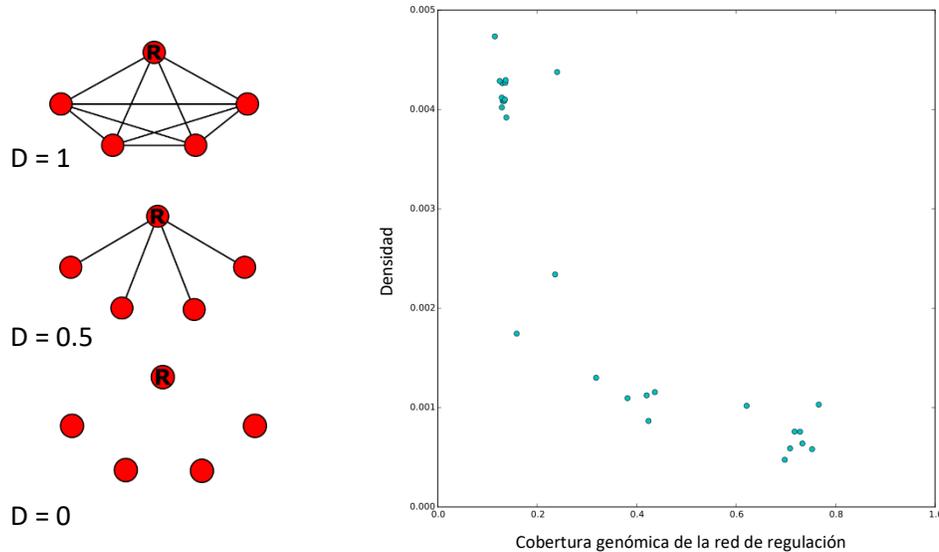


Figura 6: La propiedad de densidad en redes de regulación. (Izquierda) Representación de un grafo con diferentes densidades, de arriba hacia abajo se muestra cómo va decreciendo la densidad. (Derecha) Densidad de diferentes redes de regulación graficadas con su respectiva cobertura genómica.

Un ejemplo de propiedad que se encuentra acotada a un rango operativo es la densidad, la cual nos habla de la proporción entre nodos y aristas que existe en un grafo. En la Figura 6 (izquierda) tenemos por la parte de arriba la representación de un grafo con densidad uno, es decir, es un grafo completo con todos sus nodos conectados entre ellos; si seguimos bajando por la figura observamos que al final llegamos a una densidad de cero en la cual tenemos nodos des-conexos. En la figura 6 (derecha) observamos los valores de densidad para diferentes redes de regulación ordenados según su cobertura genómica. Por cobertura genómica entendemos que es el porcentaje o fracción del genoma incluido en nuestro modelo de red de regulación. densidad de las redes de regulación se encuentra dentro de un rango muy pequeño, entre 0.0005 y 0.005, además, pasando del 40% de cobertura genómica se muestran sus valores de densidad con tendencia a un rango aún más estrecho. Esto nos dice que al seguirse añadiendo interacciones a las redes de regulación posiblemente se tenga que mantener esa relación, de tal manera observando cómo se comporta la densidad, podemos utilizar esta propiedad para una futura estimación del número total de interacciones que puede llegar a haber en una red de regulación. Otro ejemplo sobre propiedades de las redes con un significado biológico es el diámetro. El diámetro en

una red de regulación puede decirnos de qué tamaño es la cascada de regulación más grande, es decir, con más pasos intermedios. Muchas veces el diámetro, por ejemplo, en un análisis de todas las redes de regulación con que contamos, no excede un valor de diez, lo cual quiere decir que para transmitir una señal a través de la red el peor caso que podríamos encontrar es tener que pasar por diez intermediarios.

## **Hacia el estudio de la organización y propiedades de redes de regulación**

Se ha descrito anteriormente cómo la complejidad de un sistema usualmente toma forma de jerarquía y que los sistemas jerárquicos tienen algunas propiedades en común que son independientes de su contenido específico (Simon 1962). En búsqueda del estudio de esta complejidad se han desarrollado diversos enfoques para identificar la organización jerárquico-modular de las redes de regulación (Ma, Buer et al. 2004, Resendis-Antonio, Freyre-Gonzalez et al. 2005). Sin embargo, ninguno de estos enfoques había utilizado la información o propiedades que nos proporciona el análisis de las redes mismas. Es por este motivo que, valiéndose de las propiedades de coeficiente de agrupamiento y grado de conectividad observadas en redes biológicas, nuestro grupo de trabajo desarrolló un método llamado “enfoque de descomposición natural” o “NDA” (por sus siglas en inglés), el cual deriva matemáticamente la arquitectura y componentes a nivel de sistemas (módulos) de la estructura global de una red de regulación dada (Freyre-Gonzalez, Alonso-Pavon et al. 2008) (figura 7).

El enfoque de descomposición natural es un claro ejemplo del potencial que nos brinda el estudio de las propiedades de las redes biológicas. Analizando los resultados obtenidos encontramos que no solo las distribuciones de conectividad y de coeficiente de agrupamiento siguen un comportamiento de ley de potencias, sino que también la distribución de tamaño de módulos identificados presenta el mismo comportamiento. La aparición recurrente de este comportamiento de ley de potencias

(Resendis-Antonio, Freyre-Gonzalez et al. 2005, Freyre-Gonzalez, Alonso-Pavon et al. 2008) nos sugiere que la auto-similitud es un principio que puede estar jugando un papel importante en la arquitectura de las redes de regulación.

Parte de la naturaleza auto-similar de las redes biológicas la podemos observar con la estructura anidada en que ocurren los patrones de regulación y los cuales asemejan la organización anidada propia de una *matryoshka* (figura 8). En la complejidad de las redes podemos observar cómo ésta toma forma de jerarquía y en la cual podemos ver claramente cómo los sistemas biológicos se componen de subsistemas y a su vez estos poseen sus propios subsistemas, es decir, módulos anidados en módulos.

Es precisamente por estas observaciones que nuestro proyecto, anteriormente enfocado en la exploración de varias propiedades, lo hemos dirigido principalmente al análisis de la propiedad fractal y de auto-similitud presente en las redes de regulación. Esta propiedad muestra relevancia en los mecanismos generativos de estas redes y, a diferencia de las demás propiedades propuestas para analizar, su cálculo implica un mayor esfuerzo en su implementación y es de mayor complejidad.

## **Enfoque de descomposición natural**

Aquí presentamos la serie de pasos que componen el enfoque de descomposición natural adaptado de Ibarra-Arellano, Campos-Gonzalez et al. (2016). Para aplicar el enfoque de descomposición natural se comienza con la representación una red de regulación como un grafo dirigido Figura 7 (A). En este grafo cada gen se representa con un nodo, mientras las uniones entre ellos son las interacciones regulatorias. Posteriormente, los reguladores globales son identificados por el valor kappa ( $K$ ) Figura 7 (B), el cual define un punto de equilibrio entre dos comportamientos contradictorios: importancia del nodo (alto grado de conectividad

de salida, coeficiente de agrupamiento bajo) y modularidad (baja conectividad de salida, coeficiente de agrupamiento alto). Primero la distribución del coeficiente de agrupamiento,  $C(k)$ , es computada. Después dado que  $C(k) = \gamma k^{-\alpha}$ , el valor- $K$  es luego calculado usando la formula  $K = \sqrt[\alpha+1]{\alpha\gamma} \cdot k_{max}$ .

Se identifican los módulos funcionales Figura 7 (C). Los reguladores globales (los cuales tendrán una conectividad mayor al valor de  $K$ ) y sus conexiones serán removidas de la red, de esta manera quedando revelados los módulos como islas aisladas y los genes de maquinaria basal, los cuales no codifican para reguladores, como nodos aislados. Además, se identifican los pre-submódulos que componen el mega-módulo (Figura 7, D). Aislando el mega-módulo se retiran todos los genes que no codifican para reguladores junto con sus conexiones de tal forma que podemos identificar las islas de nodos remanentes como pre-submódulos.

Identificación de los genes inter-modulares Figura 7 (E). Todos los genes que no codifican para reguladores (nodos remarcados con rojo) y sus interacciones son añadidos de nuevo siguiendo una regla: Si un gen es regulado por genes que pertenecen a dos o más sub-módulos, entonces este gen es clasificado como inter-modular (nodos amarillos) si no, este gen es añadido al mismo sub-modulo que sus reguladores.

Finalmente, los reguladores globales (nodos rojos) son añadidos de nuevo junto con sus conexiones para reconstruir la red original Figura 7 (F) y a la vez revelar su arquitectura, la cual consta de tres capas que, debido a sus proporciones, dan una forma de diamante y se encuentra compuesta a nivel de sistemas por las cuatro clases identificadas previamente: reguladores globales, módulos, maquinaria basal y genes inter-modulares.

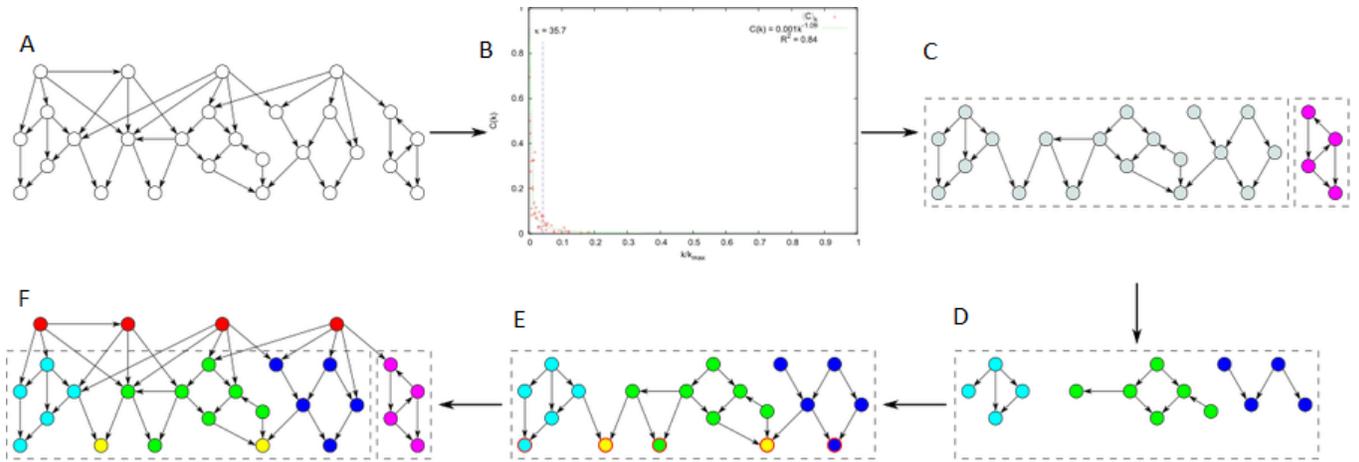


Figura 7: Enfoque de descomposición natural. Descripción de la serie de pasos que conforman la descomposición natural de redes de regulación y con la cual se obtiene la clasificación de los nodos de la red. Imagen adaptada de (Ibarra-Arellano, Campos-Gonzalez et al. 2016).

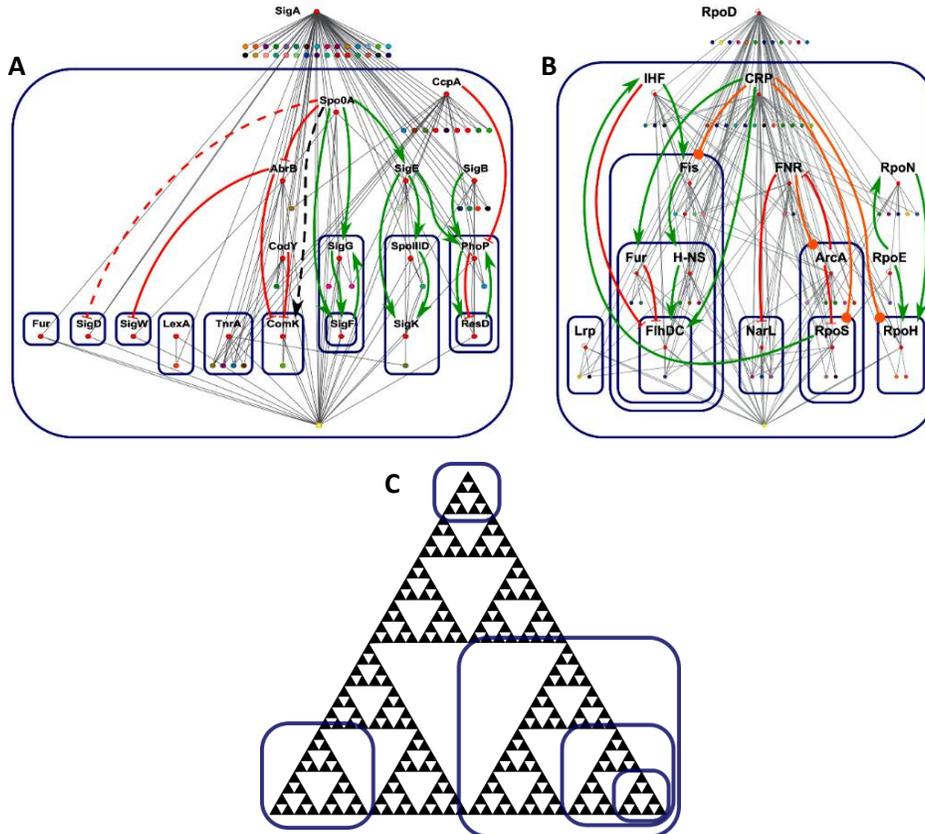


Figura 8: Redes de regulación y fractales, estructuras anidadas. (A) Red de regulación de *Bacillus subtilis*. (B) Red de regulación de *Escherichia coli*. En ambas redes de regulación los módulos fueron reducidos a un solo nodo, las flechas de colores representan las interacciones regulatorias de los reguladores globales mostrando el efecto de su interacción. Imagen tomada de (Freyre-Gonzalez, Trevino-Quintanilla et al. 2012). (C) Triangulo de Sierpinski. En las tres figuras se observan estructuras anidadas lo cual al comparar las interacciones anidadas con una figura fractal, surge la noción de la auto-similitud en redes de regulación.

## Auto-similitud

La auto-similitud o invariancia de escala es una característica que define a los fractales y se ha identificado como una propiedad general de muchos sistemas, no sólo biológicos. Tal vez uno de los mejores ejemplos para explicar el concepto de auto-similitud en sistemas biológicos es el de la fronda de un helecho, en el sentido en que cada uno de sus segmentos es similar, mas no igual a la estructura a la fronda. En el contexto de las redes biológicas moleculares, esta propiedad de auto-similitud ha sido explorada limitadamente en un par de redes biológicas (Song, Havlin et al. 2005), mas no en redes regulatorias. A pesar de ser una característica biológica relevante, se ha llegado a mencionar que las redes biológicas carecen de una estructura libre de escala, esto a su vez negando la presencia de fractalidad en redes biológicas (Lima-Mendez and van Helden 2009).

La idea de auto-similitud en redes complejas viene de la re-normalización sucesiva de estas de manera que la red termina viéndose igual bajo diferentes longitudes de escalas como en la Figura 8 (A y B) donde podemos observar la re-normalización de las redes de regulación de *E. coli* y de *B. subtilis* manteniendo el mismo patrón de estructuras regulatorias anidadas. También otra definición de la auto-similitud sería que una red es auto-similar si su distribución de grado es invariante bajo re-normalizaciones. La auto-similitud está estrechamente relacionada con la aparición de estructuras llamadas fractales.

## ¿Qué es un fractal?

El término fractal fue propuesto por el matemático polaco Benoit Mandelbrot en 1975 y significa quebrado o fracturado (Mandelbrot 1975). Un fractal podemos imaginarlo como una figura geométrica diferente a las convencionales, debido a que en ella encontramos patrones que se repiten una y otra vez a cada escala que

observemos, es por eso que se dice que se repiten en patrones auto-similares. Las propiedades fractales incluyen independencia de escala, auto-similitud y complejidad topológica.

Estas propiedades definen a un fractal a pesar de parecer subjetivas, por lo que al decidir si un objeto es fractal o no, podemos recurrir a métodos tomados de la teoría fractal que describen la irregularidad de los objetos naturales. En el análisis de fractales, se utiliza la dimensión fractal, o dimensión fraccional la cual es muchas veces utilizada como medida de complejidad (Kenkel and Walker 1996, Kim, Goh et al. 2007).

## **Dimensión fractal**

La dimensión es lo que distingue a los fractales de las figuras geométricas clásicas, y para explicar mejor el concepto de dimensión fractal es necesario primero revisar lo que entendemos por dimensión. Generalmente usamos el concepto de dimensión en dos sentidos, para referirnos, por ejemplo, a las dimensiones de un espacio euclidiano o para referirnos al número de variables que posee un sistema dinámico. Como ejemplo sencillo, en matemáticas se considera que una figura que posee dimensión de uno, cuando es seccionada, cada sección va a ser parecida a la figura original, una línea. Si tenemos una figura de dos dimensiones al seccionarla seguiremos manteniendo la representación del objeto original. Sin embargo, para los fractales las dimensiones no se comportan de esta manera. Dado que las figuras fractales presentan irregularidades, al seccionarlas nos encontramos con que presentan dimensiones intermedias, es decir fraccionarias, se encuentran entre dimensiones enteras debido a su complejidad.

Al seccionar un fractal en piezas idénticas al fractal original y escalarlas por algún factor podemos llegar a recrear el objeto original. Es decir que si fraccionamos

cada una de las figuras por un factor de  $\frac{1}{k}$  necesitaremos escalarlas por para cada dimensión a  $k^d$  copias para recrear el objeto original, donde  $k$  es el factor y  $d$  es la dimensión original del objeto. De esta manera se establece una relación tal que, si tenemos un objeto auto-similar y lo descomponemos en  $n$  piezas, cada una será una copia del objeto original escalada por un factor de  $\frac{1}{k}$  por lo tanto:

$$n = k^d$$

$$\log(n) = \log(k^d)$$

$$\log(n) = d \log(k)$$

$$d = \frac{\log(n)}{\log(k)}$$

En geometría fractal, el método de *box-counting*, que se define más adelante como *box-covering* cuando nos referimos a redes complejas, es la principal manera de evaluar la dimensión fractal de un objeto fractal. Debido a que la técnica para identificar la presencia de fractalidad en redes complejas es análoga a la de los fractales. La dificultad en este caso radica en los fractales convencionales se encuentran embebidos en un espacio Euclidiano y para las redes las métricas euclidianas resultan irrelevantes. Para resolver esta dificultad, Song, Havlin et al. (2005) propusieron el método de *box-covering* el cual toma como métrica de distancia en redes el camino mínimo (geodésico) entre dos nodos. Cabe mencionar que la dimensión fractal generalmente es usada para referirse a la auto-similitud geométrica, en este caso de redes complejas nos estamos refiriendo a una auto-similitud topológica.

Aunque se define más adelante el método de *box-covering* podemos decir que la fractalidad de una red o fractalidad topológica se refiere a la relación de ley de potencias entre el número mínimo de cajas necesarias para cubrir una red entera y el tamaño de cajas. Por lo tanto, un grafo será fractal si existe una dimensión fractal  $d_B$  finita obtenida de esta relación (Molontay 2015).

## Método de *box-covering*

La siguiente definición del método de box covering se describe en Song, Gallos et al. (2007) y Molontay (2015) de esta manera: Dada una cierta red  $G$  y un tamaño  $l_B \in \mathbb{N}, 2 \leq l_B \leq D(G)$  con el cual fijamos el tamaño de caja que particiona los nodos de la red. Una caja es un conjunto de nodos donde todas las distancias  $d(p, q)$  entre cualesquiera dos nodos  $p$  y  $q$  en la caja son menores a  $l_B$ . El número mínimo de cajas requeridas para cubrir la red completa  $G$  es designado como  $N_B(l_B)$ . En el caso en que  $l_B = 1$ ,  $N_B(l_B)$  será obviamente igual al tamaño de la red  $|N|$ , mientras que si  $l_B > D(G)$ , por ende  $N_B = 1$ . De esta manera el objetivo final de todos los algoritmos de *box-covering* es localizar la solución óptima es decir identificar el  $N_B(l_B)$  mínimo para cada tamaño  $l_B$  dado.

Ya que las cajas que se generan por el *box-covering* contienen nodos separados por una distancia  $l_B$ , podemos decir que esta es la fracción por la cual estamos partiendo nuestra red. La dimensión fractal ( $d_B$ ) se define como:

$$N_B(l_B) \approx l_B^{-d_B}$$

Es en esta relación que se observa que el número de cajas que se requiere para cubrir la red tiene la forma de una ley de potencias con respecto al tamaño de caja.

Una vez realizado el *box-covering* y para obtener el valor de la dimensión fractal se grafica en escala logaritmica  $N_B(l_B)$  normalizado por el tamaño de  $G$ , contra  $l_B$ . Si esta relación es lineal, entonces sabremos que la red tiene una dimensión fractal finita (Molontay 2015).

## Resultados y discusión

### Obteniendo las dimensiones fractales de redes de regulación

Realizamos este proyecto con ayuda del lenguaje de programación Python versión 2.7, para el cual se emplearon librerías para graficar (Matplotlib v1.4.3), análisis de redes (Networkx v1.8.1), cómputo científico (NumPy v1.11.0) y cálculos estadísticos (SciPy v0.14.0).

#### Implementación del método de *box-covering*

Realizamos un análisis de los diferentes métodos para obtener dimensiones fractales de redes complejas reportados en (Song, Gallos et al. 2007). En ellos se exponen las ventajas y desventajas de algunos métodos usados para calcular la dimensión fractal. Sin embargo, a pesar de que este artículo menciona que estos métodos son óptimos, al evaluarlos, resultan muy costosos en recursos computacionales ya que requieren de algoritmos para encontrar caminos mínimos. La búsqueda de caminos mínimos es una operación que además de utilizar mucha memoria, es tardada y dependiente del tamaño de las redes ya que implica recalcular un mismo camino más de una vez haciendo al algoritmo poco óptimo. Estos métodos por sus características son ejecutados en súper computadoras y grandes servidores. Pero dada la redundancia de operaciones, es posible ejecutarlo en equipos con menos recursos. En consecuencia, nosotros propusimos e implementamos un nuevo algoritmo para calcular la dimensión fractal que además de ser efectivo, es eficiente en el uso de recursos computacionales.

## Optimización del algoritmo de *box-covering*

El principal problema del algoritmo de *box-covering* es encontrar el camino mínimo entre dos nodos, lo cual, de repetirse para cada nodo, termina volviéndolo un problema cuya resolución toma una cantidad de tiempo que incrementa exponencialmente con el tamaño de la red. Para abordar y optimizar este problema empleamos programación dinámica. La programación dinámica u optimización dinámica es un método empleado para resolver problemas complejos descomponiéndolos en una serie de sub-problemas, de esta manera cada sub-problema se soluciona por separado y en lugar de volver a realizar cálculos, utiliza soluciones anteriores, las cuales han sido almacenadas previamente en memoria.

Proponemos realizar el pre-cálculo de matrices de adyacencia de las redes de regulación, que presentaran cada distancia posible entre nodos añadiendo aristas nuevas que representen las distancias o caminos. Por ejemplo, en un grafo de  $n$  nodos, si buscamos caminos mínimos entre todos los nodos, habría que realizar la operación  $n-1$  veces, en cambio partiendo de la matriz de adyacencia en la cual representamos nodos a una sola arista de distancia de cada uno y si queremos saber ahora cuáles nodos están a dos aristas de distancia lo único que deberíamos realizar es elevar al cuadrado la matriz y tendríamos la representación de todos los nodos que están a dos aristas de distancia, si elevamos al cubo, tendríamos distancias de tres y así en adelante, si vamos superponiendo las matrices, iríamos incluyendo las distancias anteriores, dando por resultado que cada matriz represente la adyacencia entre nodos a una distancia menor o igual a la representada en cada matriz.

Esta matriz representando la distancia máxima de  $n$  aristas o pasos entre nodos incluirá además de nodos a esta distancia, todas las distancias menores y se le llamará matriz de distancia  $i$ , donde la distancia es  $< i$  y se representa como  $l_i$ . Mostramos a continuación la representación matemática de la generación de matrices de adyacencia para cada  $l_i$ .

$$\text{Matriz de adyacencia } \mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

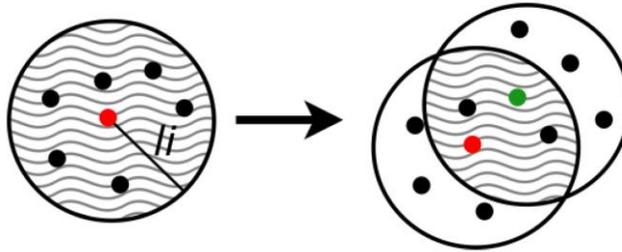
$$\mathbf{T}(\mathbf{M}) = \begin{cases} 1 & \text{si } m_{i,j} \neq 0 \\ 0 & \text{de otra manera} \end{cases}$$

$$\begin{aligned} \mathbf{M}_{l_2} &= \mathbf{A} \\ \mathbf{M}_{l_3} &= \mathbf{M}_{l_2} \vee \mathbf{T}(\mathbf{A}^2) \\ \mathbf{M}_{l_4} &= \mathbf{M}_{l_3} \vee \mathbf{T}(\mathbf{A}^3) \\ &\dots \\ \mathbf{M}_{l_n} &= \mathbf{M}_{l_{n-1}} \vee \mathbf{T}(\mathbf{A}^{n-1}) \end{aligned}$$

Enseguida, sobre estas matrices aplicamos el algoritmo de *box-covering* para generar cajas de tamaño  $l_B$  y posteriormente contar el número de cajas  $N_B$  que son necesarias para cubrir la red. Por tamaño topológico  $l_B$  tiene que cumplirse la condición de que todos los nodos que pertenezcan a una caja deben de tener entre ellos una distancia topológica menor a la establecida por el tamaño de caja  $l_B$ , es decir, su vecindario- $l$ .

### Aquí presentamos el algoritmo de *box-covering*:

1. Construimos un conjunto  $\mathbf{S}$  conteniendo todos los nodos en la red.
2. De  $\mathbf{S}$  escogemos un nodo al azar  $p$  como semilla para una nueva caja.
3. Computamos  $\mathbf{N}_p$  el vecindario de  $p$  y  $\mathbf{B1} = \mathbf{S} \cap \mathbf{N}_p$ .
4. De  $\mathbf{B1}$  escogemos un nodo al azar  $q$ . Computamos  $\mathbf{N}_q$  y  $\mathbf{B2} = \mathbf{S} \cap \mathbf{N}_q$ .
5.  $\mathbf{NB} = \mathbf{B1} \cap \mathbf{B2} \cup \{p, q\}$
6. Removemos de  $\mathbf{S}$  todos los elementos que se encuentren en  $\mathbf{NB}$ .
7. Repetimos los pasos 2 a 6 hasta que  $\mathbf{S}$  se encuentre vacío.



Formamos una nueva caja eligiendo de la red un nodo al azar al que llamamos semilla (rojo). Después consideramos a sus nodos vecinos (vecindario) dentro de una distancia  $l_i$ . En seguida tomamos otra semilla al azar (verde) del vecindario de la primera semilla. La intersección entre los vecindarios de las semillas escogidas será la caja final.

Como nuestras matrices de adyacencia representan el vecindario de cada nodo a una distancia menor a  $l_B$ , nuestro algoritmo de *box-covering* solo recorre la matriz formando cajas y ya no calculando cada camino.

## Dimensión fractal en modelos de redes biológicas

Aquí analizamos una variedad de modelos de redes de regulación bacterianos a los cuales se les aplicó el método de *box-covering* que desarrollamos para obtener su dimensión fractal.

### Obtención de redes de regulación bacterianas

Las redes fueron obtenidas del atlas de sistemas, elementos a nivel de sistemas y propiedades globales de redes bacterianas **Abasy Atlas** (<http://abasy.ccg.unam.mx/>) el cual, cuando se realizó este estudio, contaba con 50 redes de regulación, abarcando 9 organismos (*Mycobacterium tuberculosis*, *Bacillus subtilis*, *Corynebacterium*

*glutamicum, Escherichia coli, Staphylococcus aureus, Pseudomonas aeruginosa, Streptococcus pyogenes, Streptococcus pneumoniae, Streptomyces coelicolor*). Una de las ventajas que nos proporciona Abasy Atlas, aparte de abarcar un amplio rango de bacterias con los modelos de redes más completas hasta la fecha, es que todas las interacciones regulatorias se encuentran sustentadas por evidencia experimental (Ibarra-Arellano, Campos-Gonzalez *et al.* 2016). Las interacciones, sus anotaciones y la fuerza de la evidencia sustentando cada interacción están disponibles para su descarga en el formato de intercambio de información JSON (JavaScript Object Notation).

Descargamos los archivos de texto disponibles con información de interacciones regulatorias para todos los organismos de Abasy y posteriormente modelamos las redes de regulación usando la librería para manipulación y estudio de redes complejas NetworkX.

### **Cálculo de dimensión fractal**

Para obtener la dimensión fractal de las redes que modelamos, empleamos el algoritmo de *box-covering* que describimos anteriormente. Para ello generamos las matrices de adyacencia necesarias para aplicar el algoritmo de *box-covering* a cada una de las 50 redes. Debido a que el proceso de cubrir la red con cajas de diferentes tamaños es un método no determinístico, realizamos 10 iteraciones por tamaño de caja ( $l_B$ ) y almacenamos en archivos de texto los diferentes valores de número de cajas necesarias para cubrir la red ( $N_B$ ). De los valores de  $N_B$  seleccionamos los menores, los normalizamos por el tamaño de la red y los graficamos contra  $l_B$  en una escala logarítmica. Finalmente, obtuvimos la dimensión fractal por la pendiente que se obtiene de la regresión lineal robusta que se generó y que a su vez evidencia la relación que hay entre  $N_B$  y  $l_B$  que tiende a una ley de potencias.

## Análisis de la dimensión fractal a través del tiempo

Además de trabajar sobre las 50 redes de regulación obtenidas de Abasy Atlas, contamos con diferentes estadios de curación, conforme se han ido descubriendo nuevas interacciones, de la red de regulación de *E. coli* k12 MG1655. Este conjunto de datos que denominamos "a través del tiempo" abarca un intervalo de 10 años a partir del 2003. En consecuencia, aplicamos el método de *box-covering* a estas redes intermedias para observar el comportamiento de su dimensión fractal conforme se descubren nuevas interacciones.

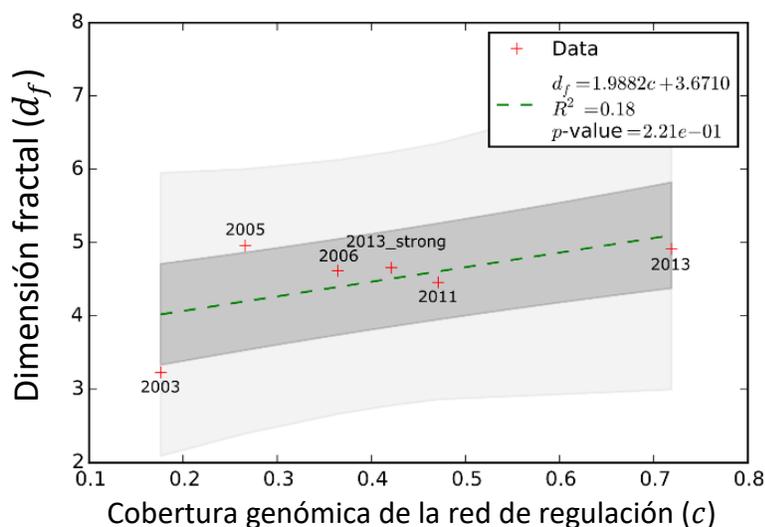


Figura 9 Cambio de la dimensión fractal de la red de *E. coli* curada a través del tiempo

Al graficar la dimensión fractal de *E. coli* a través del tiempo Figura 9, vemos cómo ésta oscila dentro de un rango restringido. Este resultado es interesante debido a que estamos viendo el comportamiento de la dimensión fractal casi sin variación desde que la red tiene menos de un 20% hasta un 70% de cobertura genómica. Aunque la red que se tenía en el 2003 tenga una dimensión fractal que se ubica por debajo del rango en el que se mueven el resto de las redes, ello no significa un mal resultado. Esto puede deberse a lo incompleto de la red en cuanto a interacciones y además a que posiblemente son muy pocos nodos los que se tienen (baja cobertura genómica) y es muy fácil a cualquier tamaño de caja cubrir con pocas cajas toda la red, como podría

ser el caso de una red aleatoria, en la cual la mayoría de los nodos se encuentran conectados entre ellos.

Este resultado nos indica que a partir del 20% de cobertura genómica, la red obtuvo una característica esencial. Posiblemente estuvo creciendo sobre una misma estructura a manera de andamiaje a la cual solo se le han ido añadiendo nodos e interacciones. Si revisamos la manera en que se va curando una red, siempre es por grupos, es decir, las redes de regulación se van construyendo a base de unidades de regulación. Los dos niveles básicos de organización en la regulación son el operón y el regulón, y los módulos están compuestos de una o más de estas unidades. De esta manera, si un módulo ya ha sido establecido, la estructura general se mantiene, aunque se le sigan añadiendo elementos. Al estar creando cajas sobre esta misma estructura es más probable que se siga agrupando a los mismos módulos, indicándonos la presencia de esta estructura sobre la cual se puede desarrollar una red.

## **Dimensión fractal en modelos de redes teóricas**

Para contrastar los resultados obtenidos de las redes biológicas y determinar su significancia, decidimos contrastarlos contra modelos teóricos. Para ello, generamos redes aleatorias usando el modelo Erdős-Rényi, así como redes libres de escala usando el modelo Barabási-Albert ambos descritos anteriormente. Estos modelos de redes poseen parámetros ajustables, los cuales determinamos mediante un análisis de propiedades de las redes biológicas con la finalidad de igualar la proporción entre nodos y aristas. Decidimos hacer comparables las redes igualando su proporción entre nodos y aristas debido a que el cálculo de la dimensión fractal es dependiente de esta relación. De esta manera comparamos los valores de dimensión fractal de redes teóricas con los obtenidos de redes biológicas, minimizando diferencias y logrando un

equivalente comparable para cada una de las redes biológicas analizadas anteriormente.

### **Generación de redes aleatorias**

El modelo Erdős-Rényi para generar redes aleatorias recibe como parámetros el número de nodos de la red resultante, equivalente a las redes biológicas y una probabilidad de crear una arista entre cada par de nodos. Esta probabilidad es equivalente a la densidad de cada red biológica ya que la densidad representa la relación entre nodos y aristas.

### **Generación de redes libres de escala**

El modelo de Barabási-Albert para generar redes libres de escala recibe como parámetros el número de nodos de la red resultante, equivalente a las redes biológicas y un factor de ramificación. El factor de ramificación se fijó en 2, debido a que el promedio de número de reguladores por gen es de aproximadamente 2.

### **Cálculo de dimensión fractal de redes teóricas**

Una vez generadas las redes teóricas con densidad equivalente a la presentada por las redes biológicas, aplicamos nuestro método de *box-covering* para analizar sus dimensiones fractales.

## **Análisis de las distribuciones de dimensión fractal**

A continuación, presentamos las distribuciones de dimensiones fractales de las redes regulatorias biológicas contra los modelos aleatorios. Graficamos las

distribuciones para observar su comportamiento y posteriormente analizamos su significancia estadística contra los modelos teóricos, Erdős-Rényi y Barabási-Albert. Comparamos tanto las distribuciones como las medias empleando las pruebas estadísticas de Kolmogórov-Smirnov y Mann-Whitney U, respectivamente. Ambos pares de análisis resultaron significativos a un nivel de significancia de  $10^{-7}$ .

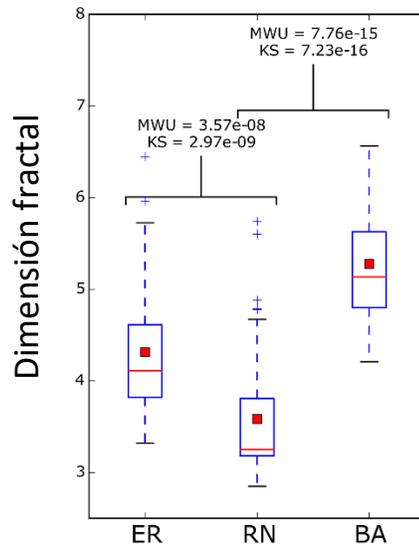


Figura 10 Graficas de caja y bigote que muestran la distribucion de dimension fractal para las diferentes redes analizadas. Erdos-Renyi (ER), Redes Biologicas (RN), y Barabasi-Albert (BA), prueba U de Mann-Whitney (MWU) y Kolmogorov-Smirnov (KS)

Debido al nivel de significancia podemos decir que las dimensiones fractales obtenidas por las redes teóricas y las redes biológicas son diferentes Figura 10. Esto puede deberse a los mecanismos de generación de cada red. Ni siquiera el modelo libre de escala Barabási-Albert tuvo dimensiones fractales cercanas a las redes biológicas, algo que se podría esperar al tener ambos la propiedad de libertad de escala, por lo que muchas veces se emplea como modelo representativo de redes biológicas por presentar la característica de leyes de potencia.

Un punto importante que debemos recalcar aquí es que, aunque se pueda obtener la dimensión fractal de una red compleja, esto no significa que ésta sea o posea características auto-similares propias de un fractal. Por lo tanto, estaríamos

descartando el modelo Erdős-Rényi de redes aleatorias ya que no presenta la propiedad de libertad de escala.

Parte de lo que estamos viendo en las distribuciones de dimensión fractal, es que las redes generadas por el modelo Barabási-Albert tienen dimensiones fractales mayores debido a que las reglas de su proceso generativo en realidad dan lugar a árboles. Por lo tanto, al aplicar el método de *box-covering*, creando una caja que trata de abarcar dentro de sí la mayor cantidad de nodos, nos topamos con el problema de que en la estructura de tipo árbol tenemos un bajo coeficiente de agrupamiento, lo cual nos lleva a tener que recurrir a un mayor número de cajas para cubrir el grafo debido a que estamos limitados a sólo poder incluir un pequeño número de nodos por caja para cualquier tamaño de caja, lo cual aumenta la pendiente en la relación de escalas, proporcionándonos una mayor dimensión fractal.

Por el lado de las redes aleatorias Erdős-Rényi podríamos esperar que, debido a que presentan un grado de nodo característico y tienen una distribución binomial de grados de conectividad, requeriríamos un menor número de cajas para cubrir la red. En este tipo de red podemos tener un mayor coeficiente de agrupamiento lo cual garantizaría que abarquemos un mayor número de nodos por caja, y que al cubrir la red cada vez con menos cajas, sea esto lo que este aumentando la dimensión fractal. Sin embargo, tampoco nos garantiza que aleatoriamente no se haya generado una estructura parecida o que se asemeje a un árbol como en el modelo Barabási-Albert, lo cual podría explicar también el por qué presenta una dimensión fractal más alta que la observada en las redes biológicas.

En la distribución de redes biológicas observamos que la dimensión fractal se mueve principalmente entre tres y cinco, dentro de este rango podríamos ubicar a las redes jerárquico-modulares y es precisamente por esta característica de presentar módulos que alcanzamos una dimensión fractal más baja que en el caso de las redes Barabási-Albert libres de escala. La presencia de módulos en estas redes permite que cuando una caja se va a crear dentro o cerca de uno, ésta pueda incluir al módulo

completo. Como el método de *box-covering* va aplicando diferentes tamaños de caja, incluso podría ser que haya un tamaño de caja óptimo el cual logre captar la misma cantidad de módulos obtenidos por el enfoque de descomposición natural. Esto podría darnos más herramientas para definir las características topológicas que representan a un módulo y cómo estas se relacionan con un mecanismo de generación de redes guiado por una estructura fractal.

## **Paisaje organizacional**

Decidimos realizar e implementar el cálculo de 29 propiedades, dentro de estas incluida la dimensión fractal. Debido a que la visualización del paisaje organizacional es complicada por el número de dimensiones que implica, además de que los valores en que se mueven las propiedades distan en magnitud y unidades, normalizamos los valores obtenidos del análisis de las redes biológicas con un escalamiento de rango, de esta manera todos los valores quedaron dentro un rango de 0 a 1. Posteriormente realizamos un análisis de *clustering* jerárquico aglomerativo con métrica euclidiana empleando el método Ward de varianza mínima para agrupar perfiles de valores de propiedades de los organismos analizados; mientras que los organismos se ordenaron de menor a mayor cobertura genómica.

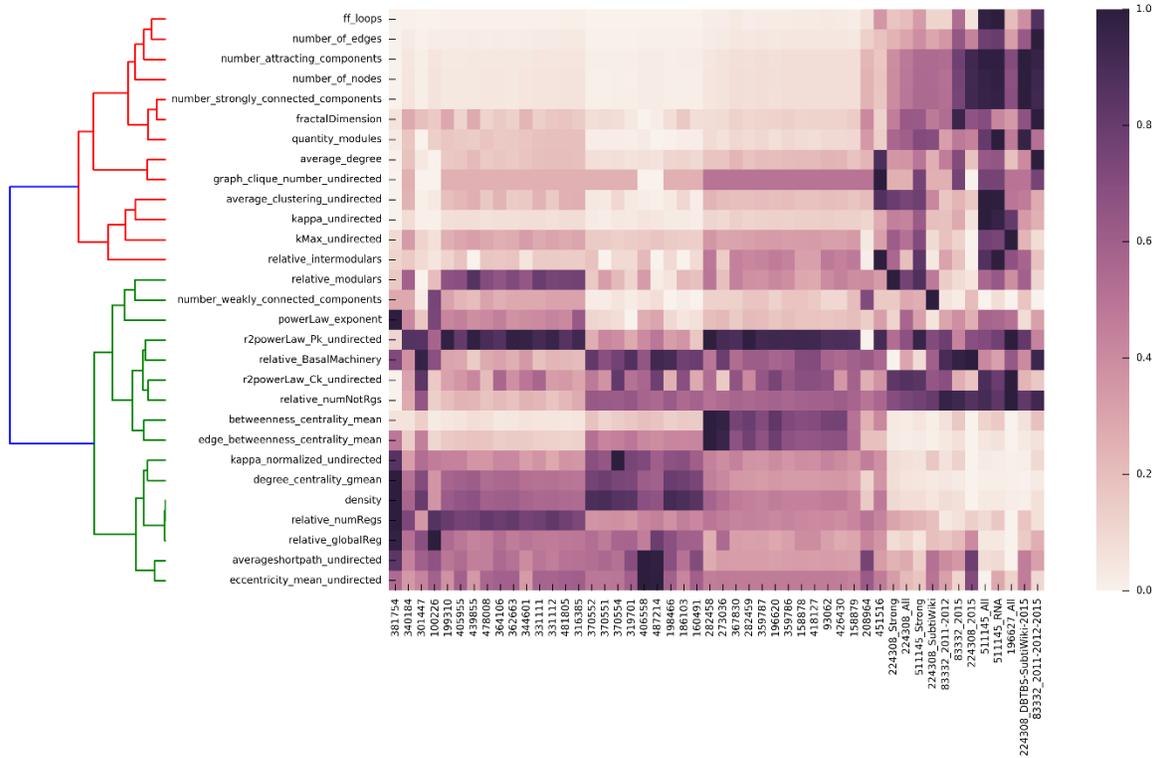


Figura 11: Heatmap y dendograma con clustering de propiedades. En el eje y tenemos las propiedades calculadas para todas las redes de regulacion. En el eje x están los identificadores de las redes de regulacion por su TaxID de NCBI ordenados por la cobertura genómica de las redes.

En el dendograma Figura 11 podemos observar cómo se forman dos grupos diferentes de propiedades relacionadas en su comportamiento con respecto a la cobertura genómica de las redes. En el primer grupo (rojo) se encuentra incluida la dimensión fractal, lo que nos llama la atención de esta observación es que la dimensión fractal está relacionada con el número de módulos y a la vez con propiedades dependientes del número de nodos y aristas. Esto implica que la dimensión fractal está relacionada con la modularidad en redes de regulación. Por otro lado, el segundo grupo (verde) tiene la propiedad de densidad, así como propiedades que surgen de las distribuciones de agrupamiento, conectividad, y propiedades que caracterizan la distribución de aristas y la navegabilidad de las redes.

## Conclusiones y perspectivas

En principio buscábamos analizar las propiedades que pueden conformar el paisaje organizacional y cómo estas podrían imponer restricciones a la evolución de redes de regulación. Al observar que la fractalidad en redes de regulación no había sido previamente analizada, abordamos el problema de la fractalidad en redes e implementamos un algoritmo para analizarla. A continuación, resumiremos las conclusiones y perspectivas de este análisis.

### Conclusiones

La presencia de fractalidad en redes de regulación nos lleva a pensar en la presencia de un andamiaje sobre el cual se desarrolla la red. Este andamiaje, también llamado esqueleto, ha sido identificado y descrito para diversas redes complejas, pero no se ha hablado sobre su presencia en redes de regulación (Goh, Salvi et al. 2006, Kim, Goh et al. 2007). La razón por la cual pudimos encontrar una dimensión fractal finita es precisamente por el esqueleto de la red, este “andamiaje” actúa como vía de comunicación de la red y es sobre él que circula todo el flujo regulatorio, es decir las señales de reguladores globales a los distintos módulos que dirigen. De esta manera la conexión entre módulos se encuentra dada por el esqueleto y las demás conexiones de la red se encuentran dentro de los módulos. Una analogía un poco burda sería imaginar el esqueleto como las ramas de un racimo de uvas y las uvas como sus módulos. Es sobre esta estructura que se podría desarrollar la fractalidad.

Con respecto a las implicaciones biológicas que conlleva la presencia de fractalidad en redes de regulación. Observamos en el dendograma de la Figura 11 que la fractalidad está altamente relacionada con la modularidad de una red. Como ya mencionamos anteriormente, una de las características de las redes de regulación es la robustez, esta robustez topológica dada por su estructura jerárquico modular puede

estar favoreciendo la aparición de fractalidad. La fractalidad puede hacer aún más robustas a las redes frente a la pérdida de “hubs”. Al estar presente una estructura a manera de andamiaje como la del esqueleto, aunado a la auto-similitud en la estructura regulatoria de la red. Podría estarse favoreciendo la dispersión de los “hubs” y módulos por la red de esta manera volviéndola menos propensa a la pérdida de nodos e interacciones. De acuerdo con Song, Havlin et al. (2006) para que una red con módulos sea robusta necesita tener una topología fractal, por lo tanto, hay una tendencia evolutiva hacia su presencia en las redes biológicas

El método optimizado de *box-covering* desarrollado en este trabajo permitió analizar las dimensiones fractales de una variedad de redes de regulación bacterianas, así como su comparación con diferentes modelos de redes teóricas. Este análisis reveló que la dimensión fractal está relacionada con el mecanismo de generación de cada modelo de red e incluso que podrían inferirse cuestiones particulares de su estructura a partir de ella.

Las conclusiones de este trabajo pueden resumirse en los siguientes puntos:

- Contrario a la generalización de que la propiedad de libertad de escala no aplica para redes biológicas y por lo tanto una red no podría presentar características fractales (auto-similitud) (Lima-Mendez and van Helden 2009) nuestros resultados y los reportados por (Freyre-Gonzalez, Trevino-Quintanilla et al. 2012, Ibarra-Arellano, Campos-Gonzalez et al. 2016), confirman que las redes de regulación bacterianas sí presentan la propiedad de libertad de escala, así como también una naturaleza fractal en su estructura.
- La fractalidad, analizada por medio de la dimensión fractal es una propiedad organizacional de redes de regulación bacteriana que no presenta una gran variación a través de distintos estadios de su curación a través del tiempo.
- La dimensión fractal es un indicador de las diferencias en mecanismos o reglas que cumplen las redes durante su proceso generativo y evolución.

- Como se muestra en el dendograma generado por la correlación entre los perfiles de propiedades de todos los organismos analizados vimos que la dimensión fractal está relacionada con la modularidad es decir, la cantidad de módulos de una red.
- La implementación de un algoritmo óptimo para *box-covering* empleando programación dinámica nos permitió analizar las dimensiones fractales de redes de regulación con un consumo mínimo de recursos computacionales (El trabajo que publica este algoritmo se encuentra en escritura).
- Hay evidencia de la existencia de un esqueleto que dota de sus propiedades tanto libres de escala como fractales a las redes de regulación.
- La alta robustez de las redes de regulación puede verse favorecida por la presencia de características fractales que fueron seleccionadas durante su evolución.

## Perspectivas

Los resultados obtenidos en este análisis abren nuevas preguntas y posibilidades para continuar con el estudio de las redes de regulación, sus características y mecanismos de evolución. Estas son perspectivas que podemos abordar en trabajos posteriores:

- Podemos utilizar las propiedades fractales de las redes de regulación para incrementar la granularidad con la que identifica sistemas el método de descomposición natural.
- Explorar las demás propiedades organizacionales para detectar invariantes que puedan mejorar los modelos de redes, así como de sus mecanismos evolutivos actuales.
- Se podría realizar un *clustering* basado en la correlación de propiedades ente organismos, esto podría darnos idea de si organismos con propiedades similares comparten un mismo nicho.
- Se podrían realizar análisis más detallados comparando propiedades obtenidas de diferentes modelos de redes para identificar y definir las características que diferencian los diferentes tipos de redes biológicas, es decir, caracterizarlas por medio de su paisaje organizacional
- Proponer y desarrollar un modelo mecanístico o generativo basado en propiedades, las cuales delimitan las posibles conformaciones que puede tener una red de regulación.

## Referencias:

- Barabási, A.-L. and R. Albert (1999). "Emergence of scaling in random networks." science **286**(5439): 509-512.
- Barabasi, A. L. and Z. N. Oltvai (2004). "Network biology: understanding the cell's functional organization." Nat Rev Genet **5**(2): 101-113.
- Bornholdt, S. (2005). "Less Is More in Modeling Large Genetic Networks." Science **310**(5747): 449-451.
- Erdos, P. and A. Rényi (1960). "On the evolution of random graphs." Publ. Math. Inst. Hung. Acad. Sci **5**(1): 17-60.
- Euler, L. (1953). "Leonhard Euler and the Königsberg bridges." Scientific American **189**(1): 66-70.
- Freyre-Gonzalez, J. A. (2008). Diseción de la arquitectura funcional de la red de regulación transcripcional de Escherichia coli: Un enfoque de descomposicion natural. Doctor en Ciencias, Universidad Nacional Autónoma de México.
- Freyre-Gonzalez, J. A., J. A. Alonso-Pavon, L. G. Trevino-Quintanilla and J. Collado-Vides (2008). "Functional architecture of Escherichia coli: new insights provided by a natural decomposition approach." Genome Biol **9**(10): R154.
- Freyre-Gonzalez, J. A., L. G. Trevino-Quintanilla, I. A. Valtierra-Gutierrez, R. M. Gutierrez-Rios and J. A. Alonso-Pavon (2012). "Prokaryotic regulatory systems biology: Common principles governing the functional architectures of Bacillus subtilis and Escherichia coli unveiled by the natural decomposition approach." J Biotechnol **161**(3): 278-286.
- García, R. (2006). "Sistemas complejos." Barcelona: Gedisa: 202.
- Goh, K.-I., G. Salvi, B. Kahng and D. Kim (2006). "Skeleton and fractal scaling in complex networks." Physical review letters **96**(1): 018701.
- Ibarra-Arellano, M. A., A. I. Campos-Gonzalez, L. G. Trevino-Quintanilla, A. Tauch and J. A. Freyre-Gonzalez (2016). "Abasy Atlas: a comprehensive inventory of systems, global network properties and systems-level elements across bacteria." Database (Oxford) **2016**.
- Kenkel, N. and D. Walker (1996). "Fractals in the biological sciences." Coenoses: 77-100.
- Kim, J. S., K. I. Goh, B. Kahng and D. Kim (2007). "Fractality and self-similarity in scale-free networks." New Journal of Physics **9**(6): 177.
- Kitano, H. (2002). "Systems biology: a brief overview." Science **295**(5560): 1662-1664.
- Lima-Mendez, G. and J. van Helden (2009). "The powerful law of the power law and other myths in network biology." Mol Biosyst **5**(12): 1482-1493.
- Ma, H. W., J. Buer and A. P. Zeng (2004). "Hierarchical structure and modules in the Escherichia coli transcriptional regulatory network revealed by a new top-down approach." BMC Bioinformatics **5**: 199.
- Mandelbrot, B. B. (1975). "Stochastic models for the Earth's relief, the shape and the fractal dimension of the coastlines, and the number-area rule for islands." Proceedings of the National Academy of Sciences **72**(10): 3825-3828.
- Molontay, R. (2015). Fractal Characterization of complex Networks. MSc, Budapest University of Technology and Economics.
- Newman, M. E. J. (2011). Networks : an introduction. Oxford, Oxford University Press.
- Price, M. N., P. S. Dehal and A. P. Arkin (2007). "Orthologous transcription factors in bacteria have different functions and regulate different genes." PLoS Comput Biol **3**(9): 1739-1750.
- Resendis-Antonio, O., J. A. Freyre-Gonzalez, R. Menchaca-Mendez, R. M. Gutierrez-Rios, A. Martinez-Antonio, C. Avila-Sanchez and J. Collado-Vides (2005). "Modular analysis of the transcriptional regulatory network of E. coli." Trends Genet **21**(1): 16-20.

Simon, H. (1962). "The architecture of complexity." Proceedings of the American Philosophical Society(106(6)): 467-482.

Song, C., S. Havlin and H. A. Makse (2005). "Self-similarity of complex networks." Nature **433**(7024): 392-395.

Song, C. M., L. K. Gallos, S. Havlin and H. A. Makse (2007). "How to calculate the fractal dimension of a complex network: the box covering algorithm." Journal of Statistical Mechanics-Theory and Experiment.

Song, C. M., S. Havlin and H. A. Makse (2006). "Origins of fractality in the growth of complex networks." Nature Physics **2**(4): 275-281.