



Análisis In Silico de la Viabilidad de la Mutación de Sistemas Biológicos

In Silic Analysis of the Feasibility of Mutation of Biological Systems

Raidel Rodríguez Romeu, ¹ Ramón Alexander Jaime Infante, ² Julián Triana Dopico, ³ Joan Vázquez Molina ⁴

¹ Ingeniero en Informática. Máster en Informática en Salud. Profesor de Informática de la Universidad de Pinar del Río Hermanos Saiz Montes de Oca. Cuba. rirro@upr.edu.cu

² Ingeniero en Informática. Máster en Nuevas Tecnologías para la Educación. Profesor de Informática de la Universidad de Pinar del Río Hermanos Saiz Montes de Oca. Cuba. ramón@upr.edu.cu

³ Licenciado en Bioquímica. Doctor en Matemáticas. Universidad Politécnica Salesiana. Guayaquil. Ecuador. j triana96@gmail.com

⁴ Ingeniero Industrial. Doctor en Matemáticas. Universidad de Leuven. Bélgica. joavzmo@upv.es

Recibido: 30 de enero de 2018

Aprobado: 03 de mayo de 2018

Citar como: Rodríguez Romeu R, Jaime Infante RA, Triana Dopico J, Vázquez Molina J. Análisis In Silico de la Viabilidad de la Mutación de Sistemas Biológicos. Rev Ciencias Médicas [Internet]. 2018 [citado: fecha de acceso]; 22(4): 719-728. Disponible en: www.revcmpinar.sld.cu/index.php/publicaciones/article/view/3473

RESUMEN

Introducción: en Bioinformática la meta de los investigadores implica la obtención de modelos computacionales para procesos biológicos que se acerquen al comportamiento real, todo esto insertado en una disciplina conocida como Biología de Sistemas. El análisis *in silico* en esta área de las ciencias biológicas se apoya en el uso de métodos matemáticos-computacionales. Dentro del análisis en sistemas biológicos es de interés poder comparar la estructura de diferentes organismos. En dicha comparación se tendrá en cuenta el metabolismo de distintos organismos, así como la topología de su red metabólica asociada. Esta comparación sirve para seleccionar las especies o cepas que mejor se adapten a una aplicación en concreto.

Objetivo: analizar la viabilidad de la mutación de un sistema biológico para asumir funciones de otro.

Método: para este análisis se usaron los sistemas informáticos BioOpt, que se centra en el análisis del balance de flujo, utilizando la programación lineal como soporte matemático, y CompNet, que compara redes metabólicas atendiendo a varias métricas de similitud y las transformaciones necesarias para llevar de una red metabólica a otra.

Resultados: se hizo una comparación entre dos sistemas biológicos, pudiendo determinar las reacciones esenciales dentro del metabolismo de estos organismos y de ellas cuáles debían ser modificadas y cuáles eliminadas para lograr la mutación de un organismo a otro.

Conclusiones: este trabajo muestra un análisis *in silico* que ayuda en la determinación de si es viable o no la manipulación genética de un organismo para que asuma funciones que estén definidas en otro.

DeCS: BIOLOGÍA DE SISTEMAS; PROGRAMAS INFORMÁTICOS; BIOLOGÍA COMPUTACIONAL.

ABSTRACT

Introduction: in Bioinformatics the goal of researchers involves obtaining computational models for biological processes that are increasingly closer to real behavior, all this inserted into a general discipline known as Systems Biology. In silico analysis in this area of the biological sciences relies on the use of mathematical-computational methods. Within the analysis in biological systems it is of interesting to be able to compare the structure of different organisms. In this comparison, the metabolism of different organisms as well as the topology of their associated metabolic network will be taken into account. This comparison serves to select the species or strains that best suit a particular application.

Objective: to analyze the viability of the mutation of a biological system to assume functions of another.

Method: for this analysis, the BioOpt computer systems were used, which focuses on the analysis of the flow balance, using linear programming as a mathematical support and CompNet solution that compares metabolic networks based on several metrics and the necessary transformations to carry from one metabolic network to another.

Results: a comparison between two biological systems was made, being able to determine the essential reactions within the metabolism of these organisms and from which they should be modified and which ones eliminated to achieve the mutation from one organism to another.

Conclusions: this work shows an *in silico* analysis that could help in the determination of whether or not genetic manipulation of an organism is viable so that it assumes function that are defined in another.

DeCS: SYSTEMS BIOLOGY; SOFTWARE; COMPUTATIONAL BIOLOGY.

INTRODUCCIÓN

La bioinformática es una ciencia interdisciplinar que vincula varios campos de estudios. La base de esta área de las ciencias biológicas se centra en la utilización de herramientas computacionales para la simulación y análisis de problemas sobre escalas de tal magnitud que sobrepasan los métodos tradicionales de análisis y experimentación.

Una constante en proyectos de bioinformática es el uso de herramientas matemáticas para extraer información útil de datos producidos por técnicas biológicas de alta productividad, como la secuenciación del genoma. ⁽¹⁾Es por eso que la investigación en biología computacional se solapa a menudo con la Biología de Sistemas.

La Biología de Sistemas analiza el funcionamiento de las células desde un punto de vista global, centrándose en el sistema en su conjunto más que en las partes. ⁽²⁾ Este enfoque permite comprender íntegramente el funcionamiento de los sistemas biológicos y profundizar en el entendimiento de cómo sus interacciones internas y con otros sistemas conllevan a la aparición de nuevas propiedades, llamadas "propiedades emergentes". Al igual que la biología computacional, el análisis *in silico* en esta área de las ciencias biológicas se apoya en el uso de métodos matemáticos-computacionales.

Todo lo anterior hace de esta incipiente rama de la ciencia un área eminentemente multidisciplinar que requiere la colaboración de científicos de diferentes campos del conocimiento como la biología, la química, la informática, la ingeniería, con el objetivo final de poder llegar, en un futuro, a ser capaces de implementar nuevas funciones en sistemas vivos de forma dirigida. ⁽³⁾ Esta idea es la que sustenta y da origen a la Biología Sintética.

Un objetivo interesante dentro de esta ciencia es analizar la estructura de diferentes organismos, por ejemplo, para seleccionar las especies o las cepas que mejor se adapten a una aplicación en concreto. En el caso del uso de cianobacterias como plataformas de producción resulta de gran interés caracterizar el metabolismo asociado a la producción de hidrógeno: saber si son o no capaces de producirlo, clasificarlas según el tipo de hidrogenasa que emplean, determinar si el hidrógeno se produce como consecuencia de algún otro proceso metabólico como algunos casos en que está asociado a la fijación de nitrógeno. ⁽³⁾

La Ingeniería Genética es la manipulación directa de los genes de un organismo usando biotecnología. Esta ciencia ha tenido un enorme impacto en la biomedicina. Mediante la inserción, eliminación y trasposición de genes en diversos organismos es posible determinar la función de determinados genes. Una manera es mediante la técnica de knock-out, en la que un gen o grupo de genes son desactivados para observar las características que cambian en el fenotipo. Los primeros descubrimientos biotecnológicos se hicieron mediante el uso de microorganismos, dando inmediatamente con resultados asombrosos como el de la insulina transgénica, que fue la primera proteína recombinante aprobada como medicamento. Los organismos transgénicos son usados, entre otras

aplicaciones, para fabricar productos destinados para uso terapéutico en humanos, productos farmacéuticos. Algunos ejemplos son la producción de insulina transgénica, insulina humana pura producida por genes humanos introducidos en bacterias.

La vacuna contra la hepatitis B se produce con levaduras a la que se les ha insertado un gen para producir el antígeno HBsAg presente en la envoltura del virus de la hepatitis B. El anticoagulante ATryn, usado para reducir el riesgo de coágulo durante operaciones quirúrgicas, se extrae de la leche de cabra modificada genéticamente o la leuencefalina, un neurotransmisor producido por la soja transgénica. Otra de las últimas grandes aplicaciones de la biotecnología es la terapia génica, que usa virus genéticamente modificados para introducir genes en el ADN humano con el fin de curar algunas enfermedades. En este sentido, son conocidos sus éxitos en el tratamiento de desórdenes genéticos como la inmunodeficiencia combinada grave o la amaurosis congénita de Leber.⁽⁴⁾

Siguiendo esta idea, uno de los análisis desarrollados por los especialistas en esta ciencia es lograr la mutación de un organismo para asumir funciones de otro. Para lograr desarrollar este objetivo sería de interés poder comparar la estructura y funcionalidad de diferentes organismos. En dicha comparación se tendría en cuenta las capacidades metabólicas de distintos organismos, así como la topología de su red metabólica asociada. Este tipo de comparación corresponde al análisis *in silico*, previo al laboratorio y serviría para determinar si es viable o no la mutación de un organismo para asumir funciones que estén definidas en otro.

Teniendo en cuenta lo antes expuesto se plantea como problema de investigación: ¿Cómo contribuir al análisis *in silico* de la viabilidad de la manipulación genética de un sistema biológico para asumir funciones de otro?

A partir de este problema se define como objetivo general analizar la viabilidad de la mutación de un sistema biológico para asumir funciones de otro.

MÉTODO

Para el desarrollo de la investigación se emplearon métodos de investigación teórico y empírico. Se empleó el método análisis y síntesis, que posibilitó la extracción de los principales elementos que se relacionan con los análisis comparativos en la biología de sistemas y las herramientas que permiten dicha comparación. Se utilizó la entrevista para intercambiar directamente con especialistas en esta área de las ciencias biológicas y conocer sobre las herramientas computacionales que son empleadas en el análisis de la biología de sistemas, así como los organismos de más interés para su estudio. El método modelación permitió representar la información referente a los análisis comparativos hechos entre los organismos seleccionados y con el uso de las herramientas informáticas.

Partiendo del objetivo de esta investigación y luego de estudiar diferentes herramientas útiles en el análisis de la Biología de Sistemas se decidió utilizar para nuestra investigación los sistemas informáticos CompNet y BioOpt. En primer lugar, se tuvo en cuenta la idea de que, si dos sistemas biológicos se parecen, atendiendo a las características topológicas de su red metabólica, entonces deben ser más factible las variaciones genéticas para lograr en un sistema funcionalidades del otro. Sobre esta idea se sustentó el desarrollo y por tanto la utilización de CompNet para comparar dos organismos y valorar sus semejanzas utilizando métricas de similitud. Otra utilidad práctica usada en el sistema informático CompNet es la facilidad de proporcionar las transformaciones necesarias, dígame mutaciones genéticas, para lograr comportamientos similares en dos especies de organismos distintas. Esta herramienta es un producto de los resultados científicos del

Grupo de Modelación Matemática y Simulaciones Numérico-Computacionales, InterTech, de la Universidad de Pinar del Río en coordinación con el Grupo de Modelización Interdisciplinar, InterTech, de la Universidad Politécnica de Valencia.

Como se puede apreciar, para el análisis deseado es necesario tomar en cuenta las posibles mutaciones genéticas que deben ser impuestas a los modelos metabólicos. Desde el análisis *in silico* que se quiere realizar, estas mutaciones pueden ser entendidas como la eliminación, adición o modificación de reacciones y/o metabolitos. Cada red metabólica está formada por un conjunto de reacciones de las cuales algunas son esenciales y no admiten ningún tipo de cambio sin que se altere el funcionamiento metabólico del organismo. La herramienta BioOpt permite predecir el comportamiento de sistemas biológicos bajo perturbaciones y por tanto fue la herramienta seleccionada para este análisis. Con el uso de BioOpt podemos saber cuáles son las reacciones que, de ser eliminadas, afectarían significativamente el comportamiento de un sistema biológico.

La implementación del software CompNet está basado en la teoría de grafos, ya que una red metabólica puede ser modelada como un grafo, más exactamente como una Red de Petri. Para poder comparar dos modelos metabólicos es necesario tener una medida que indique qué tan parecidos o qué tan diferentes son. ⁽⁵⁾Para cuantificar la similitud entre dos modelos CompNet utiliza la métrica de Baláž: ⁽⁶⁾

$$d_a(G_i, G_j) = E_i + E_j - 2E_{MCS}(G_i, G_j) + V_i - V_j,$$

donde G_i y G_j representan los grafos a comparar, E_i y E_j representan el conjunto de aristas de los grafos i y j respectivamente, V_i y V_j representan el conjunto de vértices de los grafos i y j respectivamente y $E_{MCS}(G_i, G_j)$ representa el conjunto de aristas del máximo común subgrafo entre los grafos G_i y G_j . El valor de esta métrica da una idea de la cantidad de transformaciones necesarias para pasar de un modelo metabólico a otro. Otra métrica usada es la de Bunke: ⁽⁷⁾

$$d(G_i, G_j) = 1 - \frac{MCS(G_i, G_j)}{\max(V_i, V_j)}$$

$$\max(V_i, V_j)$$

donde G_i y G_j representan los grafos a comparar, V_i y V_j representan el conjunto de vértices de los grafos i y j respectivamente y $MCS(G_i, G_j)$ representa el conjunto de vértices del máximo común subgrafo entre los grafos G_i y G_j . Esta métrica proporciona un valor en el intervalo $[0,1]$ que indicando el nivel de similitud.

La herramienta informática BioOpt se centra en el análisis del balance de flujo, utilizando la programación lineal como soporte matemático. Dado un modelo de sistema biológico, que incluye un conjunto de reacciones metabólicas, el programa es capaz de calcular todos los flujos metabólicos que maximiza una función objetivo especificada, costos reducidos y precios sombra dependiendo de las restricciones y el objetivo definido por el usuario. Este software fue desarrollado bajo licencia de código abierto y con diferentes parámetros permite al usuario obtener varios tipos de salidas que pueden ayudar en el análisis del sistema. Estos parámetros permiten el estudio más general de las perturbaciones metabólicas después de la eliminación de genes. BioOpt implementa varios análisis para tratar este tipo de problema, incluyendo una eliminación simple de genes. ⁽⁸⁾En este sentido el software proporciona una implementación del modelo matemático planteado en el análisis de balance de flujos: ^(9,10)

Max (v_i)

sujeto a: $S \cdot v_j = 0 \quad \forall j \in N$

$v_j, \text{irr} \in \mathbb{R}^+$

$v_j, \text{rev} \in \mathbb{R}$

$v_j \in \mathbb{R}, v_{\min} < v_j < v_{\max}$

donde S representa la matriz estequiométrica, V el vector de flujo de las reacciones metabólicas, el subíndice i representa el valor de flujo de la reacción que se quiere optimizar y el subíndice j representa los demás valores de flujos.

Para lograr el objetivo planteado en esta investigación se usaron dos organismos unicelulares de interés en el estudio de la biología de sistemas: la bacteria *Streptomyces coelicolor* y el hongo *Amycolatopsis balhimycina*. Este interés está marcado por lograr la síntesis de vitaminas de manera dirigida en la bacteria *Streptomyces* basándonos en el modelo metabólico del *Amycolatopsis*, ya que este produce más vitaminas como las hidrosolubles tiaminas (vitamina B1), riboflavina (vitamina B2) y biotina.

El uso de la herramienta BioOpt permitió eliminar cada reacción de los modelos metabólicos de cada organismo uno por uno y obtener los valores de flujo de la función objetivo. El hecho de eliminar estas reacciones es lo que se considera mutaciones genéticas. Cuando eliminamos una reacción estamos adicionando una restricción más al modelo matemático. Al resolver el modelo con esta nueva restricción se pueden obtener tres estados diferentes de la función objetivo: puede variar su valor, no variar o puede generarse un nuevo estado en donde sea imposible encontrar solución al modelo.

Teóricamente cuando eliminamos una reacción y el valor de la función objetivo varía o no puede encontrarse un valor, se dice que esa reacción es esencial dentro del conjunto de reacciones metabólicas y no puede ser eliminada. Estas son las reacciones de más importancia para el análisis hecho y deben ser tomadas en cuenta a la transformación de un modelo en otro.

Una vez obtenidas las reacciones esenciales por cada modelo, se procedió a hacer la comparación entre los organismos. Se tuvieron en cuenta los valores de las métricas de Baláž, la métrica Bunke y cuáles reacciones debían ser eliminadas o modificadas para lograr la mutación deseada. En esta parte de la investigación fue usada la herramienta computacional CompNet. En este proceso se buscaba fundamentalmente que entre las transformaciones necesarias para llevar de un modelo a otro no fuera necesario eliminar o modificar una de sus reacciones esenciales.

RESULTADOS

Como primer resultado de esta investigación se tiene que al aplicar la eliminación simple de genes del sistema BioOpt se obtuvieron las reacciones esenciales de cada modelo metabólico (Tabla 1). Para la muestra de organismos usados en la investigación este tipo de reacciones constituyen el 21,29 por ciento en el caso del *Amycolatopsis* y el 36,71 por ciento en el caso del *Streptomyces* del total de la red metabólica

Tabla 1. Reacciones esenciales de los modelos metabólicos.

Organismos	Total de Reacciones	Reacciones Esenciales
<i>Amycolatopsis balhimycina</i>	803	171
<i>Streptomyces coelicolor</i>	700	257

Una vez que se obtuvieron los resultados del análisis hecho con el software BioOpt, se procedió a hacer la comparación entre los organismos usando el software CompNet. La comparación se hizo usando al *Streptomyces* como primer modelo y *Amycolatopsis* como segundo modelo. O sea, que el análisis se hizo buscando las mutaciones genéticas dirigidas necesarias para obtener en la red metabólica del *Streptomyces* las mismas reacciones que en la red metabólica del *Amycolatopsis* y así lograr los similares índices de síntesis de vitamina en *Streptomyces* con respecto al *Amycolatopsis*. El primer resultado obtenido en la comparación fueron las métricas de Baláž y Bunke (Tabla 2).

Tabla 2. Métricas de la comparación.

Modelo 1	Modelo 2	Baláž	Bunke
<i>Streptomyces coelicolor</i>	<i>Amycolatopsis balhimycina</i>	2344	0,394

El segundo resultado de la comparación fue saber del total de reacciones que tiene la red metabólica del *Streptomyces* cuáles deben ser eliminadas y cuáles deben ser modificadas (Tabla 3).

Tabla 3. Cantidad de reacciones eliminadas y modificadas del modelo *Streptomyces coelicolor*.

Modelo	Total de reacciones	Reacciones eliminadas	Reacciones modificadas
<i>Streptomyces coelicolor</i>	700	75	350

Una vez obtenidos los resultados de ambas herramientas se buscó cruzar unos con otros para verificar las posibles afectaciones que podrían sufrir las reacciones esenciales en el modelo de *Streptomyces*. La tabla cuatro muestras del total de reacciones a eliminar cuántas son esenciales.

Tabla 4. Cantidad de reacciones esenciales a eliminar del modelo.

Reacciones a eliminar	Reacciones esenciales	%
75	12	16

En la tabla 5 se muestra un análisis en cuanto al total de reacciones a modificar, cuántas son esenciales.

Tabla 5. Cantidad de reacciones esenciales a modificar del modelo.

Reacciones a modificar	Reacciones esenciales	%
351	82	23,36

DISCUSIÓN

Uno de los resultados alcanzados demuestra que hay doce reacciones esenciales que según la comparación deben ser eliminadas. En la práctica esto no sería posible, pues estaríamos afectando el valor de su función objetivo, que se traduce en limitar el crecimiento del organismo. El resto de las reacciones teóricamente podrían ser eliminadas sin afectar significativamente el crecimiento del *Streptomyces*.

En el caso de las reacciones que deben ser modificadas, su análisis en cuanto a si son esenciales o no es más flexible, puesto que estas reacciones no dejarán de existir. Estas reacciones deberán sufrir modificaciones respecto a la participación de los metabolitos reaccionantes y/o productos y sus coeficientes estequiométricos.

Teniendo en cuenta los resultados de la investigación, se puede decir que resulta interesante el análisis *in silico* con respecto a funcionamiento y topología de redes metabólicas en la manipulación genética de mutaciones. Determinar la participación o no de genes esenciales en mutaciones genéticas de un organismo para que asuma funciones de otro constituye una contribución al análisis de los sistemas biológicos previo al laboratorio. Este tipo de análisis podría ser tomado en cuenta en las investigaciones de ingeniería genética con el fin de lograr nuevas funcionalidades en sistemas vivos de forma dirigida.

Del análisis realizado se puede concluir que el estudio de los referentes teóricos con respecto al análisis *in silico* de la manipulación genética de sistemas biológicos permitió conocer los fundamentos de la Biología de Sistemas y las características de las principales herramientas usadas en el análisis de esta ciencia.

A partir de la investigación realizada de las herramientas que permitan hacer un análisis comparativo de sistemas biológicos en cuanto a funcionamiento y topología de su red metabólica, se decidió usar BioOpt y CompNet en el análisis comparativo de sistemas biológicos, teniendo en cuenta la complementación de ambas y sus fundamentos matemáticos.

Mediante el análisis de sistemas biológicos atendiendo a su funcionamiento y la topología de su red metabólica se contribuye a determinar la viabilidad de la manipulación genética de un sistema biológico para asumir funciones de otro.

REFERENCIAS BIBLIOGRAFICAS

1. Raies AB, Bajic VB. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdisciplinary Reviews: Computational Molecular Science [Internet]. 2016 [citado 2017 Dic 24]; 6(2): 147-72. Disponible en: <http://onlinelibrary.wiley.com/doi/10.1002/wcms.1240/full>.
2. Wolstencroft K, Owen S, Krebs O, Nguyen Q, Stanford NJ, Golebiewski M, et al. SEEK: a systems biology data and model management platform. BMC systems biology [Internet]. 2015 [citado 2017 Dic 24]; 9(1): 33. Disponible en: <https://bmcsystbiol.biomedcentral.com/articles/10.1186/s12918-015-0174-y>.
3. Baltes NJ, Voytas DF. Enabling plant synthetic biology through genome engineering. Trends in biotechnology [Internet]. 2015 [citado 2017 Dic 24]; 33(2): 120-31. Disponible en: [http://www.cell.com/trends/biotechnology/fulltext/S0167-7799\(14\)00237-6?returnURL=http%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0167779914002376%3Fshowall%3Dtrue](http://www.cell.com/trends/biotechnology/fulltext/S0167-7799(14)00237-6?returnURL=http%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS0167779914002376%3Fshowall%3Dtrue)
4. Castilla Guerra J. La autorización de los organismos genéticamente modificados [Tesis]. España: Universidad de Sevilla; 2016. Disponible en: <https://idus.us.es/xmlui/bitstream/handle/11441/39186/TESIS.%20Autorizaci%3%b3n%20OMG..pdf?sequence=1&isAllowed=y>.
5. Infante RAJ. Empleo de las TIC´s en el aprovechamiento de la biomasa forestal en Pinar del Río. II Congreso Internacinal de Marketing Desarrollo Local y Turismo MARDELTUR 2017; Pinar Del Río. 2017. Disponible en: <https://www.responsibletravel.org/docs/MARDELTUR2017.pdf>
6. Baláž V, Koča J, Kvasnička V, Sekanina M. A metric for graphs. Časopis pro pěstování matematiky [Internet]. 1986 [citado 2017 Dic 24]; 114(2): 255-9. Disponible en: https://dml.cz/bitstream/handle/10338.dmlcz/108715/CasPestMat_114-1989-2_5.pdf
7. Bunke H, Shearer K. A graph distance metric based on the maximal common subgraph. Pattern recognition letters [Internet]. 1998 [citado 2017 Dic 24]; 19(3): 255-9. Disponible en: <http://biomet-toolbox.chalmers.se/index.php?page=downtools-bioOpt>.
8. BioOpt; 2017. Disponible en: <http://biomet-toolbox.chalmers.se/index.php?page=downtools-bioOpt>
9. Edwards J, Ramakrishna R, Schilling C, Palsson B. Metabolic flux balance analysis. BIOPROCESS TECHNOLOGY. 1997; 24:13-58.
10. Komi N, Xingxing J, Xinkai Z, Liuqing W, Jiajia L, Catherine M, Qiang H. Flux Balance Analysis Inspired Bioprocess Upgrading for Lycopene Production by a Metabolically Engineered Strain of *Yarrowia lipolytica* Metabolites [Internet]. 2015 Dec [citado 2017 Dic 24]; 5(4): 794-813. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4693195/pdf/metabolites-05-00794.pdf>

Raidel Rodríguez Romeu. Ingeniero en Informática. Máster en Informática en Salud. Profesor de Informática de la Universidad de Pinar del Río Hermanos Saiz Montes de Oca. Cuba. ***Si usted desea contactar con el autor de la investigación hágalo [aquí](#)***