

<i>Nereis. Revista Iberoamericana Interdisciplinar de Métodos, Modelización y Simulación</i>	3	25-31	Universidad Católica de Valencia "San Vicente Mártir"	Valencia (España)	ISSN 1888-8550
--	---	-------	---	-------------------	----------------

## Desarrollo de una plataforma computacional para el modelado metabólico de microorganismos

Fecha de recepción y aceptación: 09 de noviembre de 2010, 30 de noviembre de 2010

R. Reyes†, J. Garrido†, R. A. Jaime†, V. Córdova†, J. Triana†, L. Villar†, J. C. Castro†, P. Fernández de Córdoba‡, J. F. Urchuguía‡, E. Navarro‡ y A. Montagud‡

† Universidad de Pinar del Río, Cuba.

‡ Universidad Politécnica de Valencia, España.



### ABSTRACT

Synthetic biology focuses on the design and construction of artificial genetic systems that are capable of carrying out a specific function after being inserted into a living system. With the development of synthetic biology a new generation of bioengineers has appeared who develop complex, highly integrated genetic biological pathways. The improvement of this scientific discipline aims to establish a computational and conceptual framework that will support the development of modular artificial biological systems based on an engineering and systematic methodology. To achieve this, it will be necessary to provide new integrated computational tools in a common environment for the analysis of metabolic phenotypes, the design of new complex genetic pathways and the visualisation of metabolic maps to the next generation of designers in synthetic biology and future biotechnologists and biological engineers.

A result of this research is the Hydra platform (Hybrid Draw and Routes Analysis) that integrates various tools for the design, analysis, and visualisation of metabolic networks.

**KEYWORDS:** *bioinformatics, biology of systems, synthetic biology, metabolic pathway.*

### RESUMEN

La Biología Sintética (BS) se centra en el diseño y la construcción de sistemas genéticos artificiales, capaces de desarrollar una función específica después de haber sido introducidos en un sistema vivo. Con el desarrollo de la BS, se observa una nueva generación de bioingenieros que desarrollan complejos circuitos biológicos genéticos con un alto nivel de integración. La mejora de esta disciplina científica tiene por objeto establecer un marco computacional y conceptual que dé asistencia al desarrollo de sistemas biológicos artificiales modulares basándose en una metodología ingenieril y sistemática, para lo que se necesita proveer a la próxima generación de diseñadores en Biología Sintética y a los futuros biotecnólogos e ingenieros biológicos de nuevas herramientas computacionales integradas en un entorno común para el análisis de fenotipos metabólicos, el diseño de nuevos circuitos genéticos complejos y la visualización de mapas metabólicos.

Como resultado de esta investigación se obtiene la plataforma Hydra (Hybrid Draw and Routes Analysis), que integra diversas herramientas para el diseño, análisis y visualización de las redes metabólicas.

**PALABRAS CLAVE:** *bioinformática, biología de sistemas, biología sintética, ruta metabólica.*

### INTRODUCCIÓN

La bioinformática, según una de sus definiciones más sencillas, es la aplicación de la tecnología de computadores a la gestión y análisis de datos biológicos. Los términos bioinformática, biología computacional y, en ocasiones, biocomputación, utilizados en muchas situaciones como sinónimos, hacen referencia a campos de estudios interdisciplinarios muy vinculados, que requieren el uso o desarrollo de diferentes técnicas que incluyen la informática, la matemática aplicada, la estadística, las ciencias de la computación, la inteligencia artificial, la química y la bioquímica para solucionar problemas, analizar datos o simular sistemas o mecanismos, todos ellos de índole biológica, y usualmente (pero no de forma exclusiva) en el nivel molecular. El

núcleo principal de estas técnicas se encuentra en la utilización de recursos computacionales para solucionar o investigar problemas sobre escalas de tal magnitud que sobrepasan el discernimiento humano. La investigación en biología computacional se solapa a menudo con la Biología de Sistemas [1].

La Biología de Sistemas es un área de investigación científica que se preocupa del estudio de procesos biológicos. Comenzó a desarrollarse en los años sesenta del siglo XX, si bien su institucionalización académica no se produjo hasta el año 2000. Se centra en estudiar sistemas biológicos en su totalidad, usando herramientas de modelación, simulación y comparación al experimentado.

Por otro lado, surge la Biología Sintética como una evolución conceptual de la Biología de Sistemas. Esta ciencia trata de entender



el funcionamiento de las células desde un punto de vista global, centrándose en el sistema en su conjunto más que en las partes. Con el advenimiento de las técnicas experimentales de alto rendimiento se ha alcanzado un importante grado de desarrollo, hasta el punto de haber conseguido desarrollar modelos en los que el conjunto de reacciones que tienen lugar en una célula se representan de manera análoga a una red eléctrica o de comunicaciones, así como todo un conjunto de herramientas para operar sobre dichas redes [2].

Aunque la expresión Biología Sintética (BS) ha estado presente en la literatura científica y técnica desde la década de 1910, en la actualidad se ha convertido en un concepto que abarca la totalidad de las investigaciones desarrolladas en la frontera entre la Biología Molecular y la Ingeniería en sentido estricto [3]. Como tal, el término BS se está convirtiendo en un concepto cada vez más comprensivo, que:

1. Abarca nuevos marcos teóricos que se ocupan de los sistemas biológicos con las herramientas conceptuales y el lenguaje descriptivo de la Ingeniería.
2. Utiliza enfoques de diseño de organismos con propiedades novedosas inspirados en el diseño de circuitos eléctricos y mecánicos.
3. Persigue la creación de nuevos organismos con propiedades “a la carta” basada en la combinación racional de partes biológicas disociadas de su contexto natural.

Una de las características generales de los métodos automatizados es que producen conjuntos enormes de resultados que sólo se pueden analizar haciendo uso intensivo de herramientas bioinformáticas para facilitar la organización y sistematización de la información.

## MÉTODOS

La reconstrucción a escala genómica de una red metabólica es actualmente un proceso no automatizado e iterativo de toma de decisiones que fácilmente requiere el trabajo de al menos una persona/año para agrupar satisfactoriamente la lista de reacciones metabólicas para un organismo específico.

La reconstrucción de la red metabólica es un proceso de varios pasos en los que se utiliza esencialmente información genómica, metabolómica y datos fisiológicos del organismo en cuestión. El proceso implica la recopilación de todas las reacciones enzimáticas conocidas en las rutas metabólicas del sistema biológico, así como los genes asociados a la codificación de cada una de estas enzimas. Por otra parte, las publicaciones en revistas pueden ser utilizadas para identificar alguna información específica sobre las reacciones metabólicas, así como sobre sus genes relacionados.

La gran cantidad de información almacenada en bases de datos públicas, como por ejemplo información genómica [4], vías metabólicas [5], las enzimas [6] o las proteínas [7], puede ser recopilada para un organismo específico. Sin embargo, la falta de calidad debe ser considerada como uno de los principales inconvenientes de algunas de las bases de datos: falsos positivos, falsos negativos, así como objetos anotados erróneamente pueden obstaculizar los esfuerzos para reunir datos exactos [8]. En consecuencia, la reconstrucción manual a través de un control minucioso de todas y cada una de las reacciones, la ecuación de biomasa basada en moléculas constituyentes (como aminoácidos y nucleótidos), la coherencia y la integridad de la red son requisitos previos para la generación de un modelo metabólico útil y de alta calidad [9].

Aunque existen distintos software (ejemplo: Pathways Tools, Cobra, Simulink, CellDesigner) capaces de acceder a dichas bases de datos con

diferentes motivaciones, se puede decir que, aunque estos incorporan ciertos mecanismos de “depuración” y filtrado de la información, aún requieren un gran esfuerzo a nivel de desarrollo que permita reducir el tiempo que se necesita para realizar determinadas tareas. Por lo tanto, se hace necesario diseñar una base de datos que permita traducir la información biológica contenida en diversas bases de datos en un formato estándar que pueda ser usado ampliamente por la comunidad. Es el caso de SBML (del inglés, Systems Biology Markup Language), que permite su integración con distintas herramientas para el análisis y diseño de sistemas biológicos con funcionalidades novedosas, garantizando la coherencia de los datos almacenados en diferentes bases de datos. Este lenguaje es un formato legible por ordenador para la representación de modelos de procesos biológicos. Es aplicable a simulaciones de metabolismo, la señalización celular, la regulación de genes y visualización de mapas metabólicos y muchos otros temas.

La adopción de SBML ofrece muchos beneficios, algunos de los cuales son [10]:

1. Permitir el uso de múltiples herramientas sin necesidad de reescribir los modelos para cada herramienta.
2. Crear modelos para ser compartidos y publicados en una forma que otros investigadores puedan utilizar incluso en un ambiente de software diferente.
3. Garantizar la supervivencia de los modelos (y el esfuerzo intelectual puesto en ellos) más allá de la vida del software utilizado para crearlos.

### Base de datos. Contenedores de información biológica

Una base de datos o banco de datos es un conjunto de datos pertenecientes a un mismo contexto y almacenados sistemáticamente para su posterior uso. En este sentido, una biblioteca puede considerarse una base de datos compuesta en su mayoría por documentos y textos impresos en papel e indexados para su consulta. En la actualidad, y debido al desarrollo tecnológico de campos como la informática y la electrónica, la mayoría de las bases de datos están en formato digital (electrónico), que ofrece un amplio rango de soluciones al problema de almacenar datos [11]. Existen programas denominados Sistemas Gestores de Bases de Datos, abreviado SGBD, que permiten almacenar y posteriormente acceder a los datos de forma rápida y estructurada. Las propiedades de estos SGBD, así como su utilización y administración, se estudian dentro del ámbito de la informática. Las aplicaciones más usuales son para la gestión de empresas e instituciones públicas. También son ampliamente utilizados en entornos científicos con el objeto de almacenar la información experimental.

Una base de datos biológica es una biblioteca de información sobre ciencias de la vida, recogida de experimentos científicos, literatura publicada, tecnología de experimentación de alto rendimiento y análisis computacional. Contiene información de áreas de investigación como la genómica, la proteómica, la metabolómica, la expresión génica mediante microarrays y la filogenética. La información contenida en bases de datos biológicas incluye funciones, estructura y localización (tanto celular como cromosómica) de genes, efectos clínicos de mutaciones, así como similitudes de secuencias y estructuras biológicas [12].

### Modelos de datos para la información biológica

Un modelo de base de datos o esquema de base de datos es la estructura o el formato de una base de datos, descrita en un lenguaje formal y soportada por el Sistema de Gestión de Bases de Datos. En



otras palabras, un “modelo de base de datos” es la aplicación de un modelo de datos usado en conjunción con un sistema de gestión de bases de datos. Los esquemas generalmente son almacenados en un diccionario de datos. Aunque un esquema se defina en un lenguaje de base de datos de texto, el término a menudo es usado para referirse a una representación gráfica de la estructura de la base de datos. Un modelo de datos no es solamente un modo de estructurar datos, sino que también define el conjunto de las operaciones que pueden ser realizadas sobre los datos [13].

Varias técnicas son usadas para modelar la estructura de datos. La mayor parte de los sistemas de bases de datos son construidos en torno a un modelo de datos particular, aunque sea cada vez más común para ciertos productos ofrecer el apoyo a más de un modelo. La concepción teórica de un modelo pertinente de datos asegurará el diseño adecuado de las bases de datos de información biológica a través de la coherencia, estandarización y completitud de los elementos responsables de las funciones biológicas, para el diseño y creación de circuitos genéticos.

### Modelo de datos Relacional

Una base de datos relacional es una base de datos que cumple con el modelo relacional, que es el modelo más utilizado en la actualidad para modelar problemas reales y administrar datos dinámicamente. Este modelo permite establecer interconexiones (relaciones) entre los datos (que están guardados en tablas), y trabajar con ellos conjuntamente. Tras ser postuladas sus bases en 1970 por Edgar Frank Codd, de los laboratorios IBM en San José (California), no tardó en consolidarse como un nuevo paradigma en los modelos de base de datos [14].

Para manipular la información se utiliza un lenguaje relacional. Actualmente se cuenta con dos lenguajes formales: el álgebra relacional y el cálculo relacional. El álgebra relacional permite describir la forma de realizar una consulta, en cambio el cálculo relacional sólo indica lo que se desea devolver. El lenguaje más común para construir las consultas en bases de datos relacionales es SQL (Structured Query Language), un estándar implementado por los principales motores o sistemas de gestión de bases de datos relacionales. Las bases de datos relacionales pasan por un proceso que se conoce como normalización de una base de datos, que es entendido como el proceso necesario para que una base de datos sea utilizada de manera óptima. Finalmente, la base de datos diseñada se basa en un modelo relacional para almacenar y gestionar la información biológica a partir de las relaciones existentes entre cada uno de los componentes biológicos.

### Web service

Un servicio web (WS) (en inglés, Web Service) es un conjunto de protocolos y estándares que sirven para intercambiar datos entre aplicaciones. Distintas aplicaciones de software desarrolladas en lenguajes de programación diferentes y ejecutadas sobre cualquier plataforma pueden utilizar los servicios web para intercambiar datos en redes de ordenadores como Internet. La interoperabilidad se consigue mediante la adopción de estándares abiertos. Para mejorar la interoperabilidad entre distintas implementaciones de servicios web se ha creado el organismo WS-I, encargado de desarrollar diversos perfiles para definir de manera más exhaustiva estos estándares [15]. Este servicio utiliza WSDL (en inglés, Web Services Description Language), que es un formato XML para describir los servicios de red como un conjunto de variables que operan en los mensajes que contienen, ya sea orientado o procedimiento orientadas a la información del documento [15]. El lenguaje WSDL es extensible para permitir la descripción de los puntos finales y sus mensajes sin importar qué formatos de mensaje o protocolos de red se utilizan para comunicarse.

### Kegg API

Kegg proporciona valiosos medios para acceder a la información disponible en dicha base de datos, tanto para la búsqueda en los procesos bioquímicos celulares como para analizar el universo de los genes en los genomas completamente secuenciados. Los usuarios pueden acceder al servidor de Kegg API mediante la tecnología SOAP (en inglés, Simple Object Access Protocol) a través del protocolo HTTP (en español, protocolo de transferencia de hipertexto). El servidor SOAP también viene con el WSDL, lo que hace que sea fácil de construir una biblioteca de cliente para un lenguaje de programación específico [16]. Esto permite a los usuarios escribir sus propios programas para diversos propósitos y para automatizar el procedimiento de acceso al servidor Kegg API y recuperar los resultados.

Haciendo uso de Kegg API, se puede acceder a todas las funcionalidades de Kegg que fueron implementadas para el almacenamiento y la gestión de la información a diversos niveles de información biológica. De ahí que, una vez creado el Web Service Cliente para lograr una conexión entre las diferentes aplicaciones, se diseña e implementa una aplicación de escritorio, que logrará usar toda la información almacenada en Kegg para descargarla en la base de datos diseñada con este fin. La aplicación para gestionar el Web Service Cliente fue desarrollada en Java como lenguaje de programación y la información almacenada en Postgres como gestor de bases de datos. A continuación se describen algunas características de las herramientas utilizadas.

Java es un lenguaje de programación orientado a objetos, desarrollado por Sun Microsystems a principios de los años 90. El lenguaje en sí mismo toma mucha de su sintaxis de C y C++, pero tiene un modelo de objetos más simple y elimina herramientas de bajo nivel, que suelen inducir a muchos errores, como la manipulación directa de punteros o memoria. Entre diciembre de 2006 y mayo de 2007, Sun Microsystems liberó la mayor parte de sus tecnologías Java bajo la licencia GNU GPL, de acuerdo con las especificaciones del Java Community Process, de tal forma que prácticamente todo el Java de Sun es ahora software libre (aunque la biblioteca de clases de Sun que se requiere para ejecutar los programas Java aún no lo es) [17].

Postgres SQL está considerado el Sistema Gestor de Bases de Datos (SGBD) de código abierto más avanzado del mundo. Proporciona un gran número de características que normalmente sólo se encontraban en las bases de datos comerciales de alto calibre tales como Oracle.

Es un SGBD objeto-relacional, ya que aproxima los datos a un modelo objeto-relacional, y es capaz de manejar complejas rutinas y reglas. Su avanzada funcionalidad se pone de manifiesto con las consultas SQL (del inglés Structured Query Language) declarativas, el control de concurrencia multiversión, el soporte multiusuario, las transacciones, la optimización de consultas, la herencia y los valores no atómicos (atributos basados en vectores y conjuntos) [18].

La base de datos desarrollada permite suplir algunas de las deficiencias encontradas por la comunidad de Biología Sintética con relación a la coherencia, estandarización y completitud de los elementos responsables de las funciones biológicas. En este sentido, se ha logrado llevar a cabo un proceso de depuración de la información con vistas a evitar redundancias. El equipo ha trabajado en función de la estandarización de la información en aras de lograr términos comunes dentro de la comunidad de científicos. Además se trabaja en el desarrollo de algoritmos para lograr la completitud de la información, a través de otras bases de datos disponibles en la Red.



## RESULTADOS

Una vez obtenida toda la información biológica, y supliendo las deficiencias que actualmente persisten en diversas bases de datos, se genera un fichero SBML con la información referente a un organismo específico, lo que servirá de entrada a la nueva plataforma computacional para el diseño, análisis y representación de las rutas metabólicas.

Para la representación de la información usando SBML, se decide trabajar con el Nivel 2, Versión 1. El SBML Nivel 2 fue concebido en el V Taller de Plataformas de Software para Sistemas Biológicos, que se celebró en julio de 2002 en la Universidad de Hertfordshire [19]. Por esas fechas, cada vez más personas se implicaron en el grupo de colaboradores del SBML, de manera que se convirtió en un esfuerzo comunitario, con muchas herramientas nuevas que se iban mejorando para soportar SBML. Los participantes en el taller de 2002 decidieron colectivamente revisar la forma de SBML en el Nivel 2. El primer borrador de la especificación Versión 1 Nivel 2 fue liberado en agosto de 2002 y el grupo final de características se concluyó en mayo de 2003 en el VII Taller de Plataformas de Software para Biología de Sistemas en Ft. Lauderdale, Florida. Se trabaja en esta versión de Hydra con Nivel 2, Versión 1 de SBML por ser una de las más usadas en la comunidad de biólogos y por lo tanto de las más respaldadas en cuanto a pruebas realizadas.

### Hydra

La integrabilidad de herramientas es algo muy esencial hoy en día en el trabajo de la Biología Sintética, por ello nuestra plataforma informática HYDRA (Hybrid Draw and Routes Analysis) integra un grupo de herramientas tales como FBA (Flux Balance Analysis) y RON Painter (Rational Organism Network Painter), las cuales se describirán a continuación.

### FBA (Flux Balance Analysis)

La cuantificación de los flujos de la red metabólica se puede lograr a través del análisis de los flujos metabólicos. Este método usa la estequiometría de las reacciones metabólicas intracelulares, los flujos extracelulares y los flujos internos conocidos como entrada y retorno. Bajo este enfoque, se establece una relación entre las concentraciones de metabolitos intracelulares, los flujos de reacción y la estequiometría en un sistema metabólico por una ecuación de balance dinámico de masa [20].

Un sistema químico se puede describir mediante un sistema de ecuaciones diferenciales y se puede generalizar a cualquier red metabólica y representar de una manera más fácil mediante el uso de matrices, denominadas matriz estequiométrica. Un sistema está en estado estacionario cuando las tres derivadas (índices del cambio) son 0, o sea, cuando las concentraciones de estas sustancias no varían en el tiempo. El FBA implica realizar un análisis del estado estacionario característico de los sistemas biológicos, utilizando el sistema matricial que se obtiene del conjunto de reacciones.

De ahí que surja la aplicación FBA (Flux Balance Analysis), dentro de Hydra, para la gestión de una matriz estequiométrica que representa los flujos metabólicos en las rutas metabólicas. A partir de su integración en una aplicación web y sin necesidad de usar un software independiente, se accede a potentes herramientas alojadas en el servidor y de uso sencillo para gente no experta en programación ni en matemática discreta.

Con esta aplicación y su integración con Ron Painter, se podrá representar en las rutas metabólicas el valor de flujo de biomasa asociado a cada una de las reacciones.

### RON Painter

RON Painter es una aplicación web que se inserta dentro de la plataforma Hydra. Se implementó con ASP.NET, C# y XML, utilizando como entorno de desarrollo Visual Studio.NET 2008 y beneficiándonos de la programación orientada a objetos, de un fuerte chequeo estático de tipos, de los componentes estándares que ofrece la biblioteca de clases y componentes de .NET Framework. Con las potencialidades que ofrece la plataforma y haciendo uso de zonas calientes se fue interactuando entre las páginas donde se representan el mapa metabólico y las rutas que lo integran.

Para el testeo de la aplicación se utilizó el fichero SBML correspondiente a la *Synechocystis*. sp *PCC6803*, aunque se debe destacar que RON Painter tiene como proyección futura hacer la visualización no solo de esta especie, sino del resto de las especies que están reconocidas en nuestra comunidad científica, contribuyendo así al avance de las herramientas para el uso de la biología y la investigación en esta rama.

## ANÁLISIS

Hydra fue implementada haciendo uso de los siguientes lenguajes y plataformas de desarrollo, lo que favorece la construcción de una aplicación segura y compatible con el resto de las aplicaciones implementadas y de las que están aún en desarrollo.

### Xml

Es un formato de texto simple y flexible que puede ser usado como base para crear nuevos lenguajes de marca (como SBML y otros) que puedan ser usados para la publicación y el intercambio de documentos. Está basado en una serie de recomendaciones publicadas por un grupo de trabajo del W3C (World Wide Web Consortium) y por estas razones se encuentra en una inmejorable situación para ser usado en aplicaciones web [21].

### C#

Es un lenguaje de programación diseñado para construir un amplio rango de aplicaciones que se ejecutarán sobre .Net Framework. Es simple, poderoso, fuertemente tipado, con un fuerte chequeo estático de tipos y orientado a objetos. Con sus muchas innovaciones C# nos permite desarrollar aplicaciones rápidamente manteniendo la expresividad y la elegancia de los lenguajes estilo C [22].

### Visual Studio 2008

Visual Studio .Net es una suite completa de herramientas para crear, entre otras cosas, aplicaciones de escritorio y aplicaciones web. Se utiliza para desarrollar aplicaciones de alto rendimiento beneficiándonos de poderosas herramientas de desarrollo basadas en componentes y otras tecnologías para simplificar el trabajo en equipo. Visual Studio .Net soporta Visual C# con un muy cómodo Editor de Código, distintos tipos de proyectos, una confortable interfaz de diseño gráfico, un poderoso depurador fácil de usar y otras herramientas. La biblioteca de clases de .Net Framework nos permite acceder a un amplio espectro



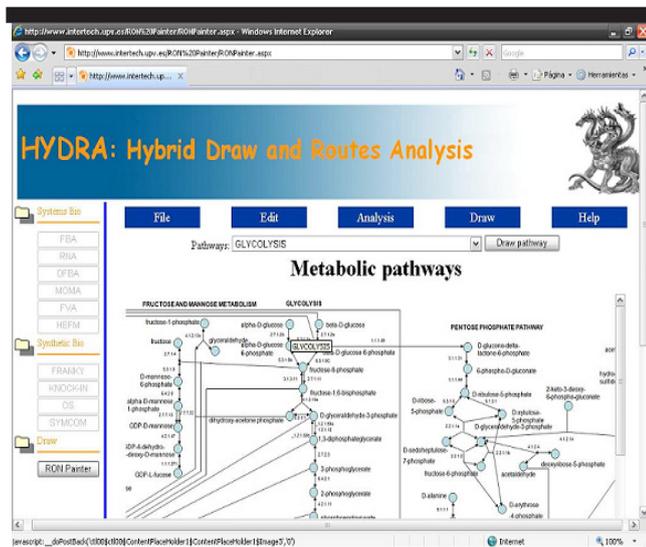


Figura 1. Visualización de una red metabólica haciendo uso de Hydra.

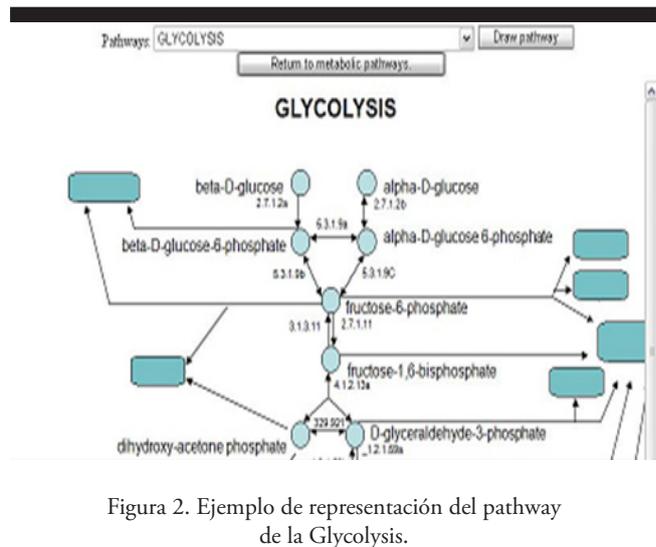


Figura 2. Ejemplo de representación del pathway de la Glycolysis.

de servicios y clases bien diseñadas que aceleran el ciclo de desarrollo significativamente [23].

**ASP.NET**

ASP.NET (parte de .NET Framework) brinda clases y herramientas para crear aplicaciones web dinámicas. Es un modelo unificado de desarrollo web que incluye los servicios que necesitamos para crear aplicaciones web profesionales con un mínimo de código y esfuerzo por parte de los programadores.

Desarrollando aplicaciones web con ASP.NET tenemos acceso a las clases de .NET Framework. Estas aplicaciones se pueden desarrollar utilizando cualquiera de los lenguajes compatibles con el CLR (Common Language Runtime), como por ejemplo C#, JScript .NET y J# entre otros. Estos lenguajes nos permiten desarrollar aplicaciones ASP.NET que se benefician del CLR, de la programación orientada a objetos, entre otros [24].

La plataforma Hydra está publicada en el sitio <www.intertech.upv.es>. En éste existe un vínculo que se denomina Links donde se encuentra el enlace a Hydra. La interfaz de la herramienta tiene un contraste de colores que identifica a nuestro grupo de desarrollo Intertech. En la parte superior se muestra el nombre de la plataforma, en su parte inferior se visualiza la última fecha de actualización de la herramienta y un link a la dirección del correo electrónico del administrador de esta. A la izquierda se muestran las herramientas que integrarán esta plataforma y la herramienta RON Painter, donde para acceder a ésta se ejecuta un clic. Una vez que se examina y se hace la selección del fichero con extensión sbml se muestra en pantalla el mapa metabólico de la bacteria, dando además la oportunidad de acceder a las rutas que componen el mapa de dos formas: una es a través de un combobox donde al desplegarse se muestran todas las rutas existentes, y la otra es haciendo clic sobre un metabolito en cuestión. Esto da la posibilidad de que se muestre la ruta donde este metabolito se expresa, como se ve en la figura 1.

El mapa metabólico contiene todas las reacciones con su identificador, el nombre de los metabolitos y el nombre de las rutas. Por otra parte, en las rutas se muestra la misma información pero, además, se muestran todas las rutas donde se expresa un metabolito dado a través de cajas

rectangulares. En el momento en que está visualizada una ruta se activa el botón FBA, y al hacer clic sobre éste, se muestra, en la ruta, el valor de flujo de biomasa asociado a cada reacción. También se puede ir desde una ruta a otras haciendo clic sobre las cajas, como se muestra en la figura 2.

**DISCUSIÓN**

Existen diversas herramientas para generar mapas metabólicos, sin embargo no preservan la forma en que tradicionalmente se representan en biología, dificultando el trabajo y reconocimiento de estos por parte de los investigadores con bagaje biológico. Algunos ejemplos son Cytoscape [25], CellDesigner [26] y Copasi [27], que hacen la representación del dibujo de una forma heurística y poco racional. Ni siquiera Kegg ha resuelto el problema, pues esta herramienta muestra los mapas y las rutas en toda su extensión, sin poder descartar los metabolitos que no están presentes en la red, lo que dificulta la claridad del modelo. Con lo cual se puede decir que Ron Painter suple estas deficiencias con la visualización de un modelo metabólico mucho más claro dentro de la comunidad de biólogos, a partir del análisis de los metabolitos que intervienen en cada una de las rutas y la generación de un modelo mucho más claro.

Por su parte, Hydra integra este conjunto de herramientas para el análisis, diseño y representación de redes metabólicas, lo que favorece el acceso rápido y directo a las distintas rutas que componen el mapa metabólico y la descripción de éstas, representando el nombre de cada metabolito, el identificador de cada reacción y la dirección de las rutas, así como mostrando el valor de flujo de cada reacción. La plataforma aún está en desarrollo, lo que facilitará la integración con nuevas herramientas en desarrollo para el análisis de los modelos metabólicos.

Es preciso destacar que la plataforma Hydra basa sus pruebas de modelado en la *Synechocystis*.sp PCC6803, por ser una cianobacteria considerada un buen candidato para la producción de una variedad de moléculas orgánicas de interés industrial, utilizando como fuente de carbono el CO<sub>2</sub> (y/o azúcares) y la luz solar (y/o azúcares) como fuente



de energía [28]. La diversidad de aplicaciones potenciales en este sentido es amplia. Han sido publicados trabajos acerca de la producción heteróloga de varios metabolitos, entre los que se encuentra el biohidrógeno [29, 30]. Este elemento es producido gracias a la presencia de enzimas que intervienen directamente en su metabolismo, como las [Ni-Fe] hidrogenasas, las [Fe-Fe] hidrogenasas y las nitrogenasas, donde el hidrógeno es producido como un subproducto de la fijación de nitrógeno. Es bien conocido que la inhibición de la actividad de las hidrogenasas por el oxígeno formado en la fotosíntesis es el factor más importante que limita una producción rentable de hidrógeno utilizando este mecanismo [31].

No obstante, la aplicación Hydra servirá para modelar comportamientos en cualquier microorganismo de estudio.

## CONCLUSIONES

En los últimos años, las aplicaciones de la biotecnología en diferentes ámbitos de la ciencia y la tecnología se han multiplicado de forma considerable en paralelo al incremento exponencial de la información que se posee sobre los organismos. La información biológica, obtenida mediante diferentes tecnologías cada vez más poderosas y eficientes está pasando a formar parte de grandes bases de datos, muchas de ellas de libre acceso. Por su parte, la Biología Sintética se ha convertido en un área emergente de alto potencial científico y tecnológico e importancia estratégica para el mundo. Sin embargo la Biología Sintética precisa un gran impulso conceptual y metodológico para alcanzar su madurez como campo de desarrollo científico y tecnológico. Con este fin se ha logrado diseñar e implementar una plataforma computacional que integra un conjunto de herramientas para el diseño, análisis y visualización de las redes metabólicas de microorganismos, basándose en las pruebas de la aplicación en la *Synechocystis*.sp PCC6803.

## AGRADECIMIENTOS

Los autores desean agradecer el soporte financiero recibido por el Ministerio de Ciencia e Innovación a través de la concesión TIN2009-12359; la Conselleria de Inmigración y Ciudadanía de la Generalitat Valenciana (concesión 3012/2009) y la Comisión Europea (Proyecto TARPOL FP7 EU KBBE 212894).

## BIBLIOGRAFÍA

1. Wikipedia, Enciclopedia Libre. Documentación libre de GNU. [Internet]. *Bioinformática*. Disponible en: [es.wikipedia.org](http://es.wikipedia.org)
2. USDOE. 2005. Genomics: GTL Roadmap: *Systems biology for energy and environment*. DOE/SC 0090. Disponible en: <http://genomicsgdl.energy.gov>. USDOE, Office of Science, Washington, DC.
3. DASIKA M., 2008 OptCircuit: An optimization based method for computational design of genetic circuits. *BMC Systems Biology* 2008, 2,24.
4. GAFFRON H. Reduction of CO<sub>2</sub> with H<sub>2</sub> in green plants. *Nature* 1939;143:204–5.
5. APPEL J., SCHULZ R., 1998. Hydrogen metabolism in organisms with oxygenic photosynthesis: hydrogenases as important regulatory devices for a proper redox poising? *Journal of Photobiology and Photochemistry* 1998;47:1–11.
6. TING CS., ROCAP G., KING J, CHISHOLM SW., 2002. Cyanobacterial photosynthesis in the oceans: the origins and significance of divergent lightharvesting strategies. *Trends Microbiol* 10: 134–142
7. KASHINO Y., LAUBER WM., CARROLL JA., WANG Q, WHITMARSH J., SATOH K, PAKRASI HB., 2002. Proteomic analysis of a highly active photosystem II preparation from the cyanobacterium *Synechocystis* sp. PCC 6803 reveals the presence of novel polypeptides. *Biochemistry* 41: 8004–8012
8. ROOSE JL., PAKRASI HB., 2004. Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into photosystem II. *J Biochemistry* 279: 45417–45422
9. SHI T., FALKOWSKI PG., 2008. Genome evolution in cyanobacteria: the stable core and the variable shell. *Proc Natl Acad Sci USA*, 105:2510-5
10. HUCKA M., 2003 *The systems biology markup language (SBML): a medium for representation*.
11. BOURNE P. *A biological database be different from a biological journal*. Biol.. Vol. 1. n.º 3. pp. 179–81. DOI
12. *Modelo de datos*. Disponible en la web en: [http://es.wikipedia.org/wiki/Modelo\\_de\\_datos](http://es.wikipedia.org/wiki/Modelo_de_datos), Abril, 2010.
13. ALBERCA, A., GÁLVEZ, J., 2008. *Modelos avanzados de Bases de Datos. Bases de Datos Orientadas a Objeto y Bases de Datos Objeto-Relacionales*.
14. Communications of the ACM, volume=13, issue=6, pages=377-387
15. *Database Driven Webservices*, 2008. Disponible en: [www.WebServices/WebService\\_with\\_database\\_Mysql,\\_Webservices\\_database,\\_Develo/index.html](http://www.WebServices/WebService_with_database_Mysql,_Webservices_database,_Develo/index.html)
16. *Manual de KEGG API*, 2010. Disponible en: [http://www.genome.jp/kegg/docs/keggapi\\_manual.html](http://www.genome.jp/kegg/docs/keggapi_manual.html)
17. *Manuales de Java en línea*. Disponibles en: <http://www.desarrolloweb.com/manuales/57/>
18. Wikipedia. La Enciclopedia Libre. *PostgreSQL*. Disponible en: <http://es.wikipedia.org/wiki/PostgreSQL>
19. SBML Team. «The 5th Workshop on Software Platforms for Systems Biology».
20. YANG C., HUA Q., SHIMIZU K., 2002: Metabolic flux analysis in *Synechocystis* using isotope distribution from <sup>13</sup>C-labeled glucose. *Meta Eng* 2002, 4:202-16
21. GARCÍA, M. I., 2010. *Curso introductorio de XML*. Universidad de Granada, España. Disponible en: <http://geneura.ugr.es/~maribel/xml/introduccion/index.shtml>
22. *Tutorial de CSharp*. Wikipedia, Disponible en: <http://msdn.microsoft.com/es-svsharp/default.aspx>
23. *Tutorial de Visual Studio 2008*. Wikipedia, Disponible en: <http://msdn.microsoft.com/es-es/vstudio/default.aspx>.
24. *Tutorial de ASP.NET*. Wikipedia, Disponible en: <http://msdn.microsoft.com/es-es/vsharp/default.aspx>
25. SHANNON P., MARKIEL A, OZIER O., 2003. Cytoscape: un ambiente de software para los modelos integrados de las redes biomoleculares de la interacción”. *Genoma Research*. 13 (11): 2498-504. doi:10.1101/gr.1239303. PMID 14597658.
26. FUNAHASHI, A.; MATSUOKA, Y.; JOURAKU, A.; MOROHASHI, M.; KIKUCHI, N.; KITANO, H., 2008. “CellDesigner 3.5: una herramienta de modelado versátil para redes bioquímicas” *Desarrollo de la IEEE* Volumen 96, Número 8 , agosto 2008 Página (s): 1254 a 1265 [doi 10.1109/JPROC.2008.925458]
27. *COPASI: Biochemical Network Simulator. Tutorial*. Disponible en: [http://www.copasi.org/tiki-view\\_articles.php](http://www.copasi.org/tiki-view_articles.php)
28. LINDBERG P, PARK S., MELIS A., 2010. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using



- Synechocystis as the model organism. *Metabolic Engineering* 2010, 12:70–79.
29. NAVARRO E., MONTAGUD A., FERNÁNDEZ DE C. P., URCHUEGUÍA JF., 2009. Metabolic flux analysis of the hydrogen production potential in *Synechocystis* sp. PCC6803, *International Journal of Hydrogen Energy* 2009, 34:8828–8838.
30. MCHUGH K., 2005. *Hydrogen production methods*. Report MPR-WP-0001. Alexandria, Virginia: MPR Associates, Inc.; 2005.
31. TAMAGNINI P., AXELSSON R., LINDBERG P., OXELFET F., WUNSCHIERS R., LINDBLAD P., 2002. Hydrogenases and hydrogen metabolism of cyanobacteria. *Microbiology and Molecular Biology Reviews* 2002; 66(1):1–20.



