



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**BIOLOGÍA SINTÉTICA, ANÁLISIS DE UNA  
DISCIPLINA EMERGENTE**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:  
BIÓLOGO**

**P R E S E N T A:**

**FEDERICO CASTRO MONZÓN**



**DIRECTOR DE TESIS:  
Dr. PABLO PADILLA LONGORIA  
2011**



Universidad Nacional  
Autónoma de México



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

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

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

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

# Índice

## Introducción

### 1-Qué es la biología sintética.

- 1.1 Definición y consenso: caracteres que definen la biología sintética o que se asocian con ella, definiciones de diversos autores, los desacuerdos y el consenso de la comunidad sobre la disciplina\_\_\_\_\_ 5
- 1.2 Los grupos dentro de la biología sintética: las diferentes aproximaciones que existen hacia la biología sintética por su inclinación, intereses o tipo de problemas que abordan\_\_\_\_\_ 5

### 2-La historia de la biología sintética.

- 2.1 Antecedentes: las investigaciones que establecieron un terreno para la disciplina\_\_\_\_\_ 7
- 2.2 Trasfondo: las condiciones asociadas al desarrollo de la biología sintética\_\_\_\_\_ 9

### 3-Las herramientas del oficio.

- 3.1 Estandarización: las características del sistema estándar de Thomas Knight y las alternativas propuestas\_\_\_\_\_ 13
- 3.2 Abstracción: partes, dispositivos y sistemas \_\_\_\_\_ 16
- 3.3 Los sistemas complejos: una breve descripción de las herramientas matemáticas utilizadas para modelar sistemas biológicos en biología sintética\_\_\_\_\_ 17

### 4-Estudio de caso: la síntesis de un genoma mínimo.

- 4.1 Definición y descripción del genoma mínimo: los caracteres que definen al genoma mínimo\_\_\_\_\_ 21
- 4.2 Justificaciones para la búsqueda del genoma mínimo\_\_\_\_\_ 21
- 4.3 Aproximaciones al diseño y construcción del genoma mínimo\_\_\_\_\_ 22
- 4.4 Avances en la construcción del genoma mínimo\_\_\_\_\_ 22
- 4.5 Implicaciones de la construcción del genoma mínimo\_\_\_\_\_ 23
- 4.6 Perspectivas\_\_\_\_\_ 24

### 5-El estado del arte y perspectivas.

- 5.1 Los problemas que enfrenta la disciplina\_\_\_\_\_ 26
- 5.2 Los trabajos de mayor impacto y la dirección de la biología sintética\_\_\_\_\_ 26
- 5.3 El desarrollo de la disciplina en la actualidad\_\_\_\_\_ 27
- 5.4 La biología sintética en México\_\_\_\_\_ 28

## Introducción

El objetivo de este trabajo es exponer la situación actual de la biología sintética, analizar las últimas tendencias en la investigación y proveer al lector de las herramientas metodológicas necesarias para introducirse de lleno en la disciplina. Para ello se examina, a lo largo de varios capítulos, el contexto histórico en el que se ha desarrollado la biología sintética, la concepción que tienen sobre esta disciplina los grupos e individuos que han contribuido a su formación, los problemas abordados, la metodología y herramientas aplicadas y las tendencias actuales en las diversas líneas de investigación. Por último, pero no menos importante, se realizará un breve análisis del estado del arte en México y el mundo.

La biología sintética ha surgido y se ha desarrollado de forma sorprendente en la última década, con el desarrollo de técnicas que permiten manipular la información genética de una forma más fácil y precisa, así como un marco teórico que permite entender cómo interactúan las partes y predecir el comportamiento de un sistema biológico complejo.

Las metas de la biología sintética son ambiciosas: incluyen, entre otras, el diseño y construcción de sistemas biológicos que no estén constreñidos por su historia evolutiva, la producción de sustancias químicas cuya síntesis actualmente no es posible o rentable, la producción de tratamientos médicos “inteligentes” donde el agente terapéutico pueda analizar la información para tomar decisiones complejas y la construcción de materiales o tejidos para los cuales se requiere que las células se organicen o diferencien formando patrones espaciales complejos.

El desarrollo de la biología sintética resulta en una mayor comprensión de los mecanismos básicos e interacciones que permiten a los organismos realizar y regular sus funciones y aumenta el grado en el cual es posible manipularlos. Los programas desarrollados dentro del ámbito de esta disciplina tienen aplicaciones médicas, industriales y agrícolas a corto y mediano plazo. Dado que la biología sintética tiene grandes implicaciones sociales, ha recibido apoyos económicos sustanciales por parte de instituciones gubernamentales y compañías privadas. Su estudio, pues, es de trascendental importancia.

Como consecuencia de la reciente formación de la biología sintética y su explosivo desarrollo, introducirse en ella puede ser una tarea abrumadora. El volumen de información a manejar es muy grande, se encuentra dispersa en un sinnúmero de publicaciones y, muy a menudo, tiene un carácter técnico y multidisciplinario. En México y Latinoamérica esto puede ser aún más difícil para el estudiante y el investigador dado que hay pocas obras de referencia, acervos bibliográficos especializados y poca o ninguna labor de docencia relacionada a este campo. Los problemas anteriores hacen necesario y valioso un trabajo que condense y relacione la información disponible, exponiéndola de una forma concisa y clara.

Este trabajo está dirigido a todo aquél que esté interesado en adentrarse en la biología sintética. Se trató de estructurar de tal forma que no se requiera un conocimiento previo del campo, algo que, dada la situación actual de la disciplina en muchos países, resulta difícil sino es que imposible. Se

supone, sin embargo, que el lector posee conocimientos básicos de biología celular, de modo que parte de la información se trata de una forma breve, señalando las fuentes a las cuales se puede acudir para clarificar o profundizar en estos aspectos.

## Capítulo 1. Qué es la biología sintética

En este capítulo se presenta de una forma muy breve las diferentes concepciones que existen sobre la biología sintética, se expondrán los caracteres asociados a dicha disciplina, los diferentes grupos que existen dentro de ésta y se discutirán las posiciones que tienen varios autores con respecto a la existencia de la biología sintética.

### 1.1 Definición y consenso.

Se concibe, de una forma general, a la biología sintética como el diseño y construcción de sistemas biológicos [1]. No obstante, los varios grupos de investigación ponen diferente énfasis en un aspecto u otro de la disciplina y no han llegado a un consenso sobre su definición. De este modo, existen investigadores que entienden a la biología sintética como un sinónimo de ingeniería genética o como el diseño y construcción de sistemas biológicos que no existen de forma natural [2].

Entre los caracteres que se han propuesto para delimitar el campo de acción de la biología sintética, está la construcción de redes genéticas artificiales y el uso de nuevas técnicas de manipulación de la información genética [3]. Sin embargo, la naturaleza de las redes genéticas y de las técnicas utilizadas para construirlas no ha cambiado de forma sustancial hasta nuestros días. Por esta razón, algunos investigadores consideran que estas técnicas no constituyen un elemento que permita definir la biología sintética como una disciplina [3].

La forma en la que se abordan los problemas en la biología sintética, más que las herramientas que se utilizan o el tipo de dispositivos que se construyen, puede ser el elemento más importante para definir la disciplina [4].

Para lidiar con la complejidad asociada a la construcción de sistemas biológicos, se intenta diseñar componentes con una estructura estandarizada, una forma de ensamblaje jerárquico, un aislamiento funcional y un comportamiento caracterizado [5]. Este tipo de elementos permite la construcción de módulos, esto es, unidades funcionales independientes que contienen todos los elementos necesarios para realizar una función. Se intenta construir estas partes de tal forma que puedan asociarse con otros módulos para formar sistemas de mayor complejidad [6].

La reducción de un sistema a módulos permite manejar un menor número de elementos e interacciones, haciendo innecesario considerar los componentes de cada uno de ellos en el proceso de diseño de redes genéticas. La reducción a módulos se puede aplicar en todos los niveles jerárquicos de estructura biológica, desde moléculas individuales a células completas, tejidos y organismos. Esto hace del diseño de sistemas biológicos un proceso sistemático y racional [7].

### 1.2 Los grupos dentro de la biología sintética.

Existen, dentro de la biología sintética, distintos grupos que se caracterizan por su inclinación, intereses o tipo de problemas que abordan. Estos grupos reflejan, en muchos casos, una asociación con una de las muchas disciplinas que han contribuido al desarrollo de la biología sintética.

El tipo de sistemas que se suele estudiar están dentro del dominio de la biología, aunque ciertamente éste se superpone a nivel molecular con el de la química orgánica. Resulta natural que exista tal relación; la biología sintética se puede considerar, en muchas formas, un sucesor intelectual de la química orgánica. En la síntesis de DNA, la producción de proteínas *in vitro*, el diseño de vesículas y, en cierta medida, en el intento por generar una célula mínima, se refleja la influencia de trabajos anteriores asociados a la química orgánica [2]. La síntesis de estructuras o sistemas biológicos extienden las implicaciones que tuvieron los trabajos que se remontan a la síntesis de la urea [8].

Algunos investigadores, particularmente aquellos que tienen una formación relacionada a las ciencias biológicas, ven en la biología sintética una herramienta que, mediante la reproducción de un proceso biológico en un sistema sintético, permite determinar si se conocen las condiciones y elementos necesarios para que tal proceso tome lugar [6]. Ésta es una estrategia que prueba de forma directa el conocimiento que se tiene sobre un fenómeno y complementa la información que puede obtenerse mediante la disección del mismo en la naturaleza.

La biología sintética se sustenta sobre investigaciones previas desarrolladas en el ámbito de la biología molecular (*ver capítulo 2*). Gran parte de los pioneros tenían una formación asociada a la ingeniería eléctrica. No es extraño, si se considera esta herencia, el énfasis que tiene la disciplina en la aplicación de principios de ingeniería a la construcción de sistemas biológicos.

## **Bibliografía**

- 1) Andrew Balmer, Paul Martin. (2008). *Synthetic Biology Social and Ethical Challenges*. BBSRC.
- 2) Luisi, P. I. (2007). *Chemical aspects of synthetic biology*. Chemistry & Biodiversity, 4(4):603-621.
- 3) Breithaupt, H. (2006). *The engineer's approach to biology*. EMBO Rep, 7(1):21-3.
- 4) Kaznessis, Y. (2007). *Models for synthetic biology*. BMC Systems Biology, 1(1).
- 5) Knight, T. F. (2005). *Engineering novel life*. Molecular Systems Biology, 1(1):msb4100028-E1.
- 6) Endy, D. (2005). *Foundations for engineering biology*. Nature, 438(7067):449-53.
- 7) Synbiology (2005) *Synbiology, an analysis of synthetic biology research in Europe and North America. Output D3: literature and statistical review*. European Commission Framework Programme 6 reference contract 15357 (NEST).
- 8) Yeh, B. J. and Lim, W. A. (2007). *Synthetic biology: lessons from the history of synthetic organic chemistry*. Nat Chem Biol, 3(9):521-525.

## Capítulo 2. La historia de la biología sintética

En este capítulo se analizan las condiciones y eventos que se asocian al desarrollo de la biología sintética. Se presta especial atención al desarrollo de investigaciones o experimentos que normalmente se incluyen dentro del dominio de otras disciplinas, pero que sentaron las bases sobre las cuales después se apoyaría la biología sintética.

Una infinidad de trabajos se vinculan con el desarrollo de la biología sintética, muchos de los cuales tienen una relación indirecta o lejana con ella. Por eso sólo se expondrán, de forma breve, las investigaciones que en los últimos 60 años han permitido manipular la información genética de una forma más precisa. La influencia de estos trabajos en la biología sintética se presenta con ejemplos en los que se ha hecho uso de herramientas desarrolladas de forma previa o que se sustentan sobre principios que establecieron investigaciones anteriores.

Asimismo, se discuten algunos trabajos cuya relación con la biología sintética no es obvia, exponiendo los argumentos que ofrecen diversos autores y se analiza el impacto que tales investigaciones han tenido sobre la disciplina.

Se presentan también el escenario en el que surgió la biología sintética y las condiciones en las cuales se desarrolló esta disciplina. Esto incluye la formación de grupos especializados, el interés por parte de instituciones y empresas por desarrollar algunas técnicas o resolver determinados problemas y la percepción del público sobre la manipulación de la información genética, así como los beneficios y riesgos que conlleva esta actividad.

### 2.1 Antecedentes: Las investigaciones que establecieron un terreno para la disciplina.

De forma regular el desarrollo de la biología sintética se asocia a las investigaciones que, en los últimos 60 años, han permitido manipular el DNA de una forma más precisa. Antes de que se desarrollaran tales técnicas se podía manipular la información genética, si bien de una forma muy rudimentaria. Por medio de agentes físicos o químicos se podían inducir mutaciones y se podían seleccionar los organismos que manifestaban un carácter específico. Sin embargo, no se podía canalizar el efecto de los mutágenos hacia una secuencia específica o para obtener una mutación particular.

Un descubrimiento que abrió las puertas a la manipulación de la información genética fue el de la transformación bacteriana, realizado por Frederick Griffith [1]. Sus trabajos mostraron que los pneumococos podían recibir un factor transformante de otros pneumococos, produciendo un cambio heredable. Posteriormente se desarrolló un método que permitía inducir la transformación, se observó que este fenómeno no era exclusivo de pneumococos, se demostró que el factor transformante estaba en el DNA [2] y que esta molécula portaba la información genética.

Otro trabajo que sentaría la base de la ingeniería genética fue el descubrimiento de las enzimas de restricción tipo II [3]; un tipo de enzimas que reconocen y cortan una secuencia específica de nucleótidos. En la naturaleza, las enzimas de restricción, también conocidas como endonucleasas,



permiten a las bacterias defenderse de infecciones de bacteriófagos. En el laboratorio, proveyeron una herramienta muy útil para aislar fragmentos de DNA e incluirlos en un plásmido.

Dos investigadores, Herbert Boyer y Stanley N. Cohen, conjugaron el conocimiento que previamente se tenía de la transformación bacteriana y del uso de enzimas de restricción, para insertar una secuencia específica de DNA en un cromosoma bacteriano [4]. Este trabajo sigue marcando la forma general en la que se producen construcciones genética hasta el día de hoy. Véanse, por ejemplo, los diferentes sistemas estandarizados de construcción descritos en el tercer capítulo de esta tesis, el proceso de construcción de un oscilador biológico o la reciente construcción de un dispositivo para la producción de terpenoides en *Escherichia coli*.

Otras investigaciones que es importante mencionar son las que François Jacob, Jacques Monod [5] y Jean Pierre Changeux [6] realizaron sobre la forma en la que se regula la expresión génica. Este proceso no es trivial y la posibilidad de entender los mecanismos subyacentes ha permitido diseñar y construir los dispositivos que hoy caracterizan a la biología sintética. Más aun, estos trabajos ya hacían patente una forma de pensar en los procesos biológicos en términos mecánicos y Monod establece de forma explícita una analogía entre los mecanismos que regulan la la expresión génica y la forma en la que funciona una compuerta lógica [7].

Un trabajo que ha complementado la tecnología de DNA recombinante de Cohen y Boyer es la síntesis de genes, desarrollada por Har Gobind Khorana [8]. Este investigador comenzó sintetizando oligonucleótidos pequeños con un método que luego desarrollaría para generar cadenas más grandes y que hoy tiene una presencia trascendental en la biología sintética.

Hasta el advenimiento de la síntesis de DNA, los genes se tenían que extraer directamente de un organismo para introducirse con poca o ninguna modificación. Esto requería una estrategia única de ensamblaje y el gen introducido podía no ser funcional si el organismo receptor usaba un código de lectura diferente al del donador. Hoy en día, la síntesis de genes tiene una presencia ubicua en la disciplina, al grado que puede pensarse que es el elemento que define a la biología sintética, aunque, como se vio en el primer capítulo de esta tesis, esto dista de la realidad.

Es notable la influencia de los trabajos de Khorana en investigaciones como la síntesis del virus de la polio, la del cromosoma de *Mycoplasma genitalium* [9] o la refactorización del bacteriófago T7 [10], en las cuales se hace uso extensivo de la síntesis de genes. El uso de esta técnica es también esencial para los varios sistemas de construcción estandarizados que, como se detalla en el siguiente capítulo, requieren de la modificación de un gen para que pueda cortarse con enzimas de restricción específicas.

Como resultado de estas investigaciones se han desarrollado y refinado diferentes técnicas que han permitido aislar, secuenciar, replicar y modificar diferentes genes. Es necesario mencionar que en este capítulo sólo se han mencionado algunos de los trabajos de mayor relevancia aunque existen una infinidad de investigaciones que dieron lugar a técnicas como la electroforesis y el PCR, entre muchas otras, que se usan de forma rutinaria en la biología sintética.

### **Línea de tiempo.**

- 1823 Friedrich Wöhler sintetiza la primera molécula orgánica: la urea.
- 1866 Gregor Mendel publica “*Versuche über Pflanzen-Hybriden*”, estableciendo los principios básicos de la herencia.
- 1876 Francis Galton ofrece una aproximación estadística para comprender la herencia genética.
- 1927 Hermann J. Müller demuestra que los rayos X inducen mutaciones.
- 1928 Frederick Griffith descubre la transformación.
- 1943 Oswald Avery, Colin MacLeod y Maclyn McCarty identifican al DNA como el factor transformante.
- 1952 Alfred Hershey y Martha Chase determinan que el DNA es la molécula hereditaria.
- 1953 James D. Watson y Francis Crick proponen un modelo de la estructura del DNA.
- 1959 El experimento de Pardee, Jacob y Monod “*PaJaMo*” revela la existencia del mRNA.
- 1961 Marshall Nirenberg, Heinrich Matthaei y Severo Ochoa descifran el código genético.
- 1961 François Jacob y Jaques Monod proponen un modelo que explica la regulación génica.
- 1970 Har Gobind Khorana sintetiza el primer gen.
- 1970 Hamilton O. Smith descubre la primera enzima de restricción.
- 1972 Paul Berg produce la primera molécula de DNA recombinante.
- 1973 Herbert Boyer y Stanley N. Cohen desarrollan la tecnología del DNA recombinante.
- 1975 Allan Maxam, Walter Gilbert y Frederick Sanger desarrollan diferentes métodos para secuenciar el DNA.
- 1975 Se lleva a cabo la conferencia de Asilomar.\*
- 1978 Se produce la primera molécula de insulina humana en una colonia de *E. coli* modificada como una aplicación de la investigación del científico mexicano Francisco Bolívar Zapata.
- 1981 Gordon y Ruddle producen el primer ratón transgénico.
- 1982 El genoma del fago lambda es secuenciado.
- 1986 Comienza el proyecto para secuenciar el genoma humano
- 1995 El genoma de *Mycoplasma genitalium* es secuenciado.
- 1996 El genoma de *Saccharomyces cerevisiae* es secuenciado.
- 1997 El genoma de *Escherichia coli* es secuenciado.
- 1998 El genoma de *Caenorhabditis elegans* es secuenciado.
- 2001 El genoma humano es secuenciado.
- 2003 Primer congreso de biología sintética.
- 2006 Síntesis del virus de la polio.
- 2006 Primer iGEM.
- 2008 Se sintetiza el genoma de *Mycoplasma genitalium*.

### **2.2 Las condiciones asociadas al desarrollo de la biología sintética.**

La biología sintética nació en un momento en que se conjugaban la disponibilidad de información y la capacidad de analizarla y modificarla, así como un interés por parte de instituciones privadas y gubernamentales por las potenciales aplicaciones médicas, agrícolas e industriales que podrían tener algunas líneas de investigación.

---

\* La conferencia de Asilomar fue una conferencia organizada por Paul Berg en 1975. Esta conferencia tenía como fin discutir los peligros que podría conllevar la manipulación del material genético y regular la ingeniería genética.

El patrón de desarrollo de la biología sintética es, según Brian J Yeh y Wendell A Lim [11] similar al que han experimentado otras disciplinas en las que la capacidad para modificar y diseñar un sistema es necesariamente precedida por la acumulación de datos experimentales y el desarrollo de modelos que expliquen tales datos.

El argumento de Yeh y Lim parece estar respaldado por el enorme volumen de información que revelaron técnicas como la secuenciación automatizada de genes y los microarreglos de DNA. Es igualmente importante señalar el desarrollo de modelos que permitieron entender las interacciones que existen entre un número grande de genes o elementos celulares. Estos modelos dieron lugar a la biología de sistemas, una disciplina que como varios investigadores señalan [12][13], ha contribuido de manera relevante al desarrollo de la biología sintética.

Otro factor que varios autores postulan que detonó el desarrollo de la biología sintética [14] es la reducción en los costos de la síntesis de DNA. El costo ha descendido de una forma sustancial en los últimos diez años, haciendo común el uso de una tecnología cuyo uso era prohibitivo para la mayor parte de las instituciones. Estas técnicas, que en un principio se utilizaban para sintetizar pequeños oligonucleótidos, resultaron particularmente útiles para generar librerías de partes estandarizadas.

Se puede argumentar que los factores mencionados anteriormente existían ya antes del desarrollo de la biología sintética. No obstante, la cantidad de información disponible, aunada a la complejidad de los modelos que se han desarrollado y la capacidad que tenemos para modificar el DNA, no tienen precedentes.

Los peligros y beneficios que la sociedad percibe que podrían acarrear el desarrollo de algunas líneas de investigación, afectaron fuertemente el ambiente en el que se desarrolló la biología sintética. Algunas investigaciones que abordaban problemas con aplicaciones inmediatas, aseguraron un respaldo sólido y han acaparado la atención de los investigadores. De una forma particular, se observa el notable desarrollo que tuvieron los trabajos dirigidos a resolver la crisis energética en los primeros años de la disciplina.

Desde su aparición, la biología sintética lidió con el mismo estigma que ha afectado a la ingeniería genética: el temor de que la tecnología sea usada inapropiadamente o que tenga efectos perjudiciales o imprevistos. Entre las respuestas a estas preocupaciones está la declaración de bioseguridad emitida en la Segunda Conferencia Internacional de Biología Sintética [15] y la propuesta de la Goldman School of Public Policy de Berkeley [16]. Estas publicaciones buscaron establecer normas de seguridad que regulen el trabajo dentro de la disciplina y alentar a las empresas que sintetizan DNA a registrar y en su caso rechazar, los pedidos de genes que pudieran resultar peligrosos.

Algunas revistas regularon rápidamente la publicación de determinados artículos asociados a la biología sintética. Tal es el caso de Nature y Proceedings of the National Academy of Sciences [17] que, ante la controversia que generó la reconstrucción del virus del polio [18] y otros artículos que manejaban información que consideraban potencialmente peligrosa, emitieron de forma conjunta un

comunicado que establecía nuevos lineamientos que tendrían que cumplir los artículos para ser publicados. Estos lineamientos consideraban, entre otras cosas, la posibilidad de excluir o modificar trabajos que a juicio del editor pudiesen ser utilizados con fines perjudiciales.

Asimismo, la regulación por parte de instituciones gubernamentales era notoria, aunque el rápido desarrollo de la disciplina parecía dejar atrás a la legislación. Sin embargo, algunas naciones ya habían comenzado a regular la investigación dentro de la biología sintética. Así, el gobierno británico estableció en el año 2003 barreras legales que restringían el acceso por parte de investigadores a cualquier clase de información que pudiese utilizarse para producir un arma biológica. De forma adicional, los fondos que las instituciones gubernamentales destinan a la investigación también se vieron afectados por la percepción que tenía la sociedad de la biología sintética en el momento [19].

Por otro lado, como se verá en el último capítulo, la difusión que recibió la biología sintética por parte de congresos como el iGEM y el SB, permitieron que esta disciplina se expandiese rápidamente; al grado en que hoy en día, pocos años después de su surgimiento, ya existen cursos en varias universidades destinados específicamente a enseñar biología sintética. La labor de investigadores comienza a ser reconocida de forma oficial y ya existen laboratorios dedicados a la investigación en esta área. De esta forma, la disciplina se encuentra en un momento de crecimiento, formulando sus bases y formando y profesionalizando un cuerpo académico.

## **Bibliografía**

- 1) Griffith, F. (1928). *The significance of pneumococcal types*. Journal of Hygiene **27**, 113–159
- 2) Avery, O. T., MacLeod, C. M., and McCarty, M. (1979). *Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type iii*. J Exp Med, 149(2):297-326.
- 3) Smith, H. O. and Wilcox, K. W. (1970). *A restriction enzyme from hemophilus influenzae. i. purification and general properties*. Biotechnology, 24:38-50.
- 4) Cohen, S. N., Chang, A. C., Boyer, H. W., and Helling, R. B. (1973). Construction of biologically functional bacterial plasmids in vitro. PNAS, 70(11):3240-3244.
- 5) Jaccob, F. and Monod, J. (1961). *Genetic regulatory mechanisms in the synthesis of proteins*. J Mol Biol, 3:318-356.
- 6) Monod, J., Wyman, J., and Changeux, J. P. (1965). *On the nature of allosteric transitions: A plausible model*. J Mol Biol, 12:88-118.
- 7) Monod, J. (1970). *Chance and Necessity*. Penguin.

- 8) Khorana, H. et al. (1971). Studies on polynucleotides: total synthesis of the structural gene for an alanine transfer ribonucleic acid from yeast. *J.Mol.Biol.* 72, 209-217
- 9) Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2008). *Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome.* *Science*, 319(5867):1215-1220.
- 10) Chan, L. Y., Kosuri, S., and Endy, D. (2005). *Refactoring bacteriophage t7.* *Molecular Systems Biology*, 1(1):msb4100025-E1-msb4100025-E10.
- 11) Yeh, B.J. and Lim, W.A. (2007). *Synthetic biology: lessons from the history of synthetic organic chemistry.* *Nat Chem Biol*, 3(9):521-525.
- 12) Andrianantoandro, E., Basu, S., Karig, D. K., and Weiss, R. (2006). *Synthetic biology: new engineering rules for an emerging discipline.* *Mol Syst Biol*, 2:msb4100073-E1-msb4100073-E14.
- 13) Serrano, L. (2007). *Synthetic biology: promises and challenges.* *Mol Syst Biol*, 3.
- 14) ETC Group. (2007.) *Extreme genetic engineering An Introduction to Synthetic Biology.* ETC Group. Online publication at [www.etcgroup.org/en/node/602](http://www.etcgroup.org/en/node/602).
- 15) Conferees, SB2.0 (2006). *Public Draft of the Declaration of the Second International Meeting on Synthetic Biology.*
- 16) Stephen M. Maurer, Keith V. Lucas & Starr Terrell. (2006). *From Understanding to Action: Community-Based Options for Improving Safety and Security in Synthetic Biology.* Goldman School of Public Policy University of California at Berkeley.
- 17) Cozzarelli, N. R. (2003). *Policy on publication of sensitive material in the life sciences.* *PNAS*, 100(4).
- 18) Cello, J., Paul, A. V., and Wimmer, E. (2002). *Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template.* *Science*, 297(5583):1016-1018.
- 19) Petsko, G. A. (2002). *An asilomar moment.* *Genome biology*, 3(10)

## **Capítulo 3**

### **Las herramientas del oficio**

Este capítulo se enfoca en analizar la forma en la que las técnicas y herramientas utilizadas de forma rutinaria en la biología sintética han permitido diseñar, modelar y construir nuevos dispositivos biológicos. Esto no es de ninguna forma un manual y algunas técnicas que son comúnmente usadas en otras disciplinas, como la transformación y la electroforesis, sólo se mencionan de una forma breve.

Se presentan las herramientas que se han desarrollado exclusivamente para la biología sintética, se analizan los diferentes sistemas estandarizados de construcción y la caracterización que se está realizando de partes y dispositivos biológicos.

Por último, se describen algunas herramientas metodológicas que permiten entender, diseñar y modelar dispositivos biológicos.

#### **3.1 Estandarización**

Antes del desarrollo de la biología sintética, los dispositivos biológicos se diseñaban de una forma artesanal, con un método único para cada parte según la secuencia original del gen de interés, lo que resultaba costoso y requería mucho trabajo. De forma adicional, las partes utilizadas en la construcción a menudo contenían varios sitios de restricción que dificultaban aislar la secuencia de interés a menos de que se incluyesen mutaciones en sitios específicos.

Los sistemas empleados actualmente en la biología sintética, permiten construir dispositivos biológicos de una forma modular, al estandarizar las uniones entre los diferentes componentes. Estos sistemas hacen uso de una combinación específica de enzimas de restricción que permite incluir genes de una forma iterativa y aprovechan la flexibilidad que da la síntesis de genes para eliminar y añadir sitios de restricción.

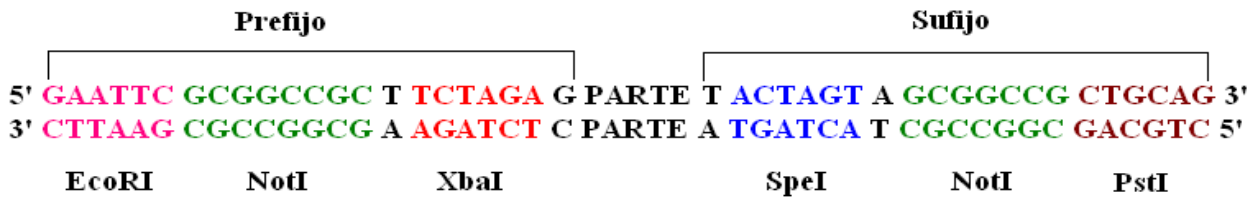
Es necesario considerar que aunque existe un sistema estandarizado de construcción que es usado en casi la totalidad de los dispositivos, la biología sintética se encuentra en un periodo de transición y no es claro todavía si se impondrá un sistema.

#### **El RFC10**

Entre los sistemas estandarizados, el RFC10 [1] (por la nomenclatura utilizada por la Biobrick Foundation, también conocido como “el estándar de Knight”), es por mucho el más utilizado. Es también uno de los primeros sistemas estandarizados de construcción, siendo solamente precedido por el sistema NOMAD [2], que no tiene un uso extendido y por prototipos del RFC10 como el RFC9.

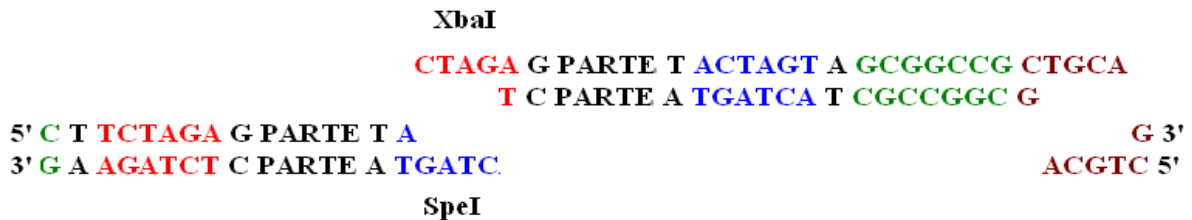
En el RFC10 las partes están flanqueadas por sitios de restricción que forman un prefijo y un sufijo. En el extremo 5' se encuentran los sitios de restricción de las enzimas EcoRI y XbaI, mientras que en el extremos 3' se encuentran los de SpeI y PstI. De forma adicional se han añadido bases entre

los sitios de restricción (*ver Fig.1*) para inhibir la metilación accidental, que afectaría la actividad de las endonucleasas.



**Fig 1.** La ubicación de los sitios de restricción dentro de un prefijo y un sufijo en una parte que utiliza el RFC10.

El diseño del sistema de Knight permite unir una secuencia de DNA en frente o detrás de otra, de manera independiente del contenido de las partes, manteniendo asimismo los sitios de restricción (*ver fig 2*). De esta forma, la construcción no es seriada y pueden realizarse varios ensamblajes simultáneamente.



**Fig 2.** El esquema muestra que los sitios de restricción XbaI y SpeI son complementarios. En este ejemplo, una parte que se libera de su vector mediante el uso de XbaI y PstI, se coloca en frente de otra cuyo plásmido ha sido abierto con SpeI y PstI. Después de la ligación, todos los sitios de restricción del RFC10 se habrán mantenido.

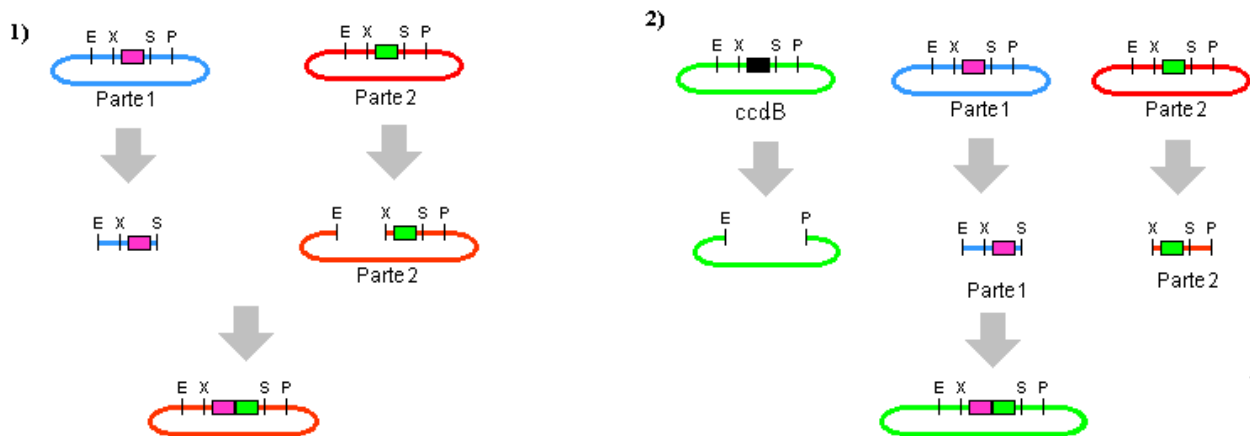
El ensamblaje en el RFC10 implica cortar los plásmidos de interés con enzimas de restricción para liberar una parte, preparar un vector, ligar esta parte con el vector, transformarlo en una cepa competente y corroborar que el proceso se haya realizado por medio de una electroforesis. Estos procesos se realizan de forma rutinaria en otras disciplinas, han sido descritos de forma extensa (*ver biografía recomendada*) y los materiales, dispositivos y reactivos necesarios para llevarlos a cabo pueden conseguirse fácilmente.

El proceso repetitivo que implica recuperar plásmidos, cortarlos, ligarlos y transformarlos puede ser automatizado. El diseño del estándar de Knight tomó en cuenta esta posibilidad, de forma que las enzimas de restricción utilizadas funcionan en el mismo buffer y a la misma temperatura, y su actividad puede inhibirse mediante un choque térmico (*ver fig3*). Sin embargo, la electroforesis requiere analizar e interpretar datos en una manera que es todavía difícil de automatizar, por lo cual no puede corroborarse la ligación y transformación.

Con el fin de automatizar el proceso de ensamblaje utilizando el RFC10, se han desarrollado plásmidos que mediante el uso de ccdB, un gen que codifica a una toxina, genera una selección negativa. Esto complementa la selección positiva de los genes de resistencia a antibióticos que se usa de forma común en el ensamblaje estándar (*ver fig4*), eliminando la posibilidad de que crezcan colonias que no hayan adquirido la construcción. Así, se hace innecesaria la electroforesis y puede automatizarse el ensamblaje con sólo unas pequeñas modificaciones (*ver fig 3*), en un proceso

conocido como el ensamblaje de tres antibióticos.

Aunque ya existen métodos para automatizar la construcción de dispositivos biológicos, su uso no se ha extendido y pocos laboratorios tienen el equipo necesario para realizar tal tarea. Sin embargo, es de esperarse que, a medida que las construcciones se vuelvan más complejas (*ver capítulo 4*) sea necesario el uso de equipos que puedan realizar los ensamblajes de forma eficiente.



**Fig 3.** En el esquema se compara:

1) El ensamblaje estándar.

2) El ensamblaje de 3 antibióticos.

A pesar de sus ventajas, el RFC10 genera un desplazamiento en el marco de lectura y un codón de terminación “TAG” en la cicatriz de marca, lo que inhabilita la adición de marcadores a una proteína o la fusión de proteínas. Silver [3] ya ha expuesto con detalle las deficiencias inherentes al sistema estándar de Knight. Parece ser el consenso de la comunidad de investigadores que trabajan con este sistema que es necesario un método que permita superar tales problemas. Sin embargo, implementarlo implicaría, en muchos casos, generar una nueva librería de partes. Esto representaría un costo muy alto si se tiene en cuenta que al día de hoy existen alrededor de 3200 partes. Por ello, es probable que el sucesor del RFC10 sea compatible con éste.

Por fortuna, desde que se implementó el estándar de Knight se ha reservado el uso de algunas enzimas de restricción que se han eliminado en todas las partes, reservando estos sitios para futuros usos. Estas enzimas son PvuII, XhoI, AvrII, NheI y SapI y son algunas de las que se están utilizando para diseñar una nueva generación de sistemas estandarizados.

Aparte del RFC10, vale la pena mencionar el RFC 12 [3], también conocido como Biobrick2 o BB2, que es un sistema que se ha diseñado de forma reciente y usa una dinámica similar a la del RFC10 para ensamblar componentes. Sin embargo, a diferencia del estándar de Knight, este sistema tiene los sitios de restricción en diferentes posiciones y sustituye XbaI por NheI.

El RFC12 permite unir dos partes creando una cicatriz entre el sitio SpeI y el de NheI que no genera un cambio en el marco de lectura ni un codón de término. Esta secuencia se traduce en el código universal como Ala-Ser, unos aminoácidos hidrofóbicos que resultan apropiados para unir dominios proteicos.



El RFC12 tiene amplia difusión, siendo segundo en popularidad ante el RFC10 y tiene la ventaja de ser compatible con este último, aunque es posible que enfrente competencia por uno de los varios estándares que se están desarrollando actualmente.

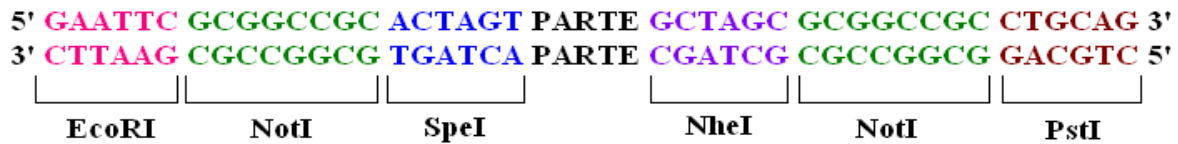


Fig 4. La ubicación de los sitios de restricción dentro de un prefijo y un sufijo en una parte que utiliza el RFC12.

### 3.2 Abstracción.

Muy a menudo el análisis de los sistemas biológicos resulta extremadamente complejo por su elevado número de elementos e interacciones. Mediante un ejercicio de abstracción es posible excluir propiedades que no son relevantes y manejar sólo una parte del sistema para entender su dinámica.

En biología sintética los sistemas se descomponen, para su análisis y diseño, en módulos, esto es, unidades funcionales independientes que contienen todos los elementos necesarios para realizar una tarea [4]. El uso de módulos permite analizar las distintas partes de manera independiente.

Así, un sistema puede ensamblarse a partir de varios dispositivos cuyo funcionamiento es conocido. Un ejemplo puede verse en el diseño de un oscilador biológico, como el represilador de Elowitz [5]. Este oscilador está compuesto por tres dispositivos que funcionan como inversores. Estos, a su vez, están conformados por diferentes partes como promotores o secuencias codificantes, que tienen una función por sí mismas y que no pueden subdividirse sin la pérdida de dicha función.

De una forma similar a la que se utiliza para describir el funcionamiento de componentes electrónicos, la actividad de diversos dispositivos biológicos se ha

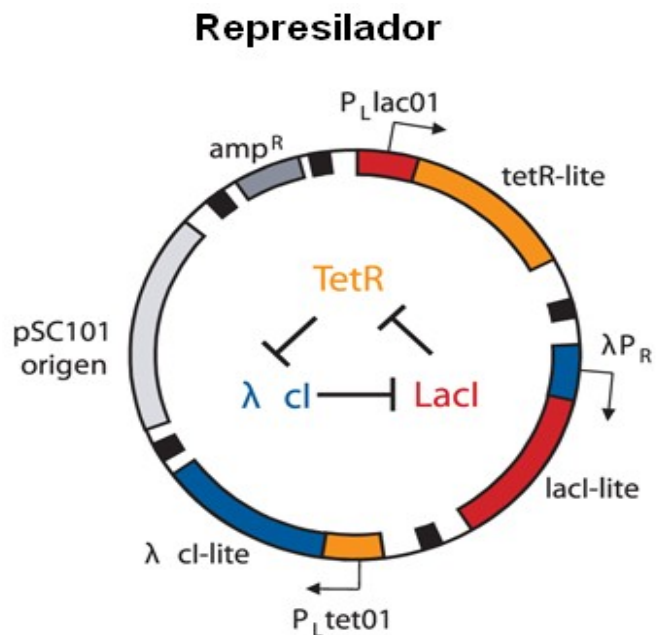


Fig 5. Una representación del represilador de Elowitz, un oscilador biológico constituido por tres inversores.

descrito por medio de lógica booleana. La forma en que se regula la expresión génica se presta a este ejercicio. La expresión de un operón es regulada por la presencia de proteínas y otras moléculas, cuyos efectos pueden registrarse en una tabla de verdad, ayudando de esta forma a

determinar el tipo de función que describe el comportamiento de un dispositivo biológico.

Glucosa	Lactosa	Expresión
0	0	0
1	0	0
1	1	0
0	1	1

**Fig 6.** Se presenta la tabla de verdad que describe el funcionamiento del operón de la lactosa. Esta tabla describe la función booleana  $\overline{AB}$

Se han diseñado gran número de dispositivos biológicos cuyo comportamiento se ajusta al de diferentes compuertas lógicas. Por ejemplo, un operón simple cuyo promotor está regulado solamente por la presencia de una proteína que actúa como represor describe la función “NOT”, mientras que otro operón puede funcionar como una compuerta lógica de tipo “AND” si su operador es inducible y el activador requiere de la presencia de un autoinductor para su correcta función.

Utilizando varias compuertas lógicas, así como fusionando dominios de diferentes promotores, se han logrado producir dispositivos cuyo funcionamiento es descrito por compuertas lógicas más complejas, como “XOR” o “NAND”.

### 3.3 Los modelos complejos.

Las abstracciones mencionadas anteriormente son increíblemente útiles para comprender la dinámica de un sistema biológico, manejando un número reducido de elementos. Sin embargo, los diferentes módulos no se encuentran separados por una barrera física, los componentes celulares tienen una multitud de interacciones y el comportamiento de los operones no se ajusta perfectamente al de una compuerta lógica [6].

Es de esperarse que el comportamiento de un sistema biológico diste de aquél descrito por medio de un análisis reduccionista [6]. Sin embargo, existen técnicas que, como se mencionó en el primer capítulo de esta tesis, complementan las herramientas que se utilizan en la biología sintética y permiten realizar modelos de sistemas complejos.

Modelar un sistema complejo es, en la mayor parte de los casos, una actividad interdisciplinaria. Se requiere un conocimiento detallado del sistema biológico a estudiar, herramientas matemáticas que permitan representar las relaciones que existen entre los elementos de un sistema y en algunos casos algunas habilidades en el manejo de computadoras.

Gran cantidad de información sobre los sistemas biológicos puede obtenerse gracias a técnicas como la secuenciación automatizada y los microarreglos de DNA. Dicha información se encuentra distribuida en una gran cantidad de bases de datos como el KEGG (Kyoto Encyclopedia of Genes and Genomes), en el Reactoma o en BioCyc y de forma importante NCBI.

Hasta hace relativamente poco, la capacidad de cómputo necesaria para analizar la gran cantidad de datos obtenida de sistemas biológicos no estaba disponible. Gracias al rápido desarrollo de los procesadores ha sido posible determinar qué genes se expresan de forma simultánea o consecutiva y, de una forma indirecta, determinar la estructura de las redes genéticas.

La forma en la que se analizan los datos puede revelar diferentes aspectos de la dinámica de un sistema. De esta manera un modelo continuo y determinista puede representar de forma fiel el comportamiento de una red genética con gran cantidad de elementos, en donde los eventos probabilísticos tienen un efecto despreciable. Por otro lado, un análisis estocástico y discreto permite analizar qué tan robusto es un sistema en condiciones limitadas, donde el ruido tiene un efecto grande. Existen una infinidad de modelos que tienen utilidad en situaciones particulares, en donde se busca representar de forma estática el funcionamiento de un dispositivo y representar su evolución en una serie de pasos discretos. En última instancia, la elección del modelo depende de la interrogante que se quiere resolver y el tipo de datos que se poseen. Estos métodos se complementan y proveen información que un solo enfoque no podría revelar.

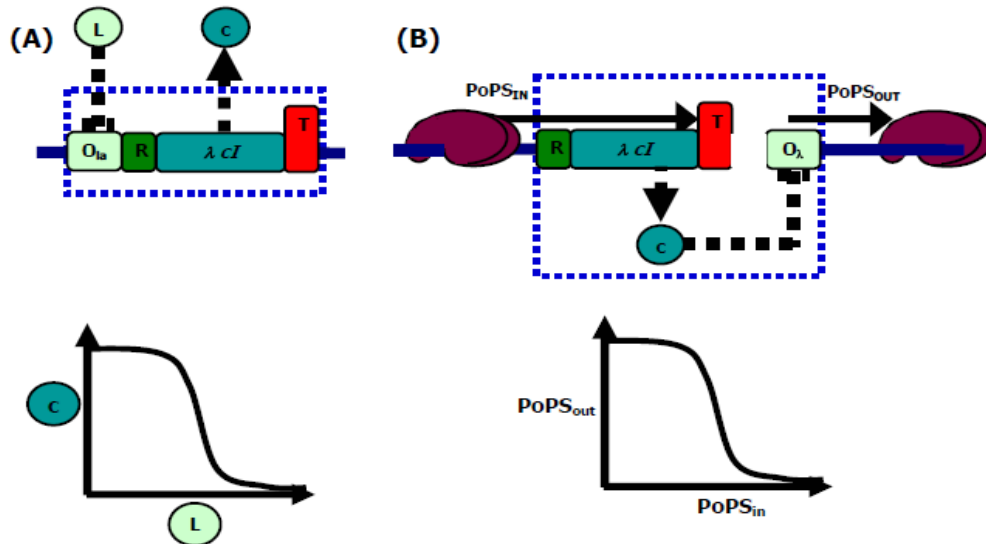
Afortunadamente, la mayor parte de las bases de datos y los programas que se utilizan para simular sistemas biológicos utilizan uno de dos formatos para codificar la información: XML o SBML. Esto hace compatible la información de las distintas bases de datos y permite reutilizar modelos, construirlos de una forma modular y promueve la cooperación de grupos de trabajo que utilizan los diferentes programas.

En muchos casos, el análisis de un sistema complejo ha revelado comportamientos o patrones insospechados. Uno de estos casos se puede observar en el análisis de las redes genéticas, que ha mostrado la existencia de patrones recurrentes en su estructura. Estos patrones, conocidos como *motivos de red* (del inglés Network Motif), tienen propiedades que emergen de la interacción de los elementos que componen la red [7]. Así, algunos de ellos pueden, entre otras cosas, reducir o amplificar el ruido en un sistema, afectar su tiempo de respuesta, llevar al sistema a expresar un comportamiento complejo (e.g. una respuesta bimodal a la expresión de una proteína en una población celular) e incluso permitirle guardar memoria de un suceso.

Ya se han desarrollado varias herramientas y bases de datos que están diseñadas específicamente para la biología sintética. Cabe mencionar, entre otros, la base de datos “Registry of Standard Biological Parts” (conocido coloquialmente como “Registry”), que mantiene una librería con partes funcionales y redes genéticas que provienen de módulos reusables; y el programa “Synthetic Biology Software Suite” que está diseñado para emplear el Registry of Standard Biological Parts y exporta la información de un modelo en formato SBML (Systems Biology Markup Language) que otros programas pueden utilizar.

Como un último inciso de esta sección, es necesario enfatizar el enorme esfuerzo que se está llevando a cabo por caracterizar las distintas partes que componen el Registry [8]. Esto no es una tarea fácil; de forma natural, los diferentes sistemas de regulación genética utilizan diversas señales para inhibir o promover la transcripción. Así, puede parecer que no existe un parámetro común que permita caracterizar la actividad de los dispositivos genéticos (*ver fig. 7*). Sin embargo, la actividad de diversos promotores y dispositivos ha podido determinarse al asociar otras partes cuyas características ya son conocidas o que permiten registrar la magnitud de la transcripción, como resultado de la expresión de un marcador, tal como una proteína fluorescente. De esta forma, se establece una unidad de medida (las polimerasas por segundo o PoPS) que permite conocer las

propiedades de una parte y a la vez establecer una señal común entre dispositivos.



**Fig 7.** Se presenta en este esquema (A) el modelo tradicional de un inversor biológico que utiliza, para su regulación, señales específicas y (B) el modelo de un inversor basado en PoPS [9].

Aunque la tasa de transcripción es uno de los parámetros que más peso tienen en el diseño de dispositivos biológicos, existen algunas características que también se han estudiado por la forma en la que afectan el desempeño final de un sistema. Entre estos caracteres, cabe resaltar la carga metabólica y evolutiva que impone un dispositivo sobre una célula y la estabilidad genética de tal dispositivo.

## Bibliografía

- 1) Knight T.F. Jr. (2003). *Idempotent Vector Design for Standard Assembly of Biobricks*. Dspace <http://hdl.handle.net/1721.1/21168>
- 2) Rebatchouk, D., Daraselia, N., and Narita, J.O. (1996). *Nomad: a versatile strategy for in vitro DNA manipulation applied to promoter analysis and vector design*. PNAS, 93(20):10891-10896.
- 3) Phillips I, Silver P. (2006) BBF RFC 23: *A new biobrick assembly strategy designed for facile protein engineering*. Dspace <http://dspace.mit.edu/handle/1721.1/32535>
- 4) Endy, D. (2005). *Foundations for engineering biology*. Nature, 438(7067):449-453.
- 5) Elowitz, M. B. and Leibler, S. (2000). *A synthetic oscillatory network of transcriptional regulators*. Nature, 403(6767):335-338.
- 6) Kaznessis, Y. (2007). *Models for synthetic biology*. BMC Systems Biology, 1(1):47.
- 7) Alon, U. (2007). *Network motifs: theory and experimental approaches*. Nature Reviews Genetics,

8(6):450-461.

8) Kelly J.R., Rubin A.J., Davis J.H., Ajo-Franklin C.M., Cumbers J, Czar M.J., de Mora K, Gliberman AL, Monie D.D., and Endy D. (2009) *Measuring the activity of Biobrick promoters using an in vivo reference standard*. J Biol Eng 20; 3 4.

9) Kelly J.R. (2005) *Design and evolution of engineered biological systems*. MIT Thesis proposal ID 1998728.

## **Capítulo 4. Estudio de caso: Construcción de un genoma mínimo**

En este capítulo se hace un estudio de caso, examinando detalladamente la forma en la que varios autores proponen la construcción de un genoma mínimo. Se define el concepto de genoma mínimo, se argumenta su asociación con la biología sintética, se presentan las justificaciones para su desarrollo, así como el impacto que podría tener este proyecto. Asimismo, se describe cómo se ha diseñado y los avances que existen relacionados con su construcción.

### **4.1 Definición y descripción del genoma mínimo**

Como consecuencia del desarrollo de la biología sintética, las células se conciben cada vez más como un conjunto de partes que pueden ensamblarse para producir un organismo. La síntesis de genes ha avanzado al grado en el que no sólo es posible sintetizar un virus, formado por una decena de genes, sino cromosomas bacterianos completos. No obstante, esto sigue siendo una labor titánica.

Los cromosomas bacterianos son complejos e incluyen una gran cantidad de genes que son innecesarios y en ocasiones no deseados para el cultivo de un organismo. Por ello se busca diseñar y construir un genoma mínimo, esto es, el conjunto más pequeño de genes requeridos para mantener una célula en condiciones particulares.

Es importante notar que el genoma mínimo puede variar dependiendo del concepto de vida o del ambiente en el que se cultiva el organismo. De esta forma, no es una estructura única sino un término que define una serie de estructuras [1].

Las características propuestas del genoma mínimo también varían según el autor. Así, puede contener entre 50 y 380 genes, diferentes tipos de membranas celulares y maquinaria de replicación o transcripción. No obstante las diferencias entre los proyectos, todos han buscado simplificar el sistema e identificar los módulos necesarios para su funcionamiento [2]; algo que permite asociar estos trabajos a la biología sintética.

### **4.2 Justificaciones para la búsqueda del genoma mínimo**

La búsqueda de un sistema biológico fácil de comprender es una de las justificaciones más frecuentes para el desarrollo de un genoma mínimo, pero no es por mucho la única. La construcción de un genoma mínimo podría revelar información sobre el origen de la vida, la evolución de la maquinaria celular y la estructura de las redes genéticas [3]. Asimismo, se ha argumentado que las células que mantienen un genoma mínimo, las células mínimas, tienen una aplicación inmediata como chasis general de alta eficiencia para una serie de dispositivos [2].

Se ha propuesto que el desarrollo de una célula mínima, al tener una mayor eficiencia que las células convencionales, permitiría, de una forma indirecta, optimizar la degradación de tóxicos, la producción de polímeros o medicamentos útiles y la de combustibles que permitiesen enfrentar la actual crisis energética [3].

Las posibles aplicaciones comerciales han dado lugar a que algunos investigadores intenten patentar un genoma mínimo [4] y no publicar los detalles de sus investigaciones a la comunidad científica.

### **4.3 Aproximaciones al diseño y construcción del genoma mínimo**

Es necesario, para el diseño y construcción de un genoma mínimo, identificar los genes esenciales para mantener un metabolismo básico y permitir la reproducción de una célula mínima. Esto se ha buscado por medio de la comparación de múltiples genomas [5] y la inactivación sistemática de genes en un organismo [6].

La comparación de genomas permite identificar genes homólogos. Se supone que estos genes, si se encuentran en organismos filogenéticamente distantes, son esenciales para mantener las funciones celulares básicas [5]. Sin embargo, en la historia evolutiva de los organismos algunos genes pueden ser sustituidos por otros que no guardan homología, de forma que genes esenciales parecen prescindibles. Por otro lado, si los organismos comparados son cercanos, éstos pueden tener muchos genes en común. Así, el análisis da la impresión de que existe una gran cantidad de genes esenciales [7].

Una aproximación complementaria ha sido la inactivación de genes. Mediante el uso de transposones es posible interrumpir la transcripción de un gen, de forma que no se exprese o lo haga de forma incorrecta. No obstante, esta metodología también es susceptible a errores, dado que muchos genes se organizan de forma policistrónica. Así, parte del transcrito no es afectado por la inserción del transposón. De forma adicional, varios genes pueden parecer prescindibles pero su efecto acumulado puede ser letal [2].

Mushegian y Koonin estimaron que se requerían 256 genes para mantener una célula mínima en condiciones permisibles, a partir de una comparación entre los genomas de *Mycoplasma genitalium* y *Haemophilus influenzae* [5]. Este número se reduce a 206 genes según estimados de Andrés Moya y su grupo, que trabajan con *Buchnera* [7], un endosimbionte obligado.

Mediante el uso de transposones, J.C. Venter y su grupo han determinado que sólo 382 de los 482 genes de *Mycoplasma genitalium* son esenciales [6]. Este número se puede reducir a 150 genes si se eliminan varias enzimas y se cultiva al organismo en un ambiente permisible. Por medio de un proceso aún más drástico, se podrían eliminar las proteínas ribosomales, reduciendo el número a 40 o 50 genes [1].

Es todavía difícil determinar cuáles son los genes necesarios para desarrollar un genoma mínimo. Ninguno de los sistemas descritos anteriormente han sido construídos. Estos trabajos permiten identificar genes aparentemente esenciales, aún cuando la función de muchos de estos sea desconocida.

### **4.4 Avances en la construcción del genoma mínimo**

Aunque aún no se ha construido un genoma mínimo, ya existen investigaciones que establecen el

terreno a partir del cual este proyecto puede desarrollarse. Por un lado, están los trabajos del grupo de Pier Luigi Luisi, que busca construir un conjunto de genes y enzimas en un compartimiento celular e incrementar de forma gradual la complejidad del sistema hasta que comience a mostrar características de vida celular [1].

El grupo de Pier Luigi Luisi utiliza liposomas como compartimientos celulares dentro de los cuales se intenta desarrollar una célula mínima. Con anterioridad, lograron generar un sistema en el cual se sintetizaba lecitina dentro de liposomas de lecitina. De esta forma, las vesículas se reproducían de forma indefinida mientras existiera el sustrato necesario en el medio. Esto está lejos de considerarse una célula mínima, pero bien puede verse como uno de los primeros pasos hacia el desarrollo de dicho proyecto. Un experimento posterior utilizó tal sistema incluyendo además la ribozima Q $\beta$  replicasa del colifago Q $\beta$ . En este experimento, una secuencia de RNA se replicaba dentro del liposoma gracias a la acción de la replicasa, aunque la división del liposoma y la replicación del RNA ocurrían de forma independiente.

El sistema mencionado anteriormente no es estable, ya que la enzima Q $\beta$  no se replica, de modo que en cada evento de división la cantidad de enzima presente en las vesículas se reduce. Eventualmente, las vesículas son incapaces de reproducir RNA.

Se han desarrollado sistemas más complejos que utilizan un extracto celular que contiene toda la maquinaria celular para la síntesis de proteínas. Sin embargo, estos trabajos están todavía lejos de producir un sistema estable.

Otra aproximación que permite obtener un vector para la construcción de un genoma mínimo es el trasplante de genomas. Esta técnica, utilizada por el grupo de Venter, involucra la inclusión de un cromosoma bacteriano en una célula hospedera para luego seleccionar las clonas que transformaron DNA de forma exitosa, que no recombinación los dos cromosomas y que al dividirse sólo incluyeron el cromosoma de interés. Este sistema se ha probado de forma exitosa trasplantando el cromosoma bacteriano de *Mycoplasma mycoides* en una célula de *Mycoplasma capricolum*, dando lugar a células de *M. mycoides*.

De forma paralela al trasplante de genomas, la síntesis de oligonucleótidos ha avanzado al grado de que no sólo es posible sintetizar secuencias que contienen unos pocos genes sino incluso genomas virales y, de forma más reciente, el genoma completo de *Mycoplasma genitalium*.

#### **4.5 Implicaciones de la construcción del genoma mínimo.**

Más allá de las posibles aplicaciones que tiene un genoma mínimo, su construcción es la culminación de un programa dentro de la biología sintética [8]. Las metas de este programa quedan patentes en la pregunta que Randy Rettberg, uno de los investigadores que organizan el iGEM, formuló: “¿Pueden construirse sistemas biológicos simples a partir de partes estandarizadas e intercambiables?”

La construcción de un sistema biológico afecta la forma en la que se conciben los organismos de



una forma similar en la que la síntesis de la urea, realizada hace casi dos siglos por Friedrich Wöhler, cambio la forma en la que se concebían las moléculas orgánicas [8].

Como se mencionó al principio de este capítulo, las células se conciben cada vez más como un conjunto de partes que pueden ensamblarse para producir un organismo. La construcción de un genoma mínimo desarrolla esta idea al establecer que las partes físicas son suficientes para construir un sistema biológico.

Al menos un grupo de investigadores ha intentado patentar un genoma mínimo [4], generando polémica sobre lo que se considera un bien común y lo que puede proteger una patente, aunque es necesario señalar que ya existían patentes sobre organismos con anterioridad. De una forma indirecta el desarrollo de un genoma mínimo también ha afectado discusiones sobre lo que constituye un organismo y lo que social o legalmente consideramos vivo.

#### **4.6 Perspectivas**

Los diversos proyectos para desarrollar un genoma mínimo aún están en sus albores; ya existen las técnicas necesarias para sintetizar e implantar un cromosoma de gran tamaño [9], pero todavía está bajo discusión cuál es la secuencia de DNA que debe usarse.

Se desconocen cuáles son las interacciones existentes entre los diversos genes, de forma que es difícil aislar los componentes de un sistema para la construcción modular que un proyecto de esta magnitud requiere. Asimismo, se desconoce la función de una tercera parte de los genes que se han catalogado como esenciales, lo que dificulta predecir el comportamiento final del genoma mínimo.

Dada la magnitud de los proyectos del genoma mínimo, se requiere probar e integrar varios sistemas que podrían tener una aplicación inmediata y revelar información útil sobre el funcionamiento de los organismos, antes de que se complete la síntesis de un genoma mínimo.

Es difícil estimar cuánto tiempo tardará en completarse un genoma mínimo o las aplicaciones derivadas de su construcción. Sin embargo, este proyecto es un paso necesario para el desarrollo de protocolos que involucren la síntesis de cromosomas de gran tamaño y el diseño radical de organismos.

#### **Bibliografía**

- 1) Luisi, P. L. (2007). Chemical aspects of synthetic biology. *Chemistry & Biodiversity*, 4(4):603-621.
- 2) Forster, A. C. and Church, G. M. (2006). *Towards synthesis of a minimal cell*. *Mol Syst Biol*, 2.
- 3) Cho, M. K., Magnus, D., Caplan, A. L., Mcgee, D. (1999). *Genetics:ethical considerations in synthesizing a minimal genome*. *Science*, 286(5447):2087-2090.

- 4) John I., Smith; Hamilton O., Hutchison; Clyde A. III ; Alperovich; Nina Y., Assad-Garcia, Nancy. (2006) US Patent Application: 20070122826
- 5) Mushegian, A. R. and Koonin, E. V. (1996). *A minimal gene set for cellular life derived by comparison of complete bacterial genomes*. PNAS, 93(19):10268-10273.
- 6) Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2006). *Essential genes of a minimal bacterium*. PNAS, 103(2):425-430.
- 7) Gil, R., F. J. Silva, J. Pereto, and A. Moya. (2004). *Determination of the core of a minimal bacterial gene set*. Microbiol. Mol. Biol. Rev. 68:518-537
- 8) Cho, M. K., Magnus, D., Caplan, A. L., McGee, D., and of Genomics Group, E. (1999). *Ethical considerations in synthesizing a minimal genome*. Science, 286(5447):2087-2090.
- 9) Lartigue, C., Glass, J. I., Alperovich, N., Pieper, R., Parmar, P. P., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2007). *Genome transplantation in bacteria: changing one species to another*. Science (New York, N.Y.), 317(5838):632-638.

## Capítulo 5

### Situación actual y perspectivas.

#### Sobre este capítulo

En este capítulo se presentan de forma breve los experimentos, artículos y conferencias recientes que han tenido un mayor impacto sobre el desarrollo de la biología sintética, se analiza cómo la han afectado, se presentan los diversos obstáculos que afectan el desarrollo de la disciplina y se discute la forma en que la misma se desarrolla en diferentes partes del mundo y en México.

#### 5.1 Los problemas que enfrenta la disciplina

El número de investigadores dentro de la biología sintética, ha aumentado de forma exponencial en los últimos años, lo que puede servir como un indicador del desarrollo de la disciplina [1]. El número de partes disponibles para la construcción de circuitos biológicos (que se encuentran en el *Registry of Standard Biological Parts*) ha crecido de manera similar en número y tamaño, aunque no en complejidad<sup>1</sup>.



**Fig 1.** La portada de la prestigiosa revista *Nature* de noviembre de 2005, que estuvo dedicada a la biología sintética. Esto refleja la importancia que había cobrado la disciplina y, de una forma indirecta, su rápido desarrollo.

Actualmente se combinan partes simples para generar módulos que incluyen generadores de pulsos [2], circuitos de efecto retardado [3], contadores [4], generadores de patrones espacio-temporales [5] y compuerta lógicas. El ensamblaje de varios módulos para generar sistemas complejos es todavía difícil de realizar, algo que varios autores [1][6], asocian con problemas que enfrenta el diseño de sistemas biológicos.

Entre las dificultades mencionadas, se encuentra la falta de caracterización de las partes, la incompatibilidad entre ellas, el comportamiento emergente que viene como resultado de las múltiples interacciones entre los componentes y la gran variabilidad que existe en los sistemas biológicos [6]. Como se ha visto en capítulos anteriores, estos problemas han sido estudiados y parcialmente resueltos aunque todavía constituyen un obstáculo para el desarrollo de la biología sintética.

Si bien es cierto que los sistemas biológicos construidos hasta el momento son relativamente simples y pocos de ellos funcionan como se espera, no debe subestimarse la importancia que tienen estos proyectos. Brian J. Yeh y Wendell A. Lim argumentan que dichos sistemas proveen información útil para generar modelos que permitan superar algunos de los problemas que actualmente enfrenta la biología sintética [7].

#### 5.2 Los trabajos de mayor impacto y la dirección de la biología sintética.

Entre los trabajos recientes que han tenido un gran impacto sobre la biología sintética, cabe mencionar el repesilador de Elowitz [5], los intentos por sintetizar un genoma mínimo que realiza

Craig Venter [8] y la modificación de *Escherichia coli* para la producción de terpenoides que tienen un uso terapéutico [9].

Estas investigaciones muestran varias direcciones en las que está desarrollándose la biología sintética. Dichos trabajos permitirían, a corto plazo, realizar experimentos sobre sistemas biológicos que no estén constreñidos por su historia evolutiva, producir sustancias químicas cuya síntesis actualmente no es posible o rentable [9], diseñar tratamientos médicos “inteligentes”, en donde el agente terapéutico pueda analizar la información para tomar decisiones complejas [10], y construir materiales o tejidos para los cuales se requiere que las células se organicen o diferencien, formando patrones espaciales complejos.

El patrón de desarrollo de la biología sintética es, en muchos aspectos, similar al que han experimentado otras disciplinas. De esta forma, cabe esperar que surjan nuevas aplicaciones que, de una forma inesperada cambien radicalmente la dirección que actualmente lleva [7] [11].

### 5.3 El desarrollo de la biología sintética

Es difícil cuantificar el desarrollo de la biología sintética, aunque se refleja en varios indicadores. El número de artículos asociados a la disciplina que se publica cada año se utilizó, en un par de reportes para la European Commission Framework Programme [12][13], para analizar la forma en la que se desarrolla esta disciplina en Europa, Estados Unidos y el resto del mundo.

Dicho reporte [12] mostraba un aumento en la publicación de artículos que se asocian con la biología sintética, indicando un crecimiento irregular que se dispara en el año 2004. Asimismo, se observó un desarrollo desigual entre distintos países [13]. El 63% de los artículos fueron publicados por organizaciones que residen en Estados Unidos, el 19% en países de la Unión Europea, 6% en Japón, 4% en Israel, y el 8% restante en Canadá, Suiza, China, Corea del Sur y otros países.

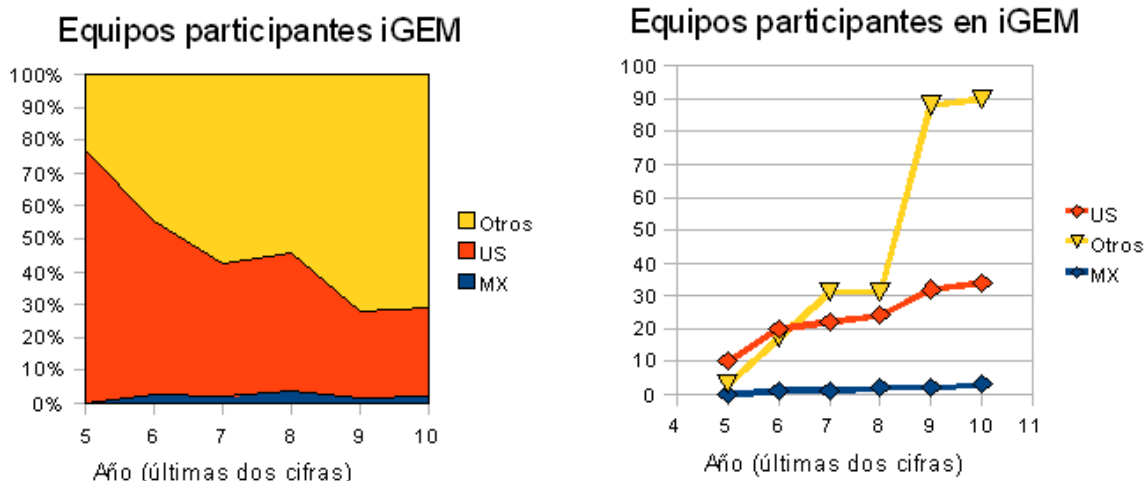
La importancia que adquieren las instituciones cuya sede se encuentra en Estados Unidos aumenta si se evalúa el impacto de los artículos usando el Thompson Scientific Index, en cuyo caso se pueden atribuir hasta el 74% de las aportaciones a instituciones de este país. El efecto se acentúa si se considera que los artículos publicados en este país se concentran en un número reducido de instituciones localizadas en Massachusetts y California (MIT, Harvard, Stanford, Berkeley, Caltech, UCLA).

Es necesario considerar que dichos reportes [12][13] incluyen una gran lista de artículos -más de mil-, que los autores asocian con la biología sintética, pero que no necesariamente se incluyen dentro de ésta. Así, los resultados anteriormente mencionados reflejan más la forma en la que se ha desarrollado la biología molecular y la ingeniería genética, que la biología sintética.

Una búsqueda rápida por artículos que al menos incluyan las palabras “biología sintética” entre el año 2002 y el 2009 en tres de las publicaciones científicas de mayor importancia, Science, Nature y PNAS, genera una lista de 179 artículos. El número de artículos que puede asociarse a la biología sintética, aún con un criterio muy laxo, es reducido y se observa un crecimiento paulatino (media

geométrica del crecimiento = 156%). No obstante, es imposible establecer una correlación entre los resultados de esta búsqueda y los datos provistos en el reporte, ya que este último no da información sobre publicaciones posteriores al año 2004.

Otro indicador de desarrollo de la biología sintética es el número de trabajos presentados en congresos informales de biología sintética como el iGEM. El número de proyectos que se presentan en esta competencia ha aumentado desde 13 en el 2005 a 127 en el 2010. La proporción de equipos que se presentan ha cambiado, siendo que los asociados a una institución estadounidense conformaban el 77% en el 2005 y en el 2010 sólo constituyeron el 27% de los equipos participantes.



Es necesario señalar que existen varios países, entre ellos México, en donde el número de trabajos publicados que los autores asocian específicamente a la biología sintética, son reducidos. No obstante, el desarrollo de la disciplina queda patente en proyectos que por sus características podrían incluirse dentro de la biología sintética o en trabajos presentados en congresos informales.

#### 5.4 La biología sintética en México

Son pocos los trabajos realizados en México que se asocian a la biología sintética de forma directa y utilizando un criterio estricto. Se puede observar el desarrollo de la disciplina en los proyectos que se han presentado en el iGEM del año 2006 a la fecha.

Entre estos trabajos, cabe mencionar el desarrollo de un dispositivo genético que permite a las células formar patrones de Turing. Este proyecto, que en el año 2009 se presentó en el iGEM obteniendo una medalla de oro por su participación, es precedido por trabajos que se remontan a la primera participación de un equipo mexicano en esta prestigiosa competencia, en el año 2006.

Ya desde el 2006, un grupo integrado por estudiantes e investigadores del IPN y de la UNAM se planteaba como objetivo el desarrollo de una red genética que fuese capaz de permitir la expresión de patrones de Turing en colonias de bacterias. No fue sino hasta el siguiente año cuando se diseñó

una construcción genética con las partes que en ese momento existían en el repositorio del MIT y hasta el 2009 se ensambló y probó un circuito genético.

El diseño de los circuitos biológicos ha cambiado a lo largo del tiempo y en todos ellos se pretendía producir patrones espacio-temporales, aunque no es seguro que estos pudiesen funcionar. Se probó exitosamente la actividad de algunas partes del circuito desarrollado en el 2009. Este circuito nunca se completó y enfrentaba problemas con los que sus antecesores habían lidiado, como lo es la larga vida media de las lactonas (24 hs), la imposibilidad de degradar lactonas fuera del medio celular y la diafonía que se produce al utilizar señales químicas con una estructura similar.

Un segundo grupo, asociado al Centro de Ciencias Genómicas de la UNAM participó en el iGEM en el 2008 presentando un dispositivo genético que permitiría a una colonia de bacterias cambiar la concentración de zinc en el medio y de esta forma utilizarlo como una señal. En el 2009 este equipo también participó presentando un circuito genético que, de funcionar, permitiría a una colonia de bacterias reconocer una invasión por fagos T7 y enviar una señal para que las células vecinas produzcan iRNA para impedir la producción de proteínas necesarias para la replicación de fagos T7 y de esta forma detener la infección.

Un tercer equipo asociado al Tecnológico de Monterrey se ha inscrito para el iGEM 2010 aunque todavía no existe información sobre el desarrollo de su proyecto.

El número de grupos de investigación dedicados a la biología sintética en México crece de forma lenta pero constante. Los investigadores provienen de diversas áreas y no ha sido hasta tiempos recientes cuando se han introducido en la disciplina, aunque un creciente número de estudiantes incluye en su formación programas específicamente dirigidos a enseñar biología sintética.

El número de artículos publicados en México que se asocian a la biología sintética es todavía muy reducido, aunque se puede esperar que esto cambie a medida que se formen más investigadores que trabajen en el campo.

## **Bibliografía**

- 1) Purnick, P. E. M. and Weiss, R. (2009). *The second wave of synthetic biology: from modules to systems*. Nature Reviews Molecular Cell Biology, 10(6):410-422.
- 2) Basu, S., Mehreja, R., Thiberge, S., Chen, M. T. & Weiss, R. (2004) *Spatiotemporal control of gene expression with pulse-generating networks*. Proc. Natl Acad. Sci. USA **101**, 6355–6360 .
- 3) Weber, W., Kramer, B. P. & Fussenegger, M. (2007) *A genetic time-delay circuitry in mammalian cells*. Biotechnol. Bioeng. **98**, 894–902 .
- 4) Friedland, A. E., Lu, T. K., Wang, X., Shi, D., Church, G., and Collins, J. J. (2009). *Synthetic gene networks that count*. Science, 324(5931):1199-1202.

- 5) Elowitz, M. B. and Leibler, S. (2000). *A synthetic oscillatory network of transcriptional regulators*. Nature, 403(6767):335-338.
- 6) Kwok, R. (2010). *Five hard truths for synthetic biology*. Nature, 463(7279):288-290.
- 7) Yeh, B. J. and Lim, W. A. (2007). *Synthetic biology: lessons from the history of synthetic organic chemistry*. Nat Chem Biol, 3(9):521-525.
- 8) Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2006). *Essential genes of a minimal bacterium*. PNAS, 103(2):425-430.
- 9) Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D., and Keasling, J. D. (2003). *Engineering a mevalonate pathway in escherichia coli for production of terpenoids*. Nature biotechnology, 21(7):796-802.
- 10) Andrianantoandro, E., Basu, S., Karig, D. K., and Weiss, R. (2006). *Synthetic biology: new engineering rules for an emerging discipline*. Molecular systems biology, 2:msb4100073-E1-msb4100073-E14.
- 11) Report of a NEST High-Level Expert Group; Directorate-General for Research. (2005). *Synthetic biology- applying engineering to biology*. European Commission. EUR 21796.
- 12) Synbiology (2005) *Synbiology, an analysis of synthetic biology research in Europe and North America. Output D3: literature and statistical review*. European Commission Framework Programme 6 reference contract 15357 (NEST).
- 13) Synbiology (2006) *Synbiology, an analysis of synthetic biology research in Europe and North America. final report on analysis of synthetic biology sector*. European Commission Framework Programme 6 reference contract 15357 (NEST).