



## **Evolución dirigida de enzimas: una revolución en Química**

**Diana M. Mate**

Departamento de Biología Molecular, Centro de Biología Molecular “Severo Ochoa”,  
Universidad Autónoma de Madrid, Madrid, España.

(\*) [diana.mate@uam.es](mailto:diana.mate@uam.es)

**Recibido:** 25/05/2019

**Aceptado:** 30/05/2019

---

### **Resumen**

La evolución dirigida de enzimas es una poderosa metodología surgida a inicios de los años 90 del pasado siglo que permite la obtención de proteínas con características mejoradas o con nuevas funciones nunca antes requeridas en la naturaleza. Esta técnica imita en el laboratorio los principios darwinianos de la evolución natural: la introducción de mutaciones al azar en el ADN que codifica para la proteína de interés y la posterior selección de aquellas variantes de la enzima que presentan mejoras en la propiedad deseada. Este proceso se repite las veces necesarias hasta lograr biocatalizadores con aplicación en biomedicina o en la industria. La relevancia de la evolución dirigida de enzimas queda reflejada por los numerosos grupos de investigación y empresas de todo el mundo que tienen líneas de trabajo sobre esta técnica, así como por el hecho de que su pionera, la profesora Frances H. Arnold del Instituto de Tecnología de California, fuera reconocida con el Premio Nobel de Química en 2018.

**Palabras claves:** Enzima; *Escherichia coli*; evolución dirigida; mutagénesis aleatoria; *Saccharomyces cerevisiae*; screening

### **Abstract**

**Directed evolution of enzymes: a revolution in Chemistry.** Directed evolution of enzymes is a powerful methodology emerged in the early 90s of the past century that allows obtaining proteins with improved characteristics or with new functions never before required in nature. This technique imitates in the laboratory the Darwinian principles of natural evolution: the incorporation of random mutations in the gene that codifies for the enzyme of interest, and subsequent selection of enzyme variants with improvements in the desired property. This process is repeated as many times as necessary until achieving biocatalysts with application in biomedicine or in industry. The relevance of this methodology is reflected by the numerous academic and company laboratories throughout the world that have lines of research on this subject. Significantly, the pioneer in directed evolution, Prof. Frances H. Arnold of the California Institute of Technology, was recognized with the Nobel Prize in Chemistry in 2018.

**Keywords:** Enzyme; *Escherichia coli*; Directed evolution; Random mutagenesis; *Saccharomyces cerevisiae*; screening.

### **Introducción**

Desde que surgieron las primeras formas de vida hace 3.700 millones de años, todos los ambientes naturales de nuestro planeta han sido colonizados por organismos, incluso aquellos con las condiciones más extremas. Así, podemos encontrar animales y plantas capaces de resistir las altas temperaturas y la sequedad del desierto, criaturas abisales adaptadas a las elevadas presiones de los océanos más profundos e incluso virus y bacterias que resisten el frío inhóspito de los Polos. Todo esto es posible porque a lo largo de millones de años la evolución natural ha sido capaz de resolver de forma gradual una serie de desafíos fisicoquímicos que han permitido la aparición de formas de vida cada vez más complejas. Uno de los retos más importantes de la vida es la transmisión de la información necesaria para perpetuarla generación tras generación. Y es aquí donde juega un papel fundamental el ADN, la conocida como molécula de la vida, que en forma de genes permite que se herede y se desarrolle generación tras generación. Los genes, a su vez, codifican las que se consideran las herramientas químicas de la vida, las proteínas, las cuales se

han optimizado y modificado, creando una diversidad incommensurable.

La investigadora Frances H. Arnold, del Instituto de Tecnología de California (Caltech) y reconocida con la mitad del Premio Nobel de Química de 2018, se inspiró en los procesos naturales implicados en la evolución natural y utilizó los mismos principios, los cambios genéticos y la selección, con el objetivo de generar nuevas enzimas que aporten beneficios a la humanidad. Este proceso se conoce como evolución dirigida de enzimas, siendo las enzimas una clase especial de proteínas caracterizadas por tener la capacidad de acelerar reacciones químicas. En 1991 Arnold publicó el primer trabajo sobre evolución dirigida de enzimas, en concreto, sobre la evolución de una proteasa para mejorar su resistencia a disolventes orgánicos<sup>1</sup>. En los casi 30 años que han pasado desde entonces, son cientos las enzimas diseñadas por evolución en el laboratorio que se utilizan para la obtención de productos muy diversos, los cuales abarcan desde fármacos hasta biocombustibles.

La otra mitad del premio Nobel fue otorgado a los investigadores George P. Smith y Gregory Winter. En 1985, George P. Smith desarrolló un elegante método denominado visualización de fagos (*phage display* en inglés), que está basado en el uso de bacteriófagos -virus que infectan a las bacterias- para el desarrollo de nuevas proteínas<sup>2</sup>. Por su parte, Gregory Winter utilizó la presentación de fagos para la evolución dirigida de anticuerpos, con el objetivo de producir nuevos fármacos<sup>3</sup>. El primero basado en este método, el adalimumab, fue aprobado por la Administración de Medicamento y Alimentos de Estados Unidos en 2002 y se utiliza para el tratamiento de enfermedades como la psoriasis, la enfermedad de Crohn o la artritis reumatoide. Desde entonces, la presentación de fagos ha producido anticuerpos que pueden neutralizar toxinas, contrarrestar enfermedades autoinmunes y curar el cáncer metastásico.

En este artículo abordaremos los aspectos generales de la evolución in vitro, los microorganismos que se utilizan como hospedadores, los métodos empleados para la generación de diversidad génica y las técnicas que permiten la exploración de las librerías de variantes enzimáticas. Por último, trataremos de dilucidar cómo diseñar una ruta evolutiva de manera eficiente para poder obtener la enzima con las propiedades deseadas.

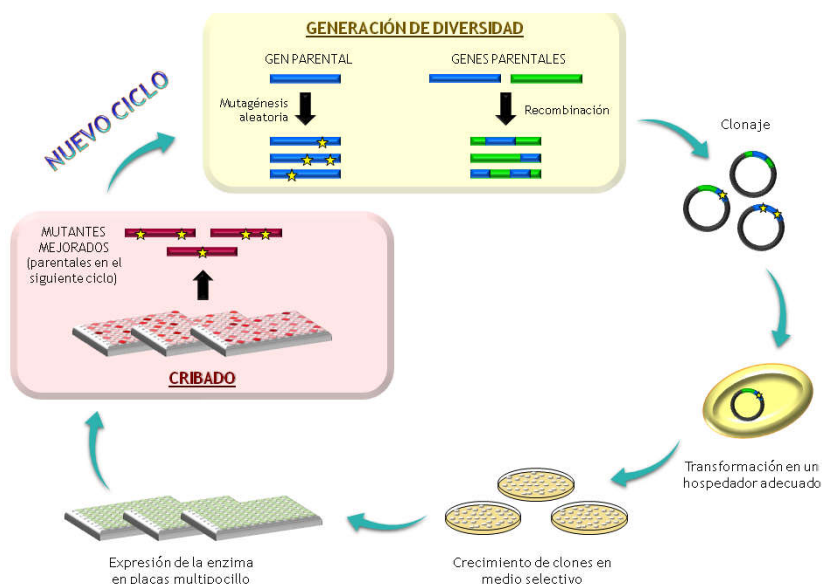
### Aspectos generales sobre evolución dirigida

La evolución molecular dirigida o evolución in vitro es una poderosa metodología que permite la obtención de proteínas con propiedades mejoradas o con nuevas funciones nunca antes requeridas en la naturaleza<sup>4,6</sup>. Antes de la aparición de esta técnica a inicios de 1990, el diseño racional o mutagénesis dirigida, consistente en el cambio de aminoácidos concretos en base a modelos tridimensionales o estructuras cristalográficas, constituía la vía más frecuente para la creación de nuevas actividades enzimáticas. A pesar de los grandes avances en la

elucidación estructural de proteínas mediante difracción de rayos X y espectroscopía de resonancia magnética nuclear, del desarrollo de métodos computacionales y predictivos para anticipar las interacciones enzima-sustrato<sup>7</sup>, o de la reconstrucción de genomas gracias a la disponibilidad de extensas bases de datos<sup>8</sup>, el conocimiento acumulado sobre el plegamiento, mecanismo y función catalítica no es garantía suficiente para el diseño racional exitoso de biocatalizadores<sup>9,10</sup>.

La evolución dirigida recrea en el laboratorio los principios darwinianos de evolución natural, permitiendo el diseño de enzimas con propiedades de alto interés biotecnológico en ausencia de información estructural o mecanística previa. En la evolución dirigida la presión selectiva es controlada por el científico en busca de una mejora en una propiedad enzimática concreta al tiempo que la escala temporal se reduce a tan sólo meses, e incluso días, de trabajo en el laboratorio<sup>11-15</sup>.

Un experimento de evolución dirigida comienza con la generación de diversidad génica mediante mutagénesis aleatoria y/o recombinación de los genes que codifican para la proteína de interés. Esta diversidad generada se expresa en un microorganismo hospedador adecuado y se explora bajo unas condiciones muy determinadas a fin de seleccionar aquellas variantes enzimáticas que presenten una mejora en la propiedad objeto de estudio. Finalmente, los genes que codifican para estas enzimas se aíslan con el fin de analizar las mutaciones introducidas y/o los procesos de recombinación que han tenido lugar. Los genes “ganadores” se someten a un nuevo ciclo de mutación y/o recombinación seguido de selección o cribado (en inglés, *screening*), acumulándose mejoras en la propiedad de interés generación tras generación (figura 1). Este proceso se repite tantas veces como sean necesarias, de manera que la introducción, recombinación y transmisión de las mutaciones adquiridas a lo largo de las generaciones conducirá a la obtención del biocatalizador con las características deseadas<sup>16,17</sup>.



**Fig. 1:** Esquema general de un experimento de evolución molecular dirigida. Las mutaciones están representadas como estrellas.

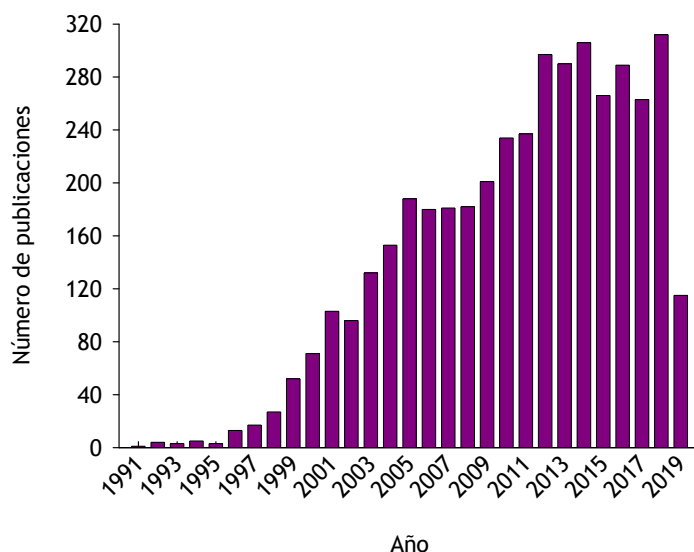
Las mutaciones beneficiosas descubiertas en experimentos de evolución *in vitro* se encuentran a menudo en zonas alejadas de los sitios catalíticos de interés, lo que permite descubrir regiones o posiciones determinadas que difícilmente podrían haberse anticipado mediante un enfoque racional<sup>18</sup>. No obstante, ambas herramientas de diseño de proteínas no son incompatibles, sino todo lo contrario. La combinación de estrategias que dirigen la mutagénesis a aminoácidos o regiones concretas de una proteína con la evolución dirigida de todo el gen puede facilitar la obtención de enzimas de gran interés biotecnológico<sup>19,20</sup>. En este sentido, se han desarrollado numerosos algoritmos bioinformáticos basados en el análisis computacional de conjuntos de datos de secuencias o de relaciones secuencia-función y que permiten el diseño de colecciones de mutantes en las que predominan las mutaciones beneficiosas frente a las perjudiciales<sup>21</sup>.

Tras casi 30 años de historia, la evolución dirigida se ha convertido en la herramienta más potente y versátil para el diseño de proteínas<sup>10</sup>. Así, haciendo uso de la evolución *in vitro* se han mejorado multitud de propiedades enzimáticas, por ejemplo, la estereo-selectividad<sup>22</sup>, la regioselectividad<sup>23,24</sup>, la quimioespecificidad<sup>25,26</sup>, la expresión funcional heteróloga<sup>27-30</sup>, la termoestabilidad<sup>31,32</sup>, la estabilidad oxidativa<sup>19,33</sup>, la actividad a pH alcalino<sup>34-36</sup>, así como la luminiscencia y la fotoestabilidad de proteínas fluorescentes<sup>37,38</sup>. La evolución dirigida también ha permitido conseguir actividad en medios no naturales para las enzimas como son los disolventes orgánicos<sup>1,39-41</sup>, los líquidos iónicos<sup>42,43</sup> o la sangre humana<sup>44</sup>. Además, en el marco de la biología sintética, la evolución *in vitro* está facilitando la ingeniería de complejas rutas metabólicas y de genomas enteros<sup>45,46</sup>. También se han desarrollado métodos de evolución *in vivo* como la plataforma de evolución continua asistida por fagos (PACE), de la que hablaremos más adelante, o técnicas basadas en el proceso de hipermutación somática que ocurre en los linfocitos B de mamíferos<sup>47</sup>. En estos métodos la función objeto de evolución está acoplada a una propiedad celular fácilmente detectable, como la supervivencia o la fluorescencia, y han permitido acelerar aún más el camino evolutivo, llevando a cabo docenas de generaciones en un solo día con una intervención mínima por parte del investigador.

La importancia de la evolución dirigida queda demostrada por la tendencia exponencial creciente en el número de trabajos científicos publicados sobre el tema desde sus inicios a principios de los años 90 del pasado siglo (figura 2), por la disponibilidad en el mercado de diversos productos desarrollados mediante evolución en el laboratorio y por la inversión en proyectos de investigación sobre evolución dirigida de enzimas que realizan importantes empresas biotecnológicas como Novozymes, Maxygen y Codexis, entre otras<sup>48</sup>.

#### *Elección del microorganismo hospedador*

La expresión funcional del gen que codifica para la proteína



**Fig. 2:** Número de artículos científicos sobre evolución dirigida publicados por año desde 1990 hasta abril de 2019 inclusive. Información obtenida de la base de datos *Scopus* introduciendo “directed evolution” como término de búsqueda.

de interés es el primer requisito para llevar a cabo un experimento de evolución dirigida. Sin embargo, seleccionar un microorganismo hospedador adecuado para la expresión de un determinado gen es a menudo difícil ya que la expresión génica está controlada por numerosos factores a nivel transcripcional y traduccional cuyos efectos sobre el gen de interés son difícil de predecir<sup>49</sup>. Aunque son muchos los microorganismos cultivables que podrían servir a priori como hospedadores en ensayos de evolución dirigida, hasta la fecha sólo se han utilizado varias especies bacterianas y de levadura así como contadas líneas celulares de insectos y de mamíferos<sup>50</sup>.

Dos de las especies bacterianas más empleadas como hospedador en estudios de evolución dirigida son *Thermus thermophilus* y *Escherichia coli*. Por un lado, la bacteria hipertermófila *Thermus thermophilus* se caracteriza por un sistema de competencia natural, una elevada temperatura de crecimiento (62-75°C dependiendo de la cepa), una rápida tasa de crecimiento, una alta eficiencia de transformación ( $10^8$ - $10^9$  colonias por  $\mu\text{g}$  de ADN) y un sistema de competencia natural que permite la transformación de ADN exógeno<sup>51</sup> (tabla 1). Es precisamente su capacidad de crecer a alta temperatura la que ha permitido su uso para la obtención de proteínas termoestables mediante evolución dirigida<sup>52</sup>. Por su parte, la bacteria mesófila *Escherichia coli* es, con diferencia, el microorganismo más utilizado como hospedador debido a su alta eficiencia de transformación ( $10^8$ - $10^{10}$  colonias por  $\mu\text{g}$  de ADN), su rápida tasa de crecimiento, su capacidad de mantener estable el ADN plasmídico, el detallado conocimiento de su genoma y la disponibilidad en el mercado de técnicas rutinarias de biología molecular que permiten su fácil manipulación<sup>50</sup> (tabla 1). No obstante, la expresión bacteriana de genes eucarióticos a menudo se ve limitada por las diferencias entre los sistemas de expresión bacterianos y nativos, por ejemplo,

**Tabla 1:** Resumen de las características y herramientas genéticas disponibles para las bacterias *Escherichia coli* y *Thermus thermophilus* y las levaduras *Saccharomyces cerevisiae* y *Pichia pastoris*.

	Competencia natural	Eficiencia de transformación (UFC <sup>a</sup> /μg de ADN)	Disponibilidad de plásmidos episómicos	Temperatura óptima de crecimiento (°C)	Tiempo de duplicación (h)	Secreción de proteínas al medio extracelular
<b>Bacteria</b>						
<i>Escherichia coli</i>	✗	10 <sup>8</sup> -10 <sup>10</sup>	✓	37	0,25-0,33	✗*
<i>Thermus thermophilus</i>	✓	10 <sup>8</sup> -10 <sup>9</sup>	✓	62-75	0,58-1	✗*
<b>Levadura</b>						
<i>Saccharomyces cerevisiae</i>	✗	10 <sup>7</sup> -10 <sup>8</sup>	✓	30	1,25-2	✓
<i>Pichia pastoris</i>	✗	10 <sup>5</sup> -10 <sup>6</sup>	✗	30	1,5-2	✓

a. UFC = unidades formadoras de colonias. \*Se requiere un paso adicional de lisis celular.

la ausencia de chaperonas y de modificaciones post-traduccionales o el distinto uso de codones (aunque existen cepas diseñadas especialmente para expresar genes con alto contenido de codones raros en *E. coli*<sup>53</sup>). Tales diferencias dan lugar a bajos niveles de expresión o a un plegamiento incorrecto de las proteínas, lo que origina su degradación o su acumulación en cuerpos de inclusión<sup>27</sup>. Estos inconvenientes pueden evitarse eligiendo hospedadores eucarióticos como las levaduras *Pichia pastoris* y *Saccharomyces cerevisiae*.

Hasta la fecha, la levadura metanotrófica *Pichia pastoris* se ha utilizado para la producción heteróloga de más de 1000 proteínas, gracias a que, al ser un organismo eucariota, tiene la capacidad de producir proteínas solubles correctamente plegadas con las modificaciones post-traduccionales adecuadas y el correspondiente procesamiento proteolítico<sup>54</sup>. Además, su cultivo en fermentadores en continuo da lugar a densidades celulares muy altas (>130 g de biomasa celular seca por litro de cultivo), permitiendo la secreción de grandes cantidades de proteína<sup>55</sup>. Sin embargo, la inmensa mayoría de los vectores disponibles para la expresión de proteínas heterólogas en *P. pastoris* son integrativos, lo que constituye el principal obstáculo para el uso de esta levadura como hospedador en evolución dirigida<sup>56</sup> (tabla 1). Aunque se han desarrollado sistemas basados en la integración de cassettes de expresión mediante PCR de extensión por solapamiento para la generación de librerías de mutantes de hidroxinitril liasas<sup>57</sup> (Liu et al., 2008), la integración del plásmido en el genoma hace más laboriosa la recuperación de los genes y, por consiguiente, dificulta el proceso evolutivo.

Estos problemas pueden resolverse haciendo uso de la levadura *Saccharomyces cerevisiae* como microorganismo hospedador ya que tiene una mayor eficiencia transformación que *P. pastoris* (hasta 10<sup>7</sup>-10<sup>8</sup> colonias por μg de ADN), también lleva a cabo modificaciones post-traduccionales, y dispone de un mecanismo muy eficaz de secreción de proteínas. Además, al contrario que *P. pastoris*, *S. cerevisiae* permite trabajar con diversos vectores episómicos multicopia que facilitan la recuperación de los genes seleccionados en el

screening<sup>50,58</sup> (tabla 1). Una ventaja muy atractiva de *S. cerevisiae* es que, a diferencia de lo que ocurre en *E. coli* y *P. pastoris*, tiene una elevada frecuencia de recombinación homóloga que facilita la creación de diversidad génica in vivo a partir de varios genes parentales<sup>56</sup>. Por otra parte, la ligación de los genes mutados en vectores de expresión es un paso crítico que requiere de PCRs adicionales pudiendo introducirse nuevas mutaciones en regiones no deseadas. El mecanismo reparador de huecos (*in vivo gap repair*) de *S. cerevisiae* permite sustituir la ligación in vitro, obteniendo *in vivo* plásmidos circulares de replicación autónoma por co-transformación del plásmido linealizado y de genes mutados que contienen secuencias homólogas (de entre 20 y 50 pares de bases, pb) con ambos extremos del plásmido<sup>59</sup>.

#### Generación de diversidad génica

El éxito de cualquier experimento de evolución dirigida depende, en gran medida, de la disponibilidad de una amplia gama de metodologías que permitan la creación de genotecas con un tamaño ajustado a la capacidad del ensayo de *screening* y con una calidad adecuada, la cual está determinada por la frecuencia de las mutaciones/recombinaciones y por el tipo de mutaciones introducidas.

Los procedimientos empleados típicamente en experimentos de evolución dirigida para generar variabilidad génica son los métodos mutagénicos y los basados en la recombinación del material genético<sup>60</sup>.

#### Métodos mutagénicos

**a) Mutagénesis aleatoria:** aunque desde un punto de vista técnico todavía no se han optimizado muchas de las condiciones deseables para experimentos de mutagénesis al azar, el objetivo ideal de estas aproximaciones es el poder sustituir de forma aleatoria cualquier aminoácido de una cadena polipeptídica por los otros 19 aminoácidos de una manera estadística y sin limitar la expresión de la proteína en el organismo hospedador. Además, sería altamente conveniente que estos métodos cumplieran los siguientes requisitos: i) tener un es-

pectro mutacional imparcial; ii) presentar una frecuencia mutacional controlable por el investigador; iii) dar lugar a sustituciones nucleotídicas consecutivas; iv) permitir mutagénesis en subconjuntos, por ejemplo, introducir principalmente aminoácidos cargados positiva o negativamente; v) ser independientes de la longitud del gen, vi) ser reproducible y sencillo; y vii) ser de bajo coste<sup>61,62</sup>.

De cara a conseguir que un proyecto de evolución dirigida sea exitoso, es esencial el ajuste preciso de la frecuencia mutacional. Un cambio aminoacídico por gen parece la tasa mutacional más adecuada para la comprensión directa de las relaciones estructura-función. Sin embargo, la mayoría de los experimentos de evolución emplean frecuencias mutacionales que inducen de uno a cuatro cambios aminoacídicos (de dos a siete cambios nucleotídicos) por cada 1.000 pb<sup>39</sup>. La selección de estas condiciones es debida a que a menudo existen limitaciones de carácter práctico, ya que los laboratorios de muchos centros de investigación sólo pueden llevar a cabo la exploración de pequeñas librerías de mutantes (de entre 2.000 y 20.000 clones). Además, dicha tasa mutacional permite la detección de mutaciones beneficiosas, mucho menos frecuentes que las perjudiciales o neutras, como se explicará en el apartado Diseño de estrategias evolutivas de este artículo<sup>63</sup>.

Con el objetivo de tener un número suficiente de clones activos, las tasas mutacionales se ajustan empíricamente para que el 40-60% de los clones retengan el 10% de la actividad del gen parental<sup>60</sup>. No obstante, existen casos excepcionales en los que se han obtenido anticuerpos con afinidades mejoradas

a partir de librerías altamente mutagenizadas (con hasta 22,5 cambios nucleotídicos por gen)<sup>64</sup>.

El método más utilizado en evolución dirigida por su sencillez y eficacia para introducir mutaciones en el material genético es la reacción de PCR propensa a error (en inglés, error-prone PCR o epPCR). Esta técnica hace uso de DNA polimerasas que exhiben baja fidelidad durante la amplificación génica. La DNA polimerasa Taq de la bacteria termófila *Thermus aquaticus* es la más empleada en reacciones de epPCR. La Taq polimerasa es una enzima robusta que carece de actividad exonucleasa 3'→5', lo que se traduce en una elevada frecuencia de error intrínseca (8,0·10<sup>-6</sup> mutaciones por nucleótido y por duplicación)<sup>65</sup>. Sin embargo, dicha tasa mutacional es insuficiente para mutagenizar de forma efectiva genes que raramente superan las 5.000 pb. Los métodos más destacados para aumentar la probabilidad de error son la adición de MnCl<sub>2</sub> y el uso de concentraciones desbalanceadas de los cuatro desoxirribonucleótidos (dNTPs) naturales<sup>65</sup>. También tienen efectos mutagénicos el incremento del tiempo de extensión y de las concentraciones de Taq polimerasa y de MgCl<sub>2</sub> en la reacción, la disminución de la concentración de DNA molde<sup>67</sup> o el uso de nucleótidos análogos a los dNTPs naturales<sup>68</sup>.

Como alternativa a la variación de parámetros, es posible aumentar la falta de exactitud de la Taq polimerasa mediante ingeniería de proteínas<sup>69</sup>, o bien emplear ADN polimerasas con diferentes espectros mutacionales como son la Mutazima®<sup>70</sup> o las DNA polimerasas humanas β, η y ι<sup>71,72</sup> (tabla 2).

**Tabla 2:** Frecuencia y tipo de mutaciones de los métodos mutagénicos basados en el empleo de ADN polimerasas. Ts/Tv: cociente transición/transversión. PolH: ADN polimerasa humana. n.d.: no determinado.

	Taq polimerasa				Mutazima® I	Mutazima® II	Método MutaGen		
	dNTPs desequilibrados <sup>a</sup>	MnCl <sub>2</sub> /dNTPs equilibrados <sup>b</sup>	MnCl <sub>2</sub> /dNTPs desequilibrados <sup>c</sup>	MnCl <sub>2</sub> /dNTPs desequilibrados <sup>d</sup>	dNTPs equilibrados <sup>e</sup>	dNTPs equilibrados <sup>f</sup>	PolH β <sup>g</sup>	PolH η <sup>h</sup>	PolH ι <sup>i</sup>
<b>Ts/Tv</b>	3,7	2,9	1,1	0,8	1,2	0,8	3,7	6,0	2,1
<b>Transiciones (%)</b>									
A→G, T→C	57,2	63,2	35,0	27,6	10,3	17,5	43,9	72,9	54,1
G→A, C→T	21,3	9,5	17,5	13,6	43,7	25,5	34,8	8,8	13,1
<b>Transversiones (%)</b>									
A→T, T→A	14,9	16,1	37,5	40,9	11,1	28,5	6,6	6,5	13,1
A→C, T→G	2,4	4,0	5,0	7,3	4,2	4,7	4,2	2,3	6,5
G→T, C→A	2,8	3,3	5,0	4,5	20,0	14,1	9,0	1,8	4,9
G→C, C→G	1,4	1,4	0,0	1,4	8,8	4,1	1,4	3,0	8,1
<b>Inserciones (%)</b>									
	1,3	n.d.	0,0	0,3	0,8	0,7	2,7	2,7	2,9
<b>Deleciones (%)</b>									
De un nucleótido	3,1	1,5	0,0	4,2	1,1	4,8	1,8	4,8	2,9
De 2 o más nucleótidos	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,1	5,8

a. [dATP] = [dGTP] = 0.2 mM y [dTTP] = [dCTP] = 1 mM (Emond *et al.*, 2008).

b. [MnCl<sub>2</sub>] = 0.25 mM y [dATP] = [dTTP] = [dGTP] = [dCTP] = 0.02 mM (Lin-Goerke *et al.*, 1997).

c. [MnCl<sub>2</sub>] = 0.15 mM y [dATP] = [dGTP] = 0.2 mM y [dTTP] = [dCTP] = 1 mM (Shafikhani *et al.*, 1997).

d. [MnCl<sub>2</sub>] = 0.5 mM y [dATP] = [dGTP] = 0.2 mM y [dTTP] = [dCTP] = 1 mM (Shafikhani *et al.*, 1997).

e. [dATP] = [dGTP] = [dTTP] = [dCTP] = 0.2 mM (Cline y Hogrefe, 2000).

f. [dATP] = [dGTP] = [dTTP] = [dCTP] = 0.2 mM (GeneMorph® II random mutagenesis kit, Stratagene).

g. Amplificación del gen entero y [dATP] = [dGTP] = 0.05 mM y [dTTP] = [dCTP] = 0.1 mM (Emond *et al.*, 2008).

h. Amplificación del gen entero y [dATP] = [dGTP] = [dTTP] = [dCTP] = 0.1 mM (Emond *et al.*, 2008).

i. Amplificación de la región del gen codificante para el dominio catalítico y [dATP] = [dGTP] = [dTTP] = [dCTP] = 0.1 mM (Emond *et al.*, 2008).

De hecho, todas las ADN polimerasas descritas tienen una gran tendencia hacia las transiciones (cambios de bases púricas por púricas y pirimidínicas por pirimidínicas) frente a las transversiones (cambios de bases púricas por pirimidínicas y a la inversa), reduciendo la variedad de los cambios nucleotídicos e impidiendo la completa exploración del espacio proteico. No obstante, la probabilidad de que tengan lugar transversiones frente a transiciones puede aumentarse mediante el empleo de la Taq en presencia de  $MnCl_2$  y concentraciones desequilibradas de dNTPs o usando la Mutazima® II (tabla 2). Además, los métodos de epPCR están limitados por la degeneración del código genético, lo que hace que a menudo un cambio de nucleótido no dé lugar a cambio de aminoácido o bien que genere una mutación conservativa, esto es, el cambio de un aminoácido por otro de características bioquímicas similares. Como consecuencia, la sustitución de un nucleótido da lugar, en promedio, a sólo 5,7 cambios aminoacídicos frente a los 19 posibles<sup>73</sup>.

**b) Mutagénesis saturada:** los métodos mutagénicos actualmente disponibles presentan la desventaja de que su uso impide la exploración de una parte importante del espacio proteico. Para paliar esta limitación se puede hacer uso de la mutagénesis saturada, estrategia consistente en el cambio de un único codón que codifica para un único aminoácido por todos los codones que codifican para los 20 aminoácidos naturales, con el objetivo de localizar el aminoácido óptimo para la función enzimática que se pretende mejorar. De esta manera se consigue aumentar el número de cambios aminoacídicos respecto a los accesibles por mutagénesis aleatoria convencional<sup>74</sup>, si bien es un método dirigido, no aleatorio y, por tanto, debe ser empleado dentro de un contexto semirracional o híbrido<sup>75,76</sup>.

Puede utilizarse también para mutagenizar varios codones

(mutagénesis saturada combinatorial, MSC), tanto en bloques contiguos como en posiciones separadas, permitiendo la exploración de todas las posibles combinaciones de los aminoácidos de interés, de cara a optimizar sus interacciones para la propiedad proteica objeto de estudio. Aunque el uso combinado de la mutagénesis saturada con métodos de cribado de alto rendimiento (en inglés *high-throughput screening*) ha permitido la mejora de diferentes propiedades enzimáticas como la termoestabilidad, la especificidad de sustrato y la enantioselectividad<sup>75</sup>, nuevamente surgen limitaciones en la exploración, ya que los experimentos de MSC generan librerías con un gran número de mutantes. Si la mutagénesis saturada de las posiciones de interés tiene lugar al nivel de DNA por una “randomización” NNN (donde N representa una mezcla de A, C, T y G), la mutagénesis saturada de un único codón generaría 64 variantes, mientras que la mutagénesis saturada de dos y tres codones produciría 4.096 y 262.144 variantes diferentes, respectivamente. Una estrategia ampliamente utilizada para reducir el número de mutantes es la utilización de los codones NNK, donde K es G o T (tabla 3). Esta restricción da lugar a 32 codones codificantes para los 20 aminoácidos y con un codón parada y sólo 11 codones redundantes, es decir, que codifican para el mismo aminoácido<sup>61</sup>. Aún más eficiente es el denominado “truco de los 22 codones”, basado en el uso de una mezcla especial de tres oligonucleótidos: uno con el codón NDT (12 codones), otro con el codón VHG (9 codones) y el último con el codón TGG<sup>77</sup>. Entre las ventajas de este método se encuentran que la mezcla no contiene codones parada, sólo es redundante para valina (GTT, GTG) y leucina (CTT, CTG), y que para la exploración por MSC de dos y tres posiciones aminoacídicas se necesitarían 1.450 y 31.899 variantes, números muy inferiores a los necesarios en el caso de usar codones NNN y NNK (tabla 3).

**Tabla 3.** Tipos de codones degenerados utilizados en mutagénesis saturada. La diversidad con NNN es de 64 codones y 20 aminoácidos, con NNK es de 32 codones y 20 aminoácidos y con el truco de los 22 codones (NDT/VHG/TGG) es de 22 codones y 20 aminoácidos.

Número de codones saturados simultáneamente	Número de codones (diversidad)	NNN		NNK		Truco de los 22 codones (NDT/VHG/TGG)	
		Clones para cubrir el 95% de la librería*	Número de codones (diversidad)	Clones para cubrir el 95% de la librería*	Número de codones (diversidad)	Clones para cubrir el 95% de la librería*	
1	64	192	32	96	22	66	
2	4096	12.288	1024	3072	484	1452	
3	262.144	768.432	32.768	98.304	10.648	31.944	
4	16.777.216	50.331.648	1.048.576	3.145.728	234.256	702.768	

\*Número de clones que hay que cribar para cubrir el 95% del total de posibles variantes en la librería.

**N** = adenina (A), timina (T), citosina (C) o guanina (G).

**K** = guanina (G) o timina (T).

**D** = adenina (A), timina (T) o guanina (G).

**V** = adenina (A), citosina (C) o guanina (G).

**H** = adenina (A), timina (T) o citosina (C).

## Métodos de recombinación del ADN

El intercambio de material genético mediante recombinación durante los procesos de reproducción sexual constituye unas de las bases de la evolución de las especies, puesto que da lugar al aumento de la variabilidad genética de una población dada y a la eliminación del material genético de las mutaciones perjudiciales, traduciéndose en una mejor adaptación de los individuos a los cambios del entorno<sup>78</sup>.

En la actualidad, existen numerosas técnicas que reproducen en el laboratorio los procesos de recombinación del ADN. En términos generales, la estrategia de recombinación de ADN es complementaria y sinérgica con la mutagénesis aleatoria, al facilitar la incorporación en un mismo gen de las mutaciones beneficiosas de diferentes genes parentales así como la eliminación de mutaciones neutras por cruzamiento de los genes descendientes con el gen parental (retrocruzamiento o *back-crossing*)<sup>59</sup>.

Los métodos de recombinación utilizados en evolución dirigida se clasifican en dos grupos, dependiendo de si la recombinación tiene lugar mediante reacciones de PCR (recombinación *in vitro*) o en el interior de un microorganismo hospedador (recombinación *in vivo*).

**a) Métodos de recombinación de ADN *in vitro*:** en 1994 el investigador y empresario neerlandés Willem P. C. Stemmer describió la primera metodología de recombinación de ADN *in vitro*: el barajado *in vitro* del DNA (en inglés, *in vitro DNA shuffling*). Esta técnica consiste en la fragmentación aleatoria de los genes que se desea recombinar con una DNAsa, seguida de una reacción de PCR en ausencia de oligonucleótidos cebadores. Los fragmentos anillan entre ellos debido al alto grado de homología y los ciclos sucesivos de anillamiento y extensión dan lugar a genes enteros cuyas secuencias son recombinaciones de las de los parentales. Estos genes enteros se utilizan como molde en una segunda reacción de PCR de cara a obtener suficiente cantidad para llevar a cabo el clonaje<sup>79</sup>. Stemmer demostró que la recombinación de los genes mediante barajado *in vitro* puede dar lugar a enzimas aún más eficientes. Tal fue la importancia de su contribución que en 2011 recibió, junto con la Prof. Arnold, el Premio Charles Stark Draper, uno de los máximos galardones en el campo de la ingeniería.

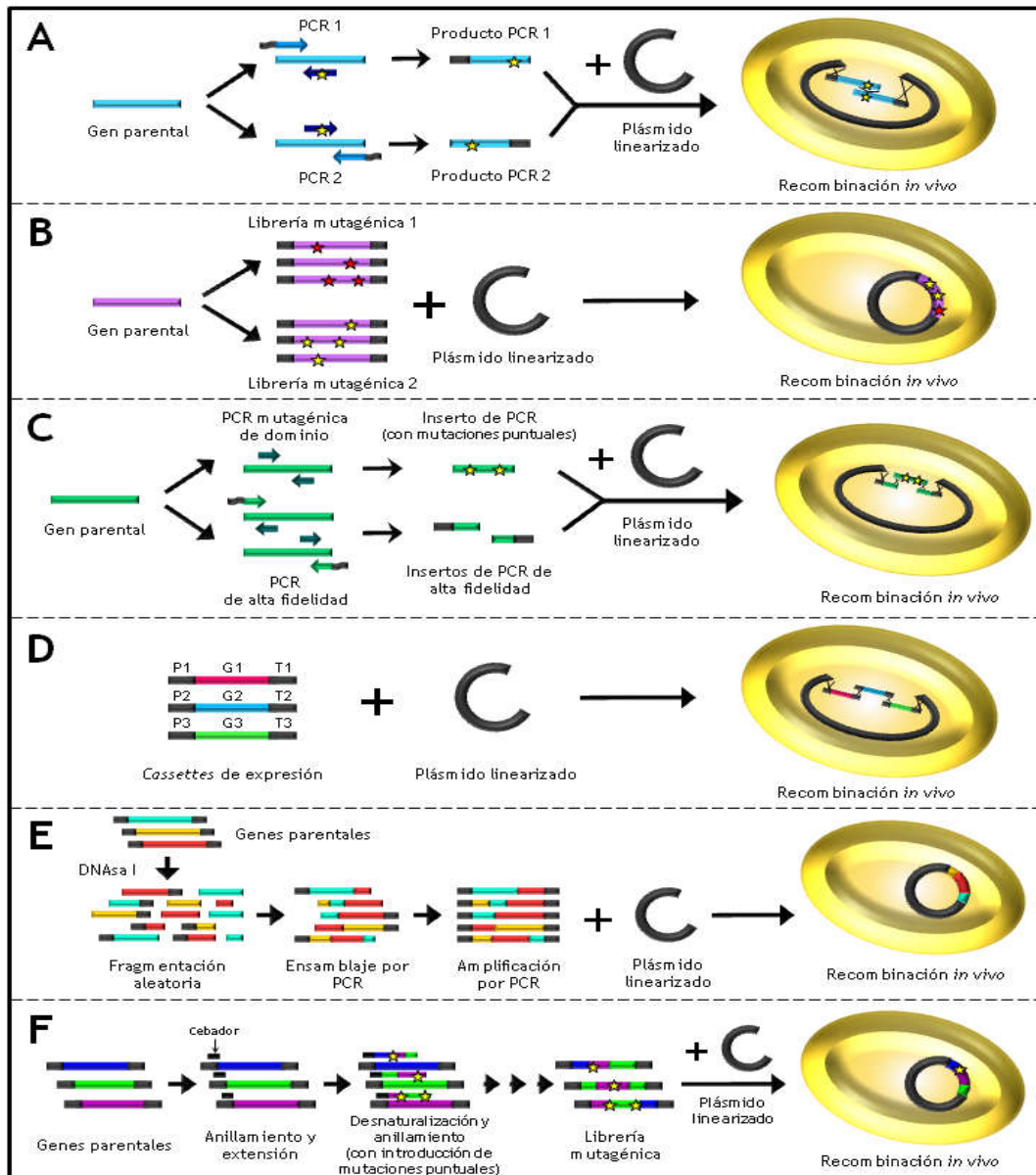
El éxito del *in vitro DNA shuffling* animó a la comunidad científica a desarrollar metodologías recombinantes diferentes, entre las que se pueden destacar: i) la recombinación homóloga mediante StEP (*Staggered Extension Process*)<sup>80</sup> y RACHITT (*Random CHimeragenesis on Transient Templates*)<sup>81</sup>; ii) la recombinación de genes de una misma familia o de genes relacionados con una homología de secuencia superior al 60% mediante DOGS (*Degenerate Oligonucleotide Gene Shuffling*)<sup>82</sup>; iii) la recombinación de genes con una homología de secuencia inferior al 60% a través del ITCHY

(*Incremental Truncation for the Creation of HYbrid enzymes*)<sup>83</sup>, del THIO-ITCHY (derivado del anterior, utiliza nucleótidos  $\alpha$ -fosforilados)<sup>84</sup>, del SHIPREC (*Sequence Homology-Independent Protein RECombination*)<sup>85</sup> y del barajado *in vitro* de exones (*in vitro exon shuffling*)<sup>86</sup> y iv) el uso combinado del ITCHY con el barajado *in vitro* del DNA a través del SCRATCHY<sup>87</sup> y del *enhanced crossover* SCRATCHY (derivado del anterior, permite obtener un mayor número de clones con múltiples entrecruzamientos<sup>88</sup>).

**b) Métodos de recombinación *in vivo*: *S. cerevisiae* como herramienta para la creación de diversidad.** Las técnicas de recombinación *in vitro* del ADN requieren de reacciones de PCR adicionales para la amplificación de los genes completos y para la ligación de estos genes en el correspondiente vector de expresión, lo que a menudo genera la introducción de mutaciones no deseadas. Esto puede evitarse haciendo uso de *S. cerevisiae*, cuyo mecanismo de replicación de alta fidelidad permite la recombinación no mutagénica de los genes parentales, proceso conocido como barajado *in vivo* del ADN o *in vivo DNA shuffling* en inglés<sup>89</sup>. Esta levadura presenta además una alta frecuencia de recombinación homóloga habiéndose demostrado entrecruzamientos de fragmentos que tienen entre 20 y 200 pb de homología<sup>90</sup>. La longitud de los extremos solapantes en la región de entrecruzamiento entre el gen de interés y el plásmido linearizado es un factor clave que determina la eficiencia de la transformación. Así, regiones de homología con al menos 40 pb dan lugar a eficiencias de recombinación de alrededor del 70%<sup>91</sup>.

Estas propiedades convierten a *S. cerevisiae* en una herramienta ideal en evolución dirigida ya que permite la clonación y la recombinación *in vivo* de genotecas mutagénicas con el plásmido linearizado<sup>56</sup>. Para ello, el único requisito es la presencia de 30-40 pb de homología entre los extremos del gen de interés (o de los diferentes genes que se quieren recombinar) y del vector de expresión linearizado<sup>92</sup>. El método IVOE o extensión por solapamiento *in vivo* (*In Vivo Overlap Extension*) permite crear librerías de mutantes para estudios de mutagénesis saturada (tanto simple como combinatorial), mutagénesis dirigida, mutagénesis de inserción y de delección. En el IVOE se realizan dos o más reacciones de PCR con oligonucleótidos degenerados en la posición/posiciones de interés empleando cebadores que contienen secuencias homólogas con el plásmido linearizado, dando lugar a fragmentos de PCR homólogos entre sí y homólogos con respecto al vector linearizado.

Finalmente se consigue la recombinación, ligación y clonaje en un único paso de transformación (figura 3A)<sup>59</sup>. El método IvAM (*In vivo Assembly of Mutant libraries with different mutational spectra*) consiste en la co-transformación en la levadura del plásmido linearizado con dos o más genotecas de mutantes obtenidas con DNA polimerasas que presentan



**Fig. 3:** Figura 3. Selección de métodos empleados en evolución dirigida y biología sintética para la creación de diversidad genética *in vivo* haciendo uso de *Saccharomyces cerevisiae*. (A) IVOE. (B) IvAM. (C) MORPHING. (D) DNA assembler. (E) CLERY. (F) StEP mutagénico combinado con barajado *in vivo* del ADN. Las mutaciones están representadas como estrellas. Los promotores, genes y terminadores que componen los diferentes *cassettes* de expresión de la ruta metabólica están denominados como P, G y T, respectivamente.

distintas predisposiciones mutacionales (diferentes relaciones transición/transversión), lo que facilita la adquisición de una mayor riqueza mutacional en la librería (figura 3B)<sup>93</sup>. Por su parte, el MORPHING (*Mutagenic Organized Recombination Process by Homologous IN vivo Grouping*) es un método que combina la epPCR en ciertos dominios con la recombinación *in vivo* intrínseca de *Saccharomyces*, permitiendo enfocar la mutagénesis en ciertos dominios (fragmentos génicos de ~90 pb) mientras que el resto del gen permanece libre de mutaciones (figura 3C)<sup>94</sup>.

En el contexto de la biología sintética, se está realizando un exhaustivo empleo de *S. cerevisiae* para el ensamblaje *in vivo* de rutas metabólicas mediante diversas técnicas. Un ejemplo

lo constituye el *DNA assembler*, el cual está basado en el auto-ensamblaje en la levadura de los diferentes *cassettes* que componen la ruta metabólica, expresando de manera funcional sus productos génicos (figura 3D)<sup>95</sup>.

Por último, también se han desarrollado métodos que combinan las técnicas de recombinación *in vitro* e *in vivo*. Por ejemplo, el método CLERY (*Combinatorial Libraries Enhanced by Recombination in Yeast*) consiste en pasos consecutivos de *in vitro* e *in vivo* *DNA shuffling* (figura 3E)<sup>96</sup>. De manera similar también se ha empleado la combinación de StEP mutagénico con barajado *in vivo* del ADN para la evolución dirigida de peroxidases (figura 3F)<sup>30</sup>.



## Exploración de las librerías de mutantes

Junto con la generación de diversidad en el gen de interés y la consiguiente expresión en un hospedador adecuado, el éxito de cualquier experimento de evolución dirigida depende del desarrollo de un método de *screening* robusto para la propiedad que se desea mejorar. El conjunto de técnicas disponibles para la exploración de librerías de variantes enzimáticas se puede clasificar en métodos de selección genética, ensayos de cribado alto rendimiento (*high throughput screening*) y métodos de cribado de ultra alto rendimiento (*ultra-high throughput screening*).

### Métodos de selección genética

Los métodos de selección genética son aquellos en los que la función enzimática objeto de estudio está vinculada a la supervivencia o velocidad de crecimiento del microorganismo hospedador. De esta manera, sólo aquellos clones que producen una enzima mejorada sobreviven o crecen más rápidamente. Los métodos de selección son compatibles con librerías muy grandes ( $\sim 10^9$  clones) y no requieren instrumentación especial<sup>97</sup>. No obstante, suele tratarse de métodos muy específicos y difíciles de implementar lo que, combinado con el hecho de que la propiedad enzimática de interés no debe interferir en el complejo metabolismo celular y que debe ser distinguible del resto de reacciones celulares, hace que su uso en experimentos de evolución *in vitro* esté reducido a la ingeniería de enzimas destoxicantes o enzimas que sintetizan nutrientes esenciales para el crecimiento y la supervivencia celular<sup>97,98</sup>.

### Ensayos de cribado de alto rendimiento

En consecuencia, el interés de los investigadores se centra en el desarrollo de protocolos de cribado o *screening* en inglés, en los que cada miembro de la librería es analizado individualmente para la función objeto de estudio, a diferencia de los métodos de selección en los que sólo se estudian los clones supervivientes. El *screening* debe asegurar la detección de las pequeñas mejoras esperadas a partir de la sustitución de un único aminoácido. Con este fin, las condiciones del ensayo se ajustan de forma que la propiedad de la enzima nativa o parental está cerca (y nunca por debajo) del límite de detección.

El *screening* de la propiedad enzimática se lleva a cabo, de forma general, mediante la detección de absorbancia o emisión de fluorescencia. Algunos métodos de *screening* se realizan en fase sólida (placas de agar), donde los cambios de color se detectan directamente por visualización directa (aparición de halos en torno a las colonias o de color en el interior del hospedador) o por análisis de imágenes digitales. Los métodos en fase sólida requieren que el sustrato pueda ser aportado como parte del medio de crecimiento y han de ser lo suficientemente sensibles como para que las colonias que expresan variantes con propiedades mejoradas muestren diferencias de color o fluorescencia con respecto a la enzima parenteral o a los clones inactivos<sup>99</sup>.



**Fig. 4:** Ejemplos de herramientas de cribado de alto rendimiento utilizadas en experimentos de evolución dirigida. (A) Robot picador de colonias. (B) Robot manipulador de líquidos. (C) Robot dispensador de líquidos. (D) Lector de microplacas. (E) Incubador de placas. (F) Citómetro de flujo.

Sin embargo, la mayoría de los protocolos de *screening* se llevan a cabo en fase líquida haciendo uso de placas multipocillo. Estos métodos permiten un rendimiento de  $10^3$ - $10^4$  variantes por ciclo y consisten en la transferencia de colonias individuales a placas multipocillo que contienen medio de cultivo, el crecimiento de las células hasta alcanzar la fase estacionaria y la inducción de la proteína de interés<sup>100</sup>. La optimización y validación de los métodos de *screening* en placas multipocillo se lleva a cabo mediante la evaluación de la linealidad del ensayo y la determinación del coeficiente de varianza (CV). Típicamente, un CV inferior al 15% se considera adecuado para experimentos de evolución dirigida. Además, el uso de equipamiento especial como lectores de microplacas y robots picadores de colonias o manipuladores de líquidos, ayuda a conseguir valores óptimos de CV evitando la detección de falsos positivos<sup>101</sup> (figura 4A-E).

### Métodos de cribado de ultra alto rendimiento

El siguiente nivel en rendimiento de cribado son los ensayos basados en la clasificación de células por activadas por fluorescencia (FACS, por sus siglas en inglés *fluorescence activated cell sorting*). Estos métodos requieren el disponer de citómetros, instrumentos que permiten determinar la fluorescencia de cada una de las partículas (por ejemplo, células, liposomas o gotas de emulsión) a velocidades extremadamente altas ( $\sim 10^7$  por hora), y aislar aquellas que muestran la señal más

alta para la propiedad deseada<sup>102,103</sup> (figura 4F). El principal inconveniente de estos métodos es que no son aplicables de manera general dado que precisan el vínculo entre el genotipo (un ácido nucleico) y el fenotipo (una propiedad observable, por ejemplo, la actividad catalítica) en un contexto de ultra alto rendimiento<sup>104</sup>. En otras palabras, las variantes enzimáticas activas deben modificar sólo las partículas en las que son producidas, algo que puede conseguirse mediante la acumulación de un producto fluorescente en la partícula. Las técnicas de *screening* basadas en FACS se pueden clasificar en cinco categorías en función del método empleado para unir genotipo y fenotipo: a) *in vivo display*, b) *in vitro display*, c) métodos basados en la proteína verde fluorescente (*green fluorescent protein*, GFP), d) atrapamiento del producto dentro de las células, y e) compartimentalización *in vitro*.

a) **Presentación de enzimas en la superficie celular (*cell surface display* o *in vivo display*)**. Esta técnica implica la unión de enzimas heterólogas a la superficie de bacterias<sup>105</sup>, levaduras<sup>106</sup> o fagos<sup>107</sup>. El *cell display* ofrece una capacidad de cribado de hasta  $10^9$  clones por ciclo de evolución y permite el contacto directo de la enzima con el sustrato, por lo que no es necesario que éste difunda a través de la membrana para llegar al citoplasma. No obstante, esta accesibilidad es también el principal inconveniente ya que el producto puede difundir fácilmente.

b) ***In vitro display***. Este método se basa en la síntesis de la proteína de interés mediante un sistema de transcripción-traducción *in vitro* de modo que el vínculo genotipo-fenotipo se consigue usando mRNA (*mRNA display*)<sup>108</sup>, ribosomas (*ribosome display*)<sup>109</sup>, liposomas (*liposome display*)<sup>110</sup>, partículas (*bead surface display*)<sup>111</sup> o la proteína iniciadora de la replicación del ADN RepA (*CIS display*)<sup>112</sup>.

c) **Métodos basados en la GFP**. Debido a sus propiedades auto-fluorescentes, la GFP y sus variantes son candidatos ideales para el análisis de la fluorescencia de las células y, por tanto, se pueden utilizar en *screenings* por citometría de flujo<sup>102</sup>. Las principales ventajas de estos métodos son que su capacidad de rendimiento es muy elevada ( $\sim 10^7$ - $10^8$  clones por ciclo) y que no necesitan sustratos fluorescentes, los cuales a menudo tienen alto coste. Sin embargo, la función enzimática de interés tiene que estar unida a la expresión de la GFP, un requisito que no se cumple en todos los casos.

d) **Atrapamiento del producto dentro de las células**. Esta estrategia debe superar la limitación de que el sustrato pueda acceder al citoplasma y que el producto fluorescente difunda dentro de la célula<sup>102</sup>. Esto puede conseguirse si la enzima cataliza reacciones en las cuales se modifica el tamaño, la polaridad u otras propiedades químicas del sustrato, dando lugar a la retención del producto dentro de la célula. La dificultad del atrapamiento de sustrato/producto en la célula ha podido solucionarse mediante el desarrollo de técnicas de micro-encapsulación que imitan la compartimentación natural

de la célula, por ejemplo, incluyendo la enzima objetivo en las gotitas de micro-emulsión junto con el sustrato y los productos<sup>104</sup>. Las principales ventajas del enfoque de célula en gota son su capacidad de ultra alto rendimiento ( $\sim 10^9$  clones por ciclo) y el pequeño volumen de las micro-emulsiones ( $\sim 5$  femtolitros), lo que reduce de manera significativa la cantidad de sustrato<sup>113</sup>. Más recientemente, se han desarrollado nuevas plataformas de cribado que utilizan células encapsuladas con polímeros tales como: i) el método de encapsulación, solubilización y cribado de alto rendimiento celular (*cellular high-throughput encapsulation, solubilization and screening*, CHESS), en el que las bacterias son encapsuladas en polímeros resistentes a detergentes y que permite el cribado de bibliotecas de  $\sim 10^8$  variantes<sup>114</sup>; y ii) el método de cáscara de polímero (*fur-shell*), basado en la formación de una capa de hidrogel fluorescente alrededor de las células de *E. coli*<sup>103</sup>.

e) **Compartimentalización *in vitro* (CIV)**. La microencapsulación se puede también llevar a cabo en ausencia de células, produciendo directamente la enzima a partir del gen mediante protocolos de transcripción-traducción *in vitro*. La CIV ofrece la ventaja de generar un gran número de gotas por volumen de emulsión ( $>10^{10}$  en 1 mL de emulsión), junto con la facilidad en la preparación de las emulsiones y su elevada estabilidad frente a cambios de temperatura, pH y concentraciones de sal<sup>113</sup>. Además, la CIV reduce el tiempo experimental al eliminar los pasos de crecimiento bacteriano asociados al clonaje, la transformación y la producción de proteínas<sup>115</sup>.

En todos los métodos anteriores un único ciclo de evolución dirigida implica la introducción de mutaciones en el gen de interés, su expresión y la detección o selección de mutantes, lo cual requiere generalmente semanas o meses con un seguimiento riguroso por parte del investigador. Este proceso se puede acelerar haciendo uso de la técnica denominada plataforma de evolución continua asistida por fagos (PACE, del inglés *phage-assisted continuous evolution*), la cual permite una evolución extremadamente rápida de biomoléculas sin apenas intervención por parte del investigador<sup>14</sup>. El PACE se basa en un recipiente de volumen fijo (una "laguna") donde un fago en replicación continua porta el gen de interés. Las células hospedadoras de *E. coli* libres de fagos se suministran constantemente a la laguna y como la capacidad de infección del fago está vinculada a la función del gen objetivo, los fagos que codifican mutantes inactivos producen células huésped no infecciosas que se diluyen fuera de la laguna. Esta estrategia permite llevar a cabo cientos de rondas de evolución en poco más de una semana.

#### *Diseño de estrategias evolutivas*

Lejos de lo trivial que pueda parecer un experimento de evolución dirigida, el éxito en este tipo de aproximaciones es directamente proporcional al planteamiento de la estrategia a emplear. La evolución dirigida no supone la mera introducción aleatoria de mutaciones o sus posibles recombinaciones,

y en muchos casos esta corriente de pensamiento puede conducir a frustraciones científicas, sobre todo a la hora de abordar metas complejas como, por ejemplo, la generación de nuevas actividades o la superación de fuertes efectos inhibitorios intrínsecos al mecanismo catalítico.

Esto se debe en parte a la infinidad de posibilidades que existen al mutar una proteína: en un polipéptido virtual de 300 aminoácidos, se generarían 5700 posibles variantes a partir de la sustitución de un único residuo mientras que se llegarían a producir más de 16 millones de variantes en el caso de abordar la mutación en dos posiciones diferentes<sup>18</sup>. En la situación ideal de disponer de las herramientas moleculares, genéticas, robóticas y computacionales adecuadas, la exploración completa de este espacio de secuencia proteica daría lugar a 20<sup>300</sup> variantes que, como subraya la Prof. Arnold en muchos de sus trabajos, es un número de combinaciones que incluso supera el número de átomos existentes en el universo<sup>18,116</sup>. Sin embargo, lo más paradójico es que este espacio de secuencia proteica multidimensional se encuentra carente de función en la mayor parte de su contenido. Un reflejo experimental de esta aseveración se puede extraer del análisis mutacional llevado a cabo en diversos experimentos de evolución *in vitro*: de manera ordinaria, en torno al 30-50% de las mutaciones introducidas de manera aleatoria son perjudiciales, un 50-70% son neutras (no aportan función) y menos del 0.5-0.01% son beneficiosas<sup>10</sup>.

Por este motivo, resulta obvio pensar que el descubrimiento de mutaciones beneficiosas es un evento poco frecuente, y la combinación de éstas, algo excepcional. Es más, de estas mutaciones beneficiosas, sólo algunas se localizan en zonas o dominios que a priori podrían anticiparse de forma racional, mientras que muchas otras aparecen en zonas muy alejadas pero que, de alguna manera, afectan a la catálisis y función enzimática a través de suaves modificaciones en la geometría global proteica y en sus propiedades electrostáticas, y de manera particular en la dinámica de los aminoácidos del centro activo. El efecto combinado de estos eventos debe influir de manera precisa en el curso de la reacción, particularmente tras la formación del complejo intermediario de alta energía. Por último, no se deben obviar las interacciones epistáticas que tienen lugar cuando la presencia de una mutación influye sobre la contribución de otra en la función enzimática. Las formas extremas de epistasia se manifiestan como efectos epistáticos negativos, en los que la combinación de determinadas mutaciones puede resultar beneficiosa, aunque de manera independiente al menos una de estas mutaciones no lo sea.

En este complejo escenario de trabajo, queda probada la importancia que tiene un diseño evolutivo apropiado para poder conseguir la función deseada. En este sentido, la tendencia más exitosa es el denominado camino adaptativo (*simple uphill walk* en inglés), que defiende la introducción exclusiva

de mutaciones beneficiosas de manera aleatoria, limitada y precisa, es decir, una por una, generación tras generación<sup>10</sup>. Empleando esta aproximación como punto de partida existen diferentes variaciones, todas ellas válidas y complementarias, para el diseño de nuevas actividades o funciones, a saber:

i) **Incremento gradual de función.** Esta opción se fundamenta en incrementar de manera progresiva la presión selectiva para alcanzar la actividad/función deseada, por ejemplo, mediante el empleo de sustratos intermediarios que pueden conducir finalmente a que la enzima trabaje eficientemente sobre el sustrato diana<sup>117</sup>.

ii) **Diseño de enzimas generalistas y re-especialización enzimática.** Las enzimas conocidas – con excepción de las enzimas sintéticas recientemente “resucitadas” de origen pre-cámbrico<sup>118</sup> – son proteínas altamente especializadas debido a años de evolución natural en estrictos ambientes metabólicos. Para generar una nueva función, la transformación de una enzima especialista en generalista (es decir, con una especificidad de sustrato más amplia que incluya una actividad marginal hacia el sustrato deseado) puede suponer una gran ventaja. Una vez ampliada su especificidad de sustrato, la actividad marginal creada se potencia, re-especializando la enzima en la dirección planeada. Cuando se obtiene la actividad en cuestión, puede ser interesante suprimir actividades laterales a través de la aplicación de procesos de *screening* o selección negativos y positivos para las actividades a anular y potenciar, respectivamente<sup>117</sup>. En este marco de trabajo puede resultar muy adecuado hacer uso de la deriva genética. Esta estrategia se basa en la acumulación de mutaciones neutras que no pongan en peligro la función proteica y que pueden conducir a un nuevo entorno evolutivo en el que se generan nuevas actividades catalíticas mediante dos vías diferentes:

a) el incremento de la estabilidad proteica que facilita la tolerancia hacia nuevas mutaciones beneficiosas pero desestabilizantes, y b) el aumento de la promiscuidad en la función enzimática generando nuevos puntos de partida para actividades no identificadas en la enzima original<sup>10,119</sup>.

iii) **Evolución dirigida sorteando “vías muertas”.** En el camino adaptativo hacia la función enzimática, es frecuente encontrarse numerosas “vías muertas” o *dead ends*, donde o bien la función se suprime o la enzima se desestabiliza siendo incapaz de tolerar nuevas mutaciones. Estas vías muertas pueden esquivarse mediante la incorporación de mutaciones estabilizantes o la exploración de rutas evolutivas múltiples y paralelas<sup>117</sup>.

iv) **Evolución dirigida enfocada.** Consiste en modificar posiciones concretas mediante mutagénesis dirigida y/o saturada en un enfoque híbrido o semirracional. Dentro de esta alternativa también se podría incluir la denominada evolución dirigida de dominios, donde se exploran librerías mutagénicas aleatorias en dominios muy específicos mediante diversas

estrategias<sup>59,61,70,120</sup>. Esta aproximación, aunque muy válida para problemas concretos, por ejemplo, el rediseño de bolsillos catalíticos o la superación de problemas mecanísticos, se encuentra impedida por el limitado conocimiento de la función proteica.

En base a la propia experiencia y a los resultados publicados por los numerosos investigadores que se dedican a la evolución dirigida, se puede asegurar, sin temor a equivocarse, que no existe un camino exclusivo para el diseño de una propiedad enzimática y que, por lo general, lo más conveniente sea la combinación de diversos enfoques en función de la proteína y del objetivo que se plantee.

## Conclusiones

La evolución dirigida es una metodología que nos permite acelerar el proceso de la evolución natural un millón de veces a nivel genético y, de ese modo, aprovechar el potencial sintético de las enzimas naturales, adaptando sus propiedades a nuestras necesidades en un período de tiempo que abarca muchas veces menos de un año. Desde su aparición hace casi tres décadas, la evolución *in vitro* se convirtió rápidamente en una herramienta de uso común para la ingeniería de biocatalizadores, anticuerpos y proteínas terapéuticas. Es tal su aplicabilidad, que la mayoría de nosotros usamos a diario numerosos productos obtenidos por ingeniería de enzimas (por ejemplo, pan, biocombustibles o compuestos quirales para la síntesis de fármacos y anticuerpos) o que contienen enzimas diseñadas por evolución (por ejemplo, detergentes o alimentos para animales).

Además, recientemente han surgido nuevos campos de investigación que permiten el diseño de enzimas para catalizar reacciones desconocidas en la naturaleza<sup>121</sup>. Ejemplos de esto son la evolución de enzimas para la formación de enlaces boro-carbono<sup>122</sup> o la catálisis híbrida, esto es, la convergencia con la química inorgánica para explorar el potencial sintético de catalizadores metálicos<sup>123</sup>.

No obstante, somos muchos los que pensamos que aún estamos en las primeras etapas de la revolución de la evolución dirigida y que, de aquí a unos años, el rango de aplicación de los biocatalizadores obtenidos por este método será aún más extenso.

## Agradecimientos

El 2018 quedará grabado para siempre en la memoria de todos los que nos dedicamos a la evolución dirigida y, especialmente, en la de aquellos que han formado parte del grupo de Frances H. Arnold o que hemos aprendido directamente de discípulos suyos.

Es por ello que me gustaría agradecer de corazón a mi director de tesis y maestro, al Dr. Miguel Alcalde del Instituto de Catálisis y Petroleoquímica, su inmenso apoyo a lo largo ya de diez años, así como su dedicación a la hora de enseñarme todo lo relacionado con la evolución dirigida tal y como él lo aprendió de su mentora Ar-

nold. Quisiera extender mi agradecimiento a otro discípulo de Arnold y mi jefe durante mi postdoctorado en Alemania, el Prof. Ulrich Schwaneberg de la Universidad Técnica de Aquisgrán, por las enriquecedoras horas de debate y *brainstorming*, y por apostar por mí para iniciar la línea de evolución dirigida de sortasas en su grupo.

Asimismo, agradecer al Dr. Aurelio Hidalgo y al Prof. José Berenguer de la Universidad Autónoma de Madrid, su calurosa acogida en su grupo de investigación y por iniciarme en el fascinante mundo de la microfluídica y de las bacterias termófilas.

También dar las gracias a mi compañero Jorge Bravo por la lectura crítica de este artículo y por su apoyo durante tantas horas en el laboratorio.

Por supuesto, gracias a la Consejería de Educación, Juventud y Deporte de la Comunidad de Madrid por concederme la Ayuda de Atracción de Talento Investigador que me permite continuar mi investigación y combinarla con tareas docentes en la Universidad Autónoma de Madrid.

Y como dicen que lo mejor se deja siempre para el final, quisiera dedicar este artículo a mis padres, a cuya recombinación genética debo mi existencia y que son sin lugar a dudas mi mayor apoyo y mi referente en este camino adaptativo al que llamamos vida.

## Bibliografía

1. K Chen, FH Arnold FH. (1991) Enzyme engineering for non-aqueous solvents: Random mutagenesis to enhance activity of subtilisin E in polar organic media. **Biotechnology**, **9**, 1073-1077 (1991).
2. GP Smith. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. **Science**, **228**, 1315-1317 (1985).
3. G Winter. Synthetic human antibodies and a strategy for protein engineering. **FEBS Lett.**, **430**, 92-94 (1998).
4. JR Cherry, AL Fidantsef. Directed evolution of industrial enzymes: an update. **Curr. Opin. Biotech.**, **14**, 438-443 (2003).
5. JD Bloom, MM Meyer, P Meinhold, CR Otey, D MacMillan, FH Arnold. Evolving strategies for enzyme engineering. **Curr. Opin. Struct. Biol.**, **15**, 447-452 (2005).
6. MT Reetz, D Kahakeaw, R Lohmer. Addressing the numbers problem in directed evolution. **ChemBioChem**, **9**, 1797-1804 (2008).
7. P Braiuca, C Ebert, A Basso, P Linda, L Gardossi. Computational methods to rationalize experimental strategies in biocatalysis. **Trends Biotechnol.**, **24**, 419-425 (2006).
8. D Floudas, M Binder, R Riley, K Barry, RA Blanchette, B Henrissat, AT Martinez, R Otilar, JW Spatafora, JS Yadav *et al.* The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. **Science**, **336**, 1715-1719 (2012).
9. S Bershtein, DS Tawfik. Advances in laboratory evolution of enzymes. **Curr. Opin. Chem. Biol.**, **12**, 151-158 (2008).

10. JD Bloom, FH Arnold. In the light of directed evolution: Pathways of adaptive protein evolution. **P. Natl. Acad. Sci. USA**, **106**, 9995-10000 (2009).
11. S Brakmann. Discovery of superior enzymes by directed molecular evolution. *ChemBioChem*, **2**, 865-871 (2001).
12. FH Arnold. How proteins adapt: lessons from directed evolution. *Cold Spring Harb. Symp. Quant. Biol.*, **74**, 41-46 (2009).
13. PA Dalby. Strategy and success for the directed evolution of enzymes. *Curr. Opin. Struc. Biol.*, **21**, 473-480 (2011).
14. KM Esvelt, JC Carlson, DR Liu. A system for the continuous directed evolution of biomolecules. **Nature**, **472**, 499-503 (2011).
15. M Alcalde. Fundamentos de evolución molecular dirigida de enzimas. En: *Monografías de la Real Academia Nacional de Farmacia, Monografía XXXV: Biocatálisis aplicada a la obtención de fármacos y productos de alto valor añadido*. Eds. JM Sánchez Montero, F Ortega Ortiz de Apodaca y A Doadrio Villarejo. Real Academia Nacional de Farmacia (Madrid). pp. 289-316 (2012).
16. C Jäckel, D Hilvert. Biocatalysts by evolution. **Curr. Opin. Biotech.**, **21**, 753-759 (2010).
17. M Wang, T Si, H Zhao. Biocatalyst development by directed evolution. *Bioresource Technol.*, **115**, 117-125 (2012).
18. FH Arnold. Directed evolution: Creating biocatalysts for the future. **Chem. Eng. Sci.**, **51**, 5091-5102 (1996).
19. JR Cherry, MH Lamsa, P Schneider, J Vind, A Svendsen, A Jones, AH Pedersen. Directed evolution of a fungal peroxidase. **Nat. Biotechnol.**, **17**, 379-384 (1999).
20. A Miele, P Giardina, E Notomista, A Piscitelli, G Sannia, V Faraco. A semirational approach to engineering laccase enzymes. **Mol. Biotechnol.**, **46**, 149-156 (2010).
21. XJ Feng, J Sanchis, MT Reetz, H Rabitz. Enhancing the efficiency of directed evolution in focused enzyme libraries by the adaptive substituent reordering algorithm. **Chem. Eur. J.**, **18**, 5646-5654 (2012).
22. MT Reetz. Laboratory evolution of stereoselective enzymes: A prolific source of catalysts for asymmetric reactions. **Angew. Chem. Int. Edit.**, **50**, 138-174 (2011).
23. MW Peters, P Meinhold, A Glieder, FH Arnold. Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3. **J. Am. Chem. Soc.**, **125**, 13442-13450 (2003).
24. P Gomez de Santos, M Cañellas, F Tieves, SHH Younes, P Molina-Espeja, M Hofrichter, F Hollmann, V Guallar, M Alcalde. Selective synthesis of the human drug metabolite 5'-hydroxy-propranolol by and evolved self-sufficient peroxygenase. **ACS Catal.**, **8**, 4789-4799 (2018).
25. A Iffland, P Tafelmeyer, C Saudan, K Johnsson. Directed molecular evolution of cytochrome c peroxidase. **Biochemistry**, **39**, 10790-10798 (2000).
26. A Glieder, ET Farinas, FH Arnold. Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. **Nat. Biotechnol.**, **20**, 1135-1139 (2002).
27. T Bulter, M Alcalde, V Sieber, P Meinhold, C Schlachtbauer, FH Arnold. Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. **Appl. Environ. Microbiol.**, **69**, 987-995 (2003).
28. JA Rakestraw, SL Sazinsky, A Piatasi, E Antipov, KD Wittrop. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*. **Biotechnol. Bioeng.**, **103**, 1192-1201 (2009).
29. D Maté, C García-Burgos, E García, A Ballesteros, S Camarero, M Alcalde. Laboratory evolution of high redox potential laccases. **Chem. Biol.**, **17**, 1030-1041 (2010).
30. E Garcia-Ruiz, D Gonzalez-Perez, FJ Ruiz-Dueñas, AT Martinez, M Alcalde. Directed evolution of a temperature-, peroxide- and alkaline pH-tolerant versatile peroxidase. **Biochem. J.**, **441**, 487-498 (2012).
31. HM Zhao, FH Arnold. Directed evolution converts subtilisin E into a functional equivalent of thermitase. **Protein Eng.**, **12**, 47-53 (1999).
32. E García-Ruiz, D Maté, A Ballesteros, AT Martinez, M Alcalde. Evolving thermostability in mutant libraries of ligninolytic oxido-reductases expressed in yeast. **Microb. Cell. Fact.**, **9**, 17 (2010).
33. D Gonzalez-Perez, E Garcia-Ruiz, FJ Ruiz-Dueñas, AT Martinez, M Alcalde. Structural determinants of oxidative stabilization in an evolved versatile peroxidase. **ACS Catal.**, **4**, 3891-3901 (2014).
34. P Torres-Salas, DM Mate, I Ghazi, FJ Plou, AO Ballesteros, M Alcalde. Widening the pH activity profile of a fungal laccase by directed evolution. **ChemBioChem**, **14**, 934-937 (2013).
35. D Gonzalez-Perez, I Mateljck, E Garcia-Ruiz, FJ Ruiz-Dueñas, AT Martinez, A Alcalde. Alkaline versatile peroxidase by directed evolution. **Catal. Sci. Technol.**, **6**, 6625-6636 (2016).
36. C Novoa, GV Dhoke, DM Mate, R Martínez, T Haarmann, M Schreiter, J Eidner, R Schwerdtfeger, P Lorenz, MD Davari, F Jakob, U Schwaneberg. KnowVolution of a fungal laccase toward alkaline pH. **ChemBioChem**, **20**, 1-10 (2019).
37. A Crameri, EA Whitehorn, E Tate, WPC Stemmer. Improved green fluorescent protein by molecular evolution using DNA shuffling. **Nat. Biotechnol.**, **14**, 315-319 (1996).
38. NC Shaner, MZ Lin, MR McKeown, PA Steinbach, KL Hazelwood, MW Davidson, RY Tsien. Improving the photostability of bright monomeric orange and red fluorescent proteins. **Nat. Methods**, **5**, 545-551 (2008).
39. TS Wong, FH Arnold, U Schwaneberg. Laboratory evolution of cytochrome P450 BM-3 monooxygenase for organic cosolvents. **Biotechnol. Bioeng.**, **85**, 351-358 (2004).

40. M Zumárraga, T Bulter, S Shleev, J Polaina, A Martínez-Arias, FJ Plou, A Ballesteros, M Alcalde. *In vitro* evolution of a fungal laccase in high concentrations of organic cosolvents. **Chem. Biol.**, **14**, 1052-1064 (2007).
41. Z Zou, H Alibiglou, DM Mate, MD Davari, F Jakob, U Schwaneberg. Directed sortase A evolution for efficient site-specific bioconjugations in organic solvents. **Chem. Commun.**, **54**, 11467-11470 (2018).
42. H Liu, L Zhu, M Bocola, N Chen, AC Spiess, U Schwaneberg. Directed laccase evolution for improved ionic liquid resistance. **Green Chem.**, **15**, 1348-1355 (2013).
43. JL Carter, M Bekhouche, A Noiriél, LJ Blum, B Doumèche. Directed evolution of a formate dehydrogenase for increased tolerance to ionic liquids reveals a new site for increasing the stability. **ChemBioChem**, **15**, 2710-2718 (2014).
44. DM Mate, D Gonzalez-Perez, M Falk, R Kittl, M Pita, AL De Lacey, R Ludwig, S Shleev, M Alcalde. Blood tolerant laccase by directed evolution. **Chem. Biol.**, **20**, 223-231 (2013).
45. TW Johannes, HM Zhao. Directed evolution of enzymes and biosynthetic pathways. **Curr. Opin. Microbiol.**, **9**, 261-267 (2006).
46. RE Cobb, T Si, H Zhao. Directed evolution: an evolving and enabling synthetic biology tool. **Curr. Opin. Chem. Biol.**, **16**, 285-291 (2012).
47. BS Majors, GG Chiang, MJ Betenbaugh. Protein and genome evolution in mammalian cells for biotechnology applications. **Mol. Biotechnol.**, **42**, 216-223 (2009).
48. M Alcalde. Evolución molecular dirigida. **Investigación y Ciencia**, **326**, 69-77 (2003).
49. MA Romanos, CA Scorer, JJ Clare. Foreign gene expression in yeast - A review. **Yeast**, **8**, 423-488 (1992).
50. A Pourmir, TW Johannes. Directed evolution: selection of the host microorganism. **Comput. Struct. Biotechnol. J.**; **2**, e201209012 (2012).
51. A Hidalgo, J Berenguer. Biotechnological applications of *Thermus thermophilus* as host. **Curr. Biotechnol.**, **2**, 304-312 (2013).
52. Al Ribeiro, M Sánchez, A Hidalgo, J Berenguer. Stabilization of enzymes by using thermophiles. En: *Microbial Steroids. Methods in Molecular Biology*. Eds. JL Barredo, I Herráiz. Humana Press (Nueva York). pp. 297-312 (2017).
53. Novagen. Competent Cells Brochure.
54. JM Cregg, JL Cereghino, JY Shi, DR Higgins. Recombinant protein expression in *Pichia pastoris*. **Mol. Biotechnol.**, **16**, 23-52 (2000).
55. JL Cereghino, JM Cregg. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. **FEMS Microbiol. Rev.**, **24**, 45-66 (2000).
56. D Gonzalez-Perez, E Garcia-Ruiz, M Alcalde. *Saccharomyces cerevisiae* in directed evolution: An efficient tool to improve enzymes. **Bioengineered**, **3**, 172-177 (2012).
57. Z Liu, B Pscheidt, M Avi, R Gaisberger, FS Hartner, C Schuster, W Skranc, K Gruber, A Glieder. Laboratory evolved biocatalysts for stereoselective syntheses of substituted benzaldehyde cyanohydrins. **ChemBioChem**, **9**, 58-61 (2008).
58. T Bulter, V Sieber, M Alcalde. *Screening mutant libraries in Saccharomyces cerevisiae*. En: *Directed Enzyme Evolution: Screening and Selection Methods. Methods in Molecular Biology*. Eds. FH Arnold y G Georgiou. Humana Press (Totowa, Nueva Jersey). pp. 99-107 (2003).
59. M Alcalde. Mutagenesis protocols in *Saccharomyces cerevisiae* by in vivo overlap extension. En: *In vitro Mutagenesis Protocols. Methods in Molecular Biology*. Ed. J. Braman. Humana Press (Totowa, Nueva Jersey). pp. 3-14 (2010).
60. TS Wong, D Roccatano, U Schwaneberg. Steering directed protein evolution: strategies to manage combinatorial complexity of mutant libraries. **Environ. Microbiol.**, **9**, 2645-2659 (2007).
61. TS Wong, D Zhurina, U Schwaneberg. The diversity challenge in directed protein evolution. **Comb. Chem. High T. Scr.**, **9**, 271-288 (2006).
62. AV Shivange, J Marienhagen, H MundhadaH, A Schenk, U Schwaneberg. Advances in generating functional diversity for directed protein evolution. **Curr. Opin. Chem. Biol.**, **13**, 19-25 (2009).
63. JC Moore, FH Arnold. Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. **Nat. Biotechnol.**, **14**, 458-467 (1996).
64. PS Daugherty, G Chen, BL Iverson, G Georgiou. Quantitative analysis of the effect of the mutation frequency on the affinity maturation of single chain Fv antibodies. **P. Natl. Acad. Sci. USA**, **97**, 2029-2034 (2000).
65. J Cline, JC Braman, HH Hogrefe. PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. **Nucleic Acids Res.**, **24**, 3546-3551 (1996).
66. PC Cirino, KM Mayer, D Umeno. Generating mutant libraries using error-prone PCR. En: *Directed Evolution Library Creation. Methods in Molecular Biology*. Eds. FH Arnold y G Georgiou. Humana Press (Totowa, Nueva Jersey). pp. 3-9 (2003).
67. DW Leung, E Chen, DV Goeddel. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. **Technique**, **1**, 11-15 (1989).
68. TS Wong, KL Tee, B Hauer, U Schwaneberg. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. **Nucleic Acids Res.**, **32**, e26 (2004).
69. PH Patel, H Kawate, E Adman, M Ashbach, KL Loeb. A single highly mutable catalytic site amino acid is critical for DNA polymerase fidelity. **J. Biol. Chem.**, **276**, 5044-5051 (2001).
70. J Cline, H Hogrefe. Randomize gene sequences with new PCR mutagenesis kit. **Strategies**, **13**, 157-161 (2000).

71. S Emond, P Mondon, S Pizzut-Serin, L Douchy, F Crozet, K Bouayadi, H Kharrat, G Potocki-Veronese, P Monsan, M Remaud-Simeon. A novel random mutagenesis approach using human mutagenic DNA polymerases to generate enzyme variant libraries. **Protein Eng. Des. Sel.**, **21**, 267-274 (2008).
72. P Mondon, D Grand, N Souyris, S Emond, K Bouayadi, H Kharrat. Mutagen: a random mutagenesis method providing a complementary diversity generated by human error-prone DNA polymerases. En: *In vitro Mutagenesis Protocols. Methods in Molecular Biology*. Ed. J Braman. Humana Press (Totowa, Nueva Jersey). pp. 373-386 (2010).
73. O Kuchner, FH Arnold. Directed evolution of enzyme catalysts. **Trends Biotechnol.**, **15**, 523-530 (1997).
74. R Georgescu, G Bandara, L Sun. Saturation mutagenesis. En: *Directed Evolution Library Creation. Methods in Molecular Biology*. Eds. FH Arnold y G Georgiou. Humana Press (Totowa, Nueva Jersey). pp. 75-83 (2003).
75. RA Chica, N Doucet, JN Pelletier. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design. **Curr. Op. Biotech.**, **16**, 378-384 (2005).
76. S Lutz. Beyond directed evolution - Semi-rational protein engineering and design. **Curr. Op. Biotech.**, **21**, 734-743 (2010).
77. S Kille, GC Acevedo-Roch, LP Parra, ZG Zhang, DJ Opperman, MT Reetz, JP Acevedo. Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis. **ACS Synth. Biol.**, **2**, 83-92 (2013).
78. PD Keightley, A Eyre-Walker. Deleterious mutations and the evolution of sex. **Science**, **290**, 331-333 (2000).
79. WPC Stemmer. Rapid evolution of a protein in vitro by DNA shuffling. **Nature**, **370**, 389-391 (1994).
80. HM Zhao, L Giver, ZX Shao, JA Affholter, FH Arnold. Molecular evolution by staggered extension process (StEP) in vitro recombination. **Nat. Biotechnol.**, **16**, 258-261 (1998).
81. WM Coco, WE Levinson, MJ Crist, HJ Hektor, A Darzins, PT Pienkos, CH Squires, DJ Monticello. DNA shuffling method for generating highly recombined genes and evolved enzymes. **Nat. Biotechnol.**, **19**, 354-359 (2001).
82. MD Gibbs, KMH Nevalainen, PL Bergquist. Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling. **Gene**, **271**, 13-20 (2001).
83. M Ostermeier, JH Shim, SJ Benkovic. A combinatorial approach to hybrid enzymes independent of DNA homology. **Nat. Biotechnol.**, **17**, 1205-1209 (1999).
84. S Lutz, M Ostermeier, SJ Benkovic. Rapid generation of incremental truncation libraries for protein engineering using alpha-phosphothioate nucleotides. **Nucleic Acids Res.**, **29**, e16 (2001).
85. V Sieber, CA Martinez, FH Arnold. Libraries of hybrid proteins from distantly related sequences. **Nat. Biotechnol.**, **19**, 456-460 (2001).
86. JA Kolkman, WPC Stemmer. Directed evolution of proteins by exon shuffling. **Nat. Biotechnol.**, **19**, 423-428 (2001).
87. S Lutz, M Ostermeier, GL Moore, CD Maranas, SJ Benkovic. Creating multiplecrossover DNA libraries independent of sequence identity. **P. Natl. Acad. Sci. USA**, **98**, 11248-11253 (2001).
88. Y Kawarasaki, KE Griswold, JD Stevenson, T Selzer, SJ Benkovic, BL Iverson, G Georgiou. Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. **Nucleic Acids Res.**, **31**, e126 (2003).
89. JSokkels. *In vivo* gene shuffling in yeast: a fast and easy method for directed evolution of enzymes. En: *Enzyme Functionality: Design, Engineering and Screening*. Ed. A Svendsen. CRC Press (Nueva York). pp. 413-424 (2004).
90. DG Gibson. Synthesis of DNA fragments in yeast by one-step assembly of overlapping oligonucleotides. **Nucleic Acids Res.**, **37**, 6984-6990 (2009).
91. KR Oldenburg, KT Vo, S Michaelis, C Paddon. Recombination-mediated PCR directed plasmid construction *in vivo* in yeast. **Nucleic Acids Res.**, **25**, 451-452 (1997).
92. P Manivasakam, SC Weber, J Mcelver, RH Schiestl. Microhomology mediated PCR targeting in *Saccharomyces cerevisiae*. **Nucleic Acids Res.**, **23**, 2799-2800 (1995).
93. M Zumárraga, S Camarero, S Shleev, A Martínez-Arias, A Ballesteros, FJ Plou, M Alcalde. Altering the laccase functionality by *in vivo* assembly of mutant libraries with different mutational spectra. **Proteins**, **71**, 250-260 (2008).
94. D Gonzalez-Perez, P Molina-Espeja, E Garcia-Ruiz, M Alcalde. Mutagenic organized recombination process by homologous *in vivo* grouping (MORPHING) for directed enzyme evolution. **PLoS ONE**, **9**, e90919 (2014).
95. ZY Shao, H Zhao, HM Zhao. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. **Nucleic Acids Res.**, **37**, e16 (2009).
96. V Abécassis, D Pompon, G Truan. High efficiency family shuffling based on multi-step PCR and *in vivo* DNA recombination in yeast: statistical and functional analysis of a combinatorial library between human cytochrome P450 1A1 and 1A2. **Nucleic Acids Res.**, **28**, e88 (2000).
97. H Leemhuis, RM Kelly, L Dijkhuizen. Directed evolution of enzymes: Library screening strategies. **IUBMB Life**, **61**, 222-228 (2009).
98. FH Arnold, AA Volkov. Directed evolution of biocatalysts. **Curr. Opin. Chem. Biol.**, **3**, 54-59 (1999).
99. AV Tobias, JM Joern. Solid-phase screening using digital image analysis. En: *Directed Enzyme Evolution: Screening and Selection Methods. Methods in Molecular Biology*. Eds.

- FH Arnold y G Georgiou. Humana Press (Totowa, Nueva Jersey). pp. 109-115 (2003).
100. O Salazar, L Sun. Evaluating a screen and analysis of mutant libraries. En: *Directed enzyme evolution: screening and selection methods. Methods in molecular biology*. Eds. FH Arnold y G Georgiou. Humana Press (Totowa, Nueva Jersey). pp. 85-97 (2003).
101. DM Mate, D Gonzalez-Perez, I Mateljak, P Gomez de Santos, AI Vicente, M Alcalde. The pocket manual of directed evolution: Tips and tricks. En: *Biotechnology of Microbial Enzymes: Production, Biocatalysis and Industrial Applications*. Ed. G Brahmachari, AL Demain y JL Adrio. Elsevier. 185-214 (2016).
102. G Yang, SG Withers. Ultrahigh-throughput FACS-based screening for directed enzyme evolution. **ChemBioChem**, **10**, 2704-2715 (2009).
103. C Pitzler, G Wirtz, L Vojcic, S Hiltl, A Böker, R Martinez, U Schwaneberg. A fluorescent hydrogel-based flow cytometry high-throughput screening platform for hydrolytic enzymes. **Chem. Biol.**, **21**, 1733-1742 (2014).
104. AD Giffiths, DS Tawfik. Miniaturising the laboratory in emulsion droplets. **Trends Biotechnol.**, **24**, 395-402 (2006).
105. E van Bloois, RT Winter, H Kolmar, MW Fraaije. Decorating microbes: surface display of proteins on Escherichia coli. **Trends Biotechnol.**, **29**, 79-86 (2011).
106. ET Boder, KD Wittrup. Yeast surface display for *screening* combinatorial polypeptides libraries. **Nat. Biotechnol.**, **15**, 553-557 (1997).
107. A Fernandez-Gacio, M Uguen, J Fastrez. Phage display as a tool for the directed evolution of enzymes. **Trends Biotechnol.**, **21**, 408-414 (2003).
108. RW Roberts, JW Szostak. RNA-peptide fusions for the in vitro selection of peptides and proteins. **Proc. Natl. Acad. Sci. USA**, **94**, 12297-12302 (1997).
109. J Hanes, A Plückthun. *In vitro* selection and evolution of functional proteins by using ribosome display. **P. Natl. Acad. Sci. USA**, **94**, 4937-4942 (1997).
110. S Fujii, T Matsumura, T Sunami, T Nishikawa, Y Kazuta, T Yomo. Liposome display for in vitro selection and evolution of membrane proteins. **Nat. Protoc.**, **9**, 1578-1591 (2014).
111. L Diamante, P Gatti-Lafranconi, Y Schaerli, F Hollfelder. *In vitro* affinity screening of protein and peptide binders by megavalent bead surface display. **Protein Eng. Des. Sel.**, **26**, 713-724 (2013).
112. R Odegrip, D Coomber, B Eldridge, R Hederer, PA Kuhlman, C Ullman, K FitzGerald, D McGregor. CIS display: In vitro selection of peptides from libraries of protein/DNA complexes. **P. Natl. Acad. Sci. USA**, **101**, 2806-2810 (2004).
113. A Aharoni, AD Griffiths, DS Tawfik. High-throughput screens and selections of enzyme-encoding genes. **Curr. Opin. Chem. Biol.**, **9**, 210-216 (2005).
114. KJ Yong, DJ Scott. Rapid directed evolution of stabilized proteins with cellular high-throughput encapsulation solubilization and screening (CHESS). **Biotechnol. Bioeng.**, **112**, 438-446 (2015).
115. R Martínez, U Schwaneberg. A roadmap to directed enzyme evolution and screening systems for biotechnological applications. **Biol. Res.**, **46**, 395-405 (2013).
116. FH Arnold. Fancy footwork in the sequence space shuffle. **Nat. Biotechnol.**, **24**, 328-330 (2006).
117. CA Tracewell, FH Arnold. Directed enzyme evolution: climbing fitness peaks one amino acid at a time. **Curr. Opin. Chem. Biol.**, **13**, 3-9 (2009).
118. R Perez-Jimenez, A Ingles-Prieto, ZM Zhao, I Sánchez-Romero, J Alegre-Cebollada, P Kosuri, S Garcia-Manyes, T Kappock, M Tanokura, A Holmgren, JM Sánchez-Ruiz, EA Gaucher, JM Fernandez. Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. **Nat. Struct. Mol. Biol.**, **18**, 592-596 (2011).
119. JD Bloom, PA Romero, Z Lu, FH Arnold. Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. **Biol. Direct**, **2**, 17 (2007).
120. A Hidalgo, A Schliessmann, R Molina, J Hermoso, UT Bornscheuer. A one-pot, simple methodology for cassette randomisation and recombination for focused directed evolution. **Protein Eng. Des. Sel.**, **21**, 567-576 (2008).
121. IUT Bornscheuer, GW Huisman, RJ Kazlauskas, S Lutz, JC Moore, K Robins. Engineering the third wave of biocatalysis. **Nature**, **485**, 185-194 (2012).
122. SBJ Kan, X Huang, Y Gumulya, K Chen, FH Arnold. Genetically programmed chiral organoborane synthesis. **Nature**, **552**, 132 (2017).
123. AR Grimm, DF Sauer, MD Davari, L Zhu, M Bocola, S Kato, A Onoda, T Hayashi, J Okuda, U Schwaneberg. Cavity size engineering of a  $\beta$ -barrel protein generates efficient biohybrid catalysts for olefin metathesis. **ACS Catal.**, **8**, 3358-3364 (2018).