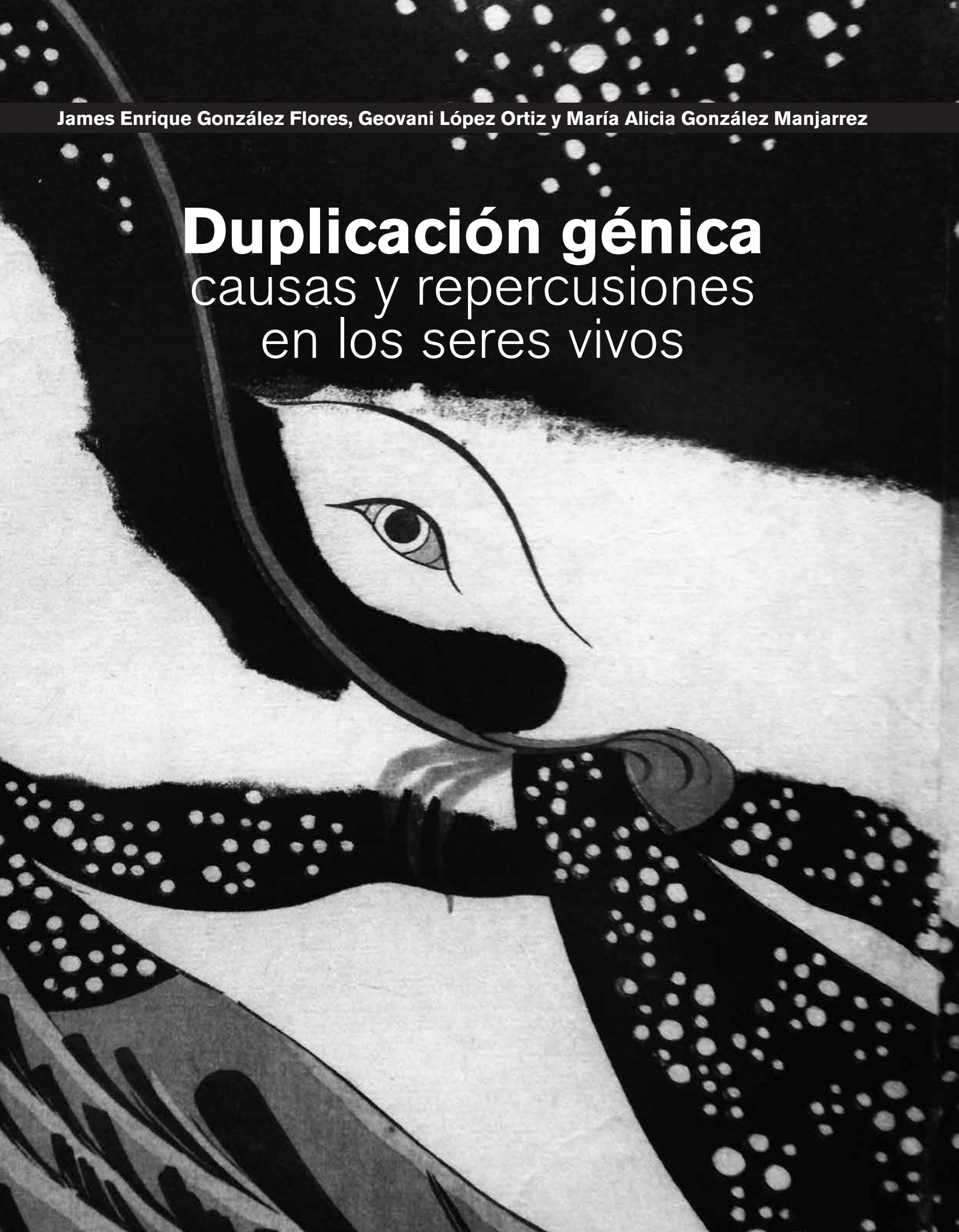


James Enrique González Flores, Geovani López Ortiz y María Alicia González Manjarrez

Duplicación génica

causas y repercusiones
en los seres vivos



Se entiende por duplicación génica el aumento en el número de genes a partir de secuencias preexistentes, y tales duplicaciones pueden comprender unos cuantos nucleótidos, genes, fragmentos de cromosomas, cromosomas e incluso genomas completos. La historia de su estudio se remonta a principios del siglo pasado, cuando en 1911 el botánico Yoshinari Kuwada encontró que uno de los cromosomas del maíz se encontraba duplicado. Posteriormente, en 1918, Calvin Bridges consideró la posibilidad de que fragmentos redundantes de cromosomas pudieran modificarse a lo largo del tiempo y llevar a cabo funciones distintas a las que desarrollaban originalmente. En 1932, el genetista John Burdon Haldane, en su libro *Las causas de la evolución*, propone que los eventos de duplicación génica podrían influir en la evolución de los organismos, en ese sentido, especies con más de una copia de un gen serían menos propensas a posibles efectos perjudiciales de mutaciones sencillas pues éstas se repartirían entre las diversas copias existentes. En 1970, Susumu Ohno publicó su libro *Evolution by gene duplication*, en el que estableció la importancia de la duplicación génica en la evolución de las especies. Sin embargo, fue hasta la década de los noventa cuando gracias al enorme auge que adquirió la secuenciación de genomas completos se encontró que la duplicación de genes era abundante en todos los organismos.

En la actualidad se ha determinado que la gran mayoría de las especies presentan un alto porcentaje de genes du-

plicados, y se considera que la duplicación constituye una de las principales fuentes de variación genética; es por lo tanto un elemento clave en la generación de la diversidad biológica (figura 1).

El origen de los genes duplicados

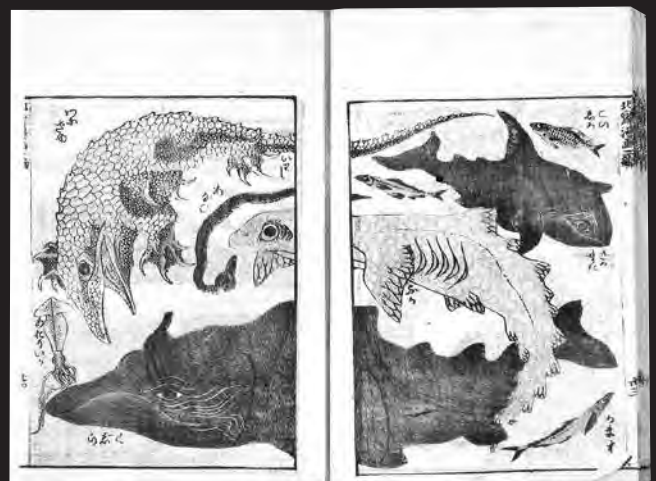
¿Cómo surgen nuevos genes y qué repercusión tienen en la fisiología de los organismos? Ésta es una de las preguntas más importantes que se ha planteado la genética molecular y cuya respuesta se encuentra justamente en el estudio de los genes duplicados o genes parálogos, los cuales pueden originarse mediante los siguientes procesos: a) entrecruzamiento de regiones homólogas del genoma o *crossing over*,

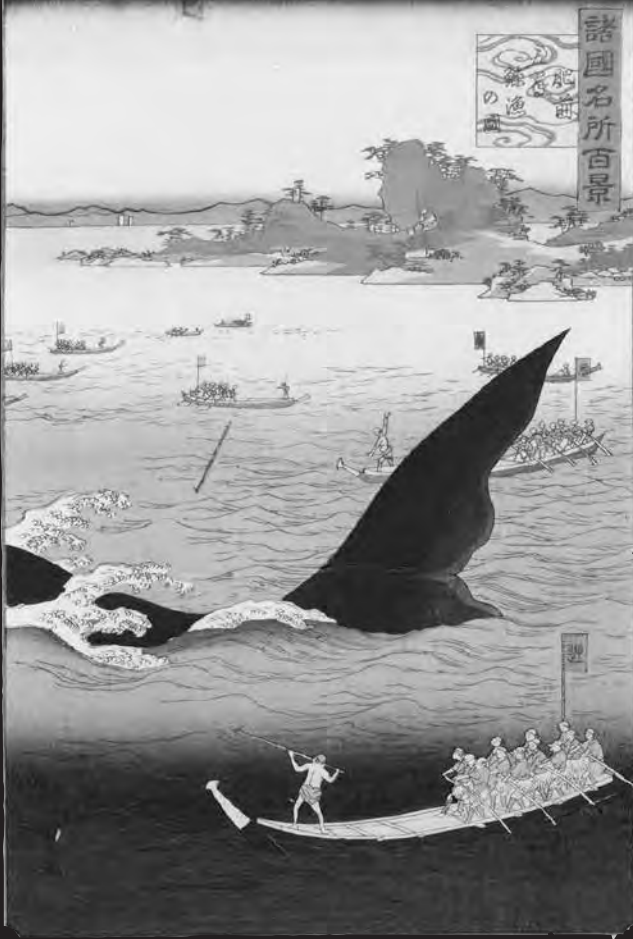
que suele generar genes repetidos en “tándem”, esto es, varios genes duplicados que se encuentran en posiciones adyacentes en el genoma —dependiendo del posicionamiento del entrecruzamiento, la región duplicada puede contener parte del gen, el gen entero o varios genes; b) retrotranscripción, que ocurre cuando un ARN mensajero es retrotranscrito a ADN complementario y después es insertado en el genoma, lo cual suele generar pérdida o ganancia de segmentos de ADN, que pueden ser insertados e interrumpir de manera aleatoria cualquier secuencia del genoma; y c) duplicaciones cromosómicas o del genoma, que pueden ocurrir durante la división celular, después de la replicación del ADN; un tipo de duplicación a gran escala, muy frecuente en plantas, pero rara vez se presenta en animales.



*En un sentido estricto,
nada en la evolución se crea de nuevo,
los genes surgen de genes preexistentes.*

SUSUMU OHNO





Destino evolutivo de los genes duplicados

Al término de la década de los setentas surgió un planteamiento importante sobre el destino evolutivo de los genes duplicados. El modelo de diversificación funcional sugerido por Ohno propone que los pares de genes duplicados evolucionan a diferentes velocidades, de suerte que una de las copias evoluciona lentamente, manteniendo la función original, mientras que la otra copia cambia rápidamente, adquiriendo con ello una nueva función. Este modelo ha sido analizado con gran detalle y se ha comprobado que tales escenarios únicamente se cumplen en ciertos casos. Un modelo alternativo conocido como duplicación-degeneración-complementación establece que, dado que muchos genes pueden tener más de una función, después de una duplicación ambas copias pueden mutar libremente, de manera tal que los dos genes que se han degenerado son necesarios para llevar a cabo la función ancestral, que realizaba el gen original antes de la duplicación.

En términos generales, los genes duplicados pueden tener cuatro posibles destinos evolutivos: 1) la conservación de la función ancestral, que ocurre cuando la dosis adicional representa una ventaja y se mantiene por conversión génica constante, y es común en genes que codifican para

proteínas que forman complejos macromoleculares y en donde es importante mantener la estequiometría de las subunidades del complejo; 2) la pseudogenización, que es cuando una de las copias se pierde, ya sea porque acumuló mutaciones sobre su región regulatoria que impiden la expresión del gen o bien por mutaciones en la región codificante que imposibilitan la generación de un producto funcional —tras una duplicación génica, los procesos de pseudogenización son los que ocurren con mayor frecuencia, y en especies como *Saccharomyces cerevisiae* llegan a ser cerca de 90%; 3) la subfuncionalización, cuando ambos genes son necesarios para llevar a cabo la función original debido a que han acumulado mutaciones sobre la región codificante o regulatoria, es decir, la función que previamente llevaba un solo gen se repartirá entre las copias que resultan de tal duplicación —esta división de funciones tiene múltiples implicaciones y hay ocasiones en que sólo es evidente bajo un contexto metabólico peculiar; y 4) la neofuncionalización, cuando mutaciones a lo largo de la región regulatoria o codificante inciden directamente sobre los patrones de expresión o sobre las propiedades bioquímicas del producto, facilitando la adquisición de nuevas funciones —los procesos de neofuncionalización de proteínas regulatorias incluyen el surgimiento de nuevos sitios de unión al ADN, y cuando los cambios ocurren sobre las secuencias regulatorias presentes en el ADN, éstas pueden ser reconocidas por elementos que no afectaban al gen ancestral. La neofuncionalización propicia la modificación de aminoácidos que alterarán la especificidad, la afinidad o las propiedades cinéticas de las enzimas, así como el establecimiento de nuevas

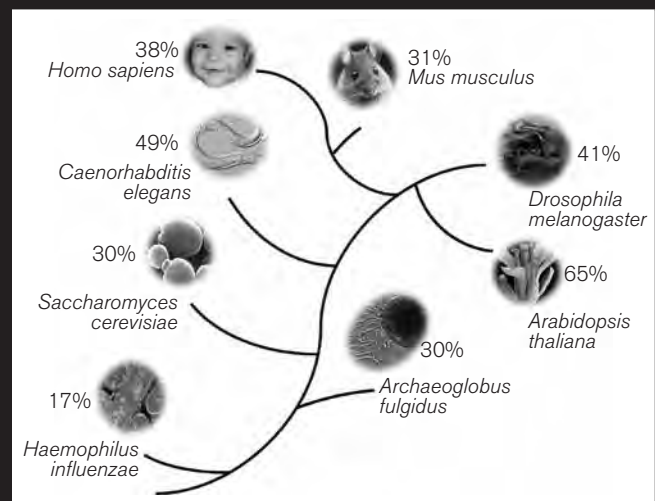


Figura 1. Representación del árbol de la vida junto con el porcentaje de genes duplicados en cada especie.



interacciones de proteínas debido a cambios en la secuencia codificante o de localización subcelular.

Un ejemplo de neofuncionalización en la naturaleza lo presentan los peces zoárcidos antárticos, que tienen la capacidad de sobrevivir a temperaturas cercanas a cero grados centígrados. Los ancestros de estos peces tenían un gen que codificaba para una enzima denominada ácido siálico sintasa; después de un evento de duplicación de este gen, una de las copias resultantes acumuló mutaciones que le permitieron codificar para una proteína que es secretada al exterior y cuya función ahora es evitar la formación de cristales de hielo, los cuales podrían ser fatales para estos peces; este hecho influyó de manera notoria en la adaptación de los mismos a ambientes fríos.

La importancia de la duplicación génica

La duplicación génica es un evento que está presente en los tres dominios del árbol de la vida: bacterias, arqueas y eu-

cariontes, es decir, constituye una característica vigente y concurrente en todos los seres vivos (cuadro 1). Su importancia radica en tres postulados clave: 1) representa un mecanismo principal en la innovación de genes y de funciones asociadas a sus productos; 2) ha sido un factor determinante en el surgimiento y evolución de vías metabólicas y por lo tanto en el origen de los primeros seres vivos; y 3) la diversidad biológica es el resultado de múltiples procesos de duplicación génica, desde la duplicación de pequeños fragmentos de un gen, hasta de genes y genomas completos.

Como se mencionó, la mayoría de los genes duplicados pueden acumular mutaciones rápidamente hasta perder su función; sin embargo, una fracción pequeña pero importante de genes duplicados (10%) puede permanecer en el genoma y formar familias completas, esto es, grupos de genes que provienen de un mismo gen ancestral. La familia de la tripsina, presente en el genoma de la mosca *Drosophila melanogaster*, es una de las familias más grandes de genes parálogos constituida por 111 miembros. Otro caso intere-

DOMINIOS	NÚMERO TOTAL DE GENES	NÚMERO DE GENES DUPLICADOS
BACTERIA		
<i>Mycoplasma pneumoniae</i>	677	298
<i>Helicobacter pylori</i>	1 590	266
<i>Haemophilus influenzae</i>	1 709	284
ARQUEAS		
<i>Archaeoglobus fulgidus</i>	2 436	719
EUCARIONTES		
<i>Saccharomyces cerevisiae</i>	6 241	1 858
<i>Caenorhabditis elegans</i>	18 424	8 971
<i>Drosophila melanogaster</i>	13 601	5 536
<i>Arabidopsis thaliana</i>	25 498	16 574
<i>Homo sapiens</i>	40 580	15 343

Cuadro 1. La predominancia de la duplicación génica en los tres dominios de la vida.



sante es la familia de genes parálogos que codifican para los receptores olfatorios en el ratón, la cual tiene cerca de 1 500 genes y se ha postulado que la mayoría de éstos emergieron a partir de unas cuantas copias ancestrales, mismas que a lo largo del tiempo se duplicaron de manera reiterada y generaron toda una gama de receptores con una mayor especificidad para el reconocimiento de los olores.

La robustez génica puede ser definida como: la habilidad que tiene un sistema biológico para resistir mutaciones y perturbaciones sin alterar su funcionamiento; esto puede ser propiciado por la redundancia génica, lo cual implica que los genes duplicados pueden balancear la pérdida de funciones de aquellas copias que reciben mutaciones que afectan directamente su función. La importancia de la robustez en un sistema biológico es que permite crear nuevas redes biológicas, rutas metabólicas y redes regulatorias alternativas; por lo tanto, la robustez permite la innovación evolutiva y la diversidad fenotípica.

Genes duplicados en la levadura de cerveza

El análisis de la secuencia genómica de la levadura *Saccharomyces cerevisiae* puso de manifiesto la existencia y la mag-



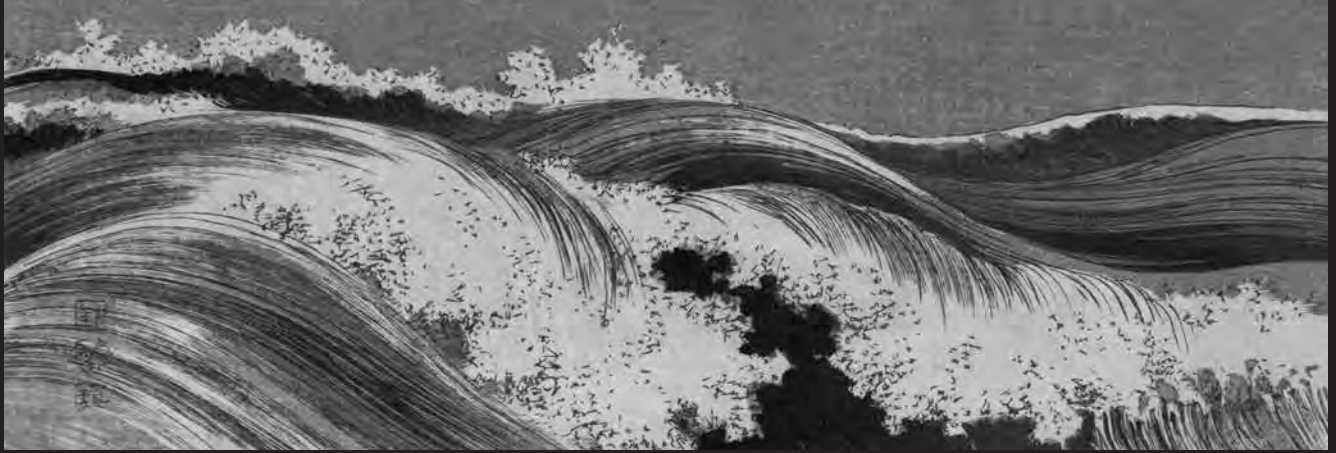
nitud de la redundancia génica; alrededor de 30% de su genoma está constituido por genes presentes en más de una copia. Estos genes parálogos son el resultado tanto de una duplicación total del genoma, la cual involucró la pérdida y retención selectiva de algunos de ellos, así como de eventos independientes de duplicación génica a menor escala. El hallazgo permitió identificar un gran número de genes parálogos que codifican para enzimas implicadas en el metabolismo del carbono y del nitrógeno.

Nuestro grupo de investigación inició el estudio de la diversificación de genes y proteínas parálogas utilizando como mo-

delo algunos pares de genes cuyos productos participan en la biosíntesis de aminoácidos en *S. cerevisiae*. La selección de tales genes está relacionada con enzimas que participan en la biosíntesis de aminoácidos y durante el metabolismo central del carbono (glucólisis y ciclo de los ácidos tricarbónicos), por tanto su especialización podría resultar en el desarrollo de funciones importantes en la utilización de compuestos de carbono relacionados con la generación de energía como el adenosíntrifosfato o ATP.

La levadura *S. cerevisiae* tiene la capacidad de crecer y generar energía en forma de ATP en condiciones fermentativas o respiratorias (metabolismo facultativo). Se ha pro-





puesto que la retención selectiva de genes parálogos facilitó la adquisición de un metabolismo facultativo en esta levadura, por tal motivo nuestro grupo se ha centrado en el estudio de la diversificación o redundancia de los genes parálogos, así como de sus productos. Para ello se analiza el papel de cada uno de los dos genes mientras la levadura se cultiva en condiciones fermentativas y respiratorias. Asimismo, hemos comparado la función de los genes parálogos de *S. cerevisiae* con la de los genes ortólogos presentes en otras especies denominadas “tipo ancestral”.

Algunas de las preguntas que se pretenden contestar al estudiar la función de los parálogos de la levadura son: ¿la duplicación de genes constituye una redundancia funcional?, ¿cómo diversifican los genes parálogos a lo largo de la evolución?, ¿qué mecanismos moleculares están involucrados en la diversificación funcional?, ¿cómo se diversifican las enzimas parálogas?, ¿qué ventajas adaptativas y fisiológicas representa el preservar tales genes duplicados?, ¿la conservación de genes parálogos constituyó una adaptación al metabolismo facultativo?

Entre las parejas de genes parálogos que nuestro grupo ha estudiado destacan genes implicados en la biosíntesis de aminoácidos tales como: GDH1/GDH3, LYS20/LYS21, BAT1/BAT2, ALT1/ALT2 y LEU4/LEU9, las cuales se conservaron después de la duplicación completa del genoma de *S. cerevisiae*. Éstos y los productos codificados se han estudiado utilizando metodologías que nos permiten establecer su función y compararla con la de los genes ortólogos de organismos tipo ancestral, un enfoque que nos permite determinar si la función tipo ancestral se ha conservado en los parálogos, si es redundante o se ha diversificado. El análisis se ha centrado en determinar las propiedades enzimáticas, la localización subcelular y las interacciones proteína-proteína en los parálogos, así como en su regulación transcripcional y la organización de su cromatina.

El caso que se ha estudiado con mayor detalle lo constituyen las proteínas parálogas GDH1/GDH3, las cuales asi-

milan nitrógeno y sintetizan glutamato a partir de amonio y α -cetoglutarato. En condiciones fermentativas, la formación de glutamato depende exclusivamente de la acción de GDH1, sin embargo, en condiciones respiratorias se requiere la expresión simultánea de GDH1/GDH3, lo que conlleva a la formación de isozimas heteroligoméricas para llevar a cabo su función en tales condiciones. Por lo tanto, la retención de dichos genes permite la formación de isozimas homohexaméricas (GDH1 y GDH3) y de isozimas heterohexaméricas en las que se combinan monómeros de GDH1 y GDH3. Cada una de las isozimas de esta familia tiene una capacidad peculiar de utilizar α -cetoglutarato, GDH1 posee





la mayor capacidad, en tanto que GDH3 tiene una menor capacidad. El crecimiento óptimo en etanol depende de la presencia de enzimas heterohexaméricas de capacidad intermedia. En glucosa la expresión de GDH3 se reprime por remodelación de cromatina y por tanto la isozima prevalente en esta condición es la GDH1 homohexamérica, lo que permite que en esta condición las levaduras crezcan a una velocidad mayor a la observada en etanol.

Otro caso interesante es el de las isozimas homocitrato sintasa codificadas por los parálogos LYS20/LYS21, las cuales intervienen en el primer paso de la síntesis del aminoácido lisina. El análisis de estas enzimas parálogas demostró que se requieren ambas isozimas para poder crecer en condiciones fermentativas, es decir, en esta condición la función de LYS20 y LYS21 es redundante. Sin embargo, la diversificación de las propiedades cinéticas de las proteínas LYS20/LYS21 fue esencial para el metabolismo respiratorio, ya que en tal condición la biosíntesis de lisina depende exclusivamente de LYS21. La inhibición diferencial por lisina de cada una de las dos isozimas y la regulación de la concentración intracelular de cada isozima resulta en un delicado sistema regulatorio que coordina la eficiencia con la que el α -cetoglutarato se utiliza para la síntesis de lisina o para la síntesis de algún otro intermediario. Por lo tanto, la retención y diversificación funcional de LYS20 y LYS21 con papeles parcialmente redundantes ha contribuido a la adquisición del metabolismo facultativo.

Otro ejemplo lo ilustran las aminotransferasas de cadena ramificadas (BCAATs, por sus siglas en inglés), las cuales son codificadas por los parálogos BAT1/BAT2 y catalizan el

primer paso de la biosíntesis y el último del catabolismo de los aminoácidos de cadena ramificada mediante la transferencia del grupo amino entre los aminoácidos valina, isoleucina y leucina y el α -ceto ácido correspondiente. Estudios recientes muestran que el único ortólogo KLBAT1 en la levadura tipo ancestral *K. lactis* es una enzima bifuncional que participa en la biosíntesis y el catabolismo de valina, isoleucina y leucina. Esta función dual ha sido distribuida en las proteínas parálogas BAT1/BAT2 de *S. cerevisiae*, apoyando el modelo de subespecialización que explica la evolución de los genes duplicados.



Hasta ahora, los resultados publicados han demostrado que la enzima ancestral distribuyó su función mediante la divergencia de expresión de BAT1/BAT2 y la localización subcelular diferencial de los productos. BAT1 se expresa en condiciones biosintéticas, mientras que BAT2 en catabólicas; BAT1 tiene una localización mitocondrial, mientras que BAT2 una localización citoplasmática. Este estudio demostró que la expresión diferencial y la localización subcelular de BAT1/BAT2 resultó en la distribución del papel biosintético y catabólico del gen tipo ancestral de *K. lactis* en dos isozimas que constituyen una adaptación al metabolismo facultativo.

Conclusiones

El estudio del destino evolutivo de algunos genes duplicados presentes en *S. cerevisiae* cuyos productos participan en la biosíntesis de aminoácidos indica que la retención y la diversificación de los mismos ha tenido un profundo impacto en la adaptación de la levadura *S. cerevisiae* al metabolismo facultativo, desempeñando un papel importante en la adquisición de un "nuevo" modo de vida que depende de la capacidad de crecer y obtener energía en condiciones fermentativas o respiratorias. 🌍



James Enrique González Flores

Geovani López Ortiz

María Alicia González Manjarrez

Instituto de Fisiología Celular,

Universidad Nacional Autónoma de México.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología CB-2014-239492. Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica PAPIIT IN201015.

REFERENCIAS BIBLIOGRÁFICAS

Force, A. *et al.* 1999. "Preservation of duplicate genes by complementary, degenerative mutations", en *Genetics*, vol. 151, núm. 4, pp. 1531-1545.

Gu, Z. *et al.* 2002. "Extent of gene duplication in the genomes of *Drosophila*, nematode and yeast", en *Molecular Biology and Evolution*, vol. 19, núm. 3, pp. 256-262.

Levasseur, Anthony y Pierre Pontarotti. 2011. "The role of duplications in the evolution of genomes highlights the need for evolutionary-based approaches in comparative genomics", en *Biology Direct*, vol. 6, pp. 1-12.

Lilly, M. *et al.* 2006. "The effect of increased branched-chain amino acid transaminase activity in yeast on the production of higher alcohols and on the flavour profiles of wine and distillates", en *FEMS Yeast Research*, núm. 5, pp. 726-743.

Lynch, M. y J. S. Conery. 2003. "The origins of genome complexity", en *Science*, núm. 302, pp. 1401-1404.

_____. 2000. "The evolutionary fate and consequences of duplicate genes", en *Science*, núm. 290, pp. 1151-1155.

Mombaerts, Peter. 2001. "The human repertoire of odorant receptor genes and pseudogenes", en *Annual Review of Genomics and Human Genetics*, vol. 2, pp. 493-510.

Ohno, Susumu. 1967. *Sex chromosomes and sex-linked genes*. Springer, Berlín.

Wolfe, K. H. y D. C. Shields. 1997. "Molecular evidence for an ancient duplication of the entire yeast genome", en *Nature*, núm. 387, pp. 708-713.

Zhang, J. 2003. "Evolution by gene duplication: an update", en *Trends in Ecology and Evolution*, núm. 18, pp. 292-298.

IMÁGENES

Pp. 116 y 120: Utagawa Kuniyoshi, *Miyamoto Musashi y la ballena*, 1848-1852. P. 117: No identificado, *Sol y olas*. Katsushika Hokusai: p. 117: *Ballena*, 1815; p. 120: *Rompiendo las olas*, 1847; p. 121 y 122: *Caza de ballenas en las Islas Goto*, s. XIX; p. 122: *Ola masculina*, s. XIX. Utagawa Hiroshige: p. 118 y 119: *La caza de ballenas en Goto en la provincia de Hizen*, s. XIX. P. 119: Itō Jakuchū, *Elefante y ola*, s. XVII. P. 121: Uehara Konen, *Olas*, ca. 1910.

GENE DUPLICATION: CAUSES AND REPERCUSSIONS IN LIVING THINGS

Palabras clave. Duplicación génica, evolución, diversificación funcional.

Key words. Gene duplication, evolution, functional diversification.

Resumen. La redundancia génica es una característica de todos los seres vivos y la duplicación de genes funcionales representa una fuente de material genético para el origen de capacidades nuevas o especializadas. Una duplicación genómica ofrece la posibilidad de que todas las proteínas codificadas por el ADN de un organismo evolucionen y que potencialmente éste se adapte a un "nuevo" modo de vida.

Abstract. Genetic redundancy is a characteristic of all living things and duplication of functional genes represents a source of genetic material for the emergence of new or specialized capabilities. Gene duplication offers the possibility of all proteins encoded by an organism's DNA evolving and the organism's potentially adapting to a "new" way of life.

James Enrique González Flores es biólogo por la Facultad de Estudios Superiores Iztacala y maestro en ciencias bioquímicas por la Facultad de Química de la UNAM; actualmente está por obtener el grado de doctor en el posgrado de Ciencias Bioquímicas de la misma casa de estudios.

Geovani López Ortiz es biólogo experimental por la Universidad Autónoma Metropolitana y doctor en ciencias bioquímicas por la Facultad de Química de la UNAM. Labora en la Subdivisión de Medicina Familiar en la División de Estudios de Posgrado en la Facultad de Medicina de la UNAM.

María Alicia González Manjarrez es bióloga por la Facultad de Ciencias de la UNAM y realizó su doctorado en investigación biomédica básica en la misma casa de estudios. Actualmente es investigadora del Instituto de Fisiología Celular de la UNAM, pertenece al Sistema Nacional de Investigadores (nivel III) y es jefa del Departamento de la División Académica de Ciencias Básicas, Bioquímica y Biología Estructural del mismo instituto.

Recibido el 7 de abril de 2015; aceptado el 18 de agosto de 2015.