



# Científicos españoles descifran el genoma completo del olivo, el árbol más emblemático del Mediterráneo

Un equipo de científicos de tres centros españoles ha logrado, tras tres años de trabajo, secuenciar por primera vez el genoma completo del olivo (*Olea europaea*), todo un hito en el estudio y aprovechamiento de este árbol milenario. Los resultados, publicados en la revista de acceso abierto *GigaScience*, abren la puerta a la mejora genética del olivo en lo que se refiere tanto a su capacidad productiva como a su capacidad para defenderse de amenazas externas causadas por hongos (*Verticillium dahliae*) y otros patógenos (por ejemplo, la temida *Xylella fastidiosa*, que ha arrasado zonas enteras en Italia). El ejemplar secuenciado es un olivo de más de 1.300 años de antigüedad de la variedad farga, una de las más importantes del este de España.



El olivo cuyo genoma completo ha sido secuenciado es un ejemplar de la variedad farga que supera los 1.300 años de antigüedad.

[Texto : Alfredo Briega Martín]

El olivo es uno de los primeros árboles que se domesticaron en la historia de la humanidad, seguramente hace unos 6.000 años. Emblema por excelencia del Mediterráneo, tiene una gran importancia en la economía española y de otros países de la zona (Italia, Grecia, Portugal). No en vano, Es-

paña es el máximo productor de aceite de oliva del mundo. Cada año se elaboran tres millones de toneladas de aceite, tanto para consumo local como para exportación, de las cuales España produce un tercio. Del mismo modo, nuestro país lidera la investigación acerca de diversos aspectos del olivo, si bien Francia, Italia, Portugal y Tur-

quía también son referentes en la investigación genética del olivo.

Sin embargo, hasta el momento se desconocía el genoma del olivo que regula, por ejemplo, las diferencias entre las variedades, tamaños y sabor de las aceitunas, por qué son tan longevos, o las claves de su adaptación al secano, aspectos muy valiosos para la producción de aceite.



El equipo del Centro de Regulación Genómica (CRG) que ha llevado a cabo la secuenciación del genoma del olivo. De izqda. a dcha., Marina Marcet-Houben, Toni Gabaldón y la primera autora del trabajo, Irene Julca.



El conjunto de computadoras de alta capacidad, también llamado clúster, en el que se almacenan los datos del genoma del olivo y se llevan a cabo los cálculos bioinformáticos.

[El proceso de anotación del genoma identificó la presencia de más de 56.000 genes, el doble que el genoma humano ]

Un equipo de investigadores del Centro de Regulación Genómica (CRG) de Barcelona, del Real Jardín Botánico de Madrid (CSIC-RJB) y del Centro Nacional de Análisis Genómico de Barcelona (CNAG-CRG) ha arrojado luz sobre el puzzle genético del olivo al lograr secuenciar por primera vez el genoma completo de esta especie. Los resultados de este trabajo, financiado íntegramente por el Banco Santander, se han recogido en la vanguardista revista de acceso abierto *GigaScience*. Esta publicación abre las puertas a nuevos trabajos de investigación que servirán tanto para profundizar en el desarrollo de la capacidad productiva del olivo como para ayudar a protegerle de infecciones que en la actualidad están causando verdaderos estragos, como son los ataques de bacterias (*Xylella fastidiosa*) y hongos (*Verticillium dahliae*, su principal amenaza en España).

“Es, sin duda, un árbol emblemático cuya mejora vegetal resulta muy difícil puesto que hay que esperar al menos 12 años para ver qué características morfológicas tendrá y si resulta o no interesante para hacer, por ejemplo, cruces”, destaca Toni Gabaldón, profesor de investigación de ICREA y jefe del Grupo de Genómica Comparativa del CRG, que ha liderado este trabajo. “Conocer la información genética del olivo nos permitirá contribuir a la mejora de la produc-

ción de aceites y aceitunas, de gran relevancia en la economía española”, añade.

#### Financiación privada para apoyar la ciencia pública

La historia de este trabajo empieza con un reconocimiento, una casualidad y un reto. Hace cuatro años, Gabaldón participaba junto a Pablo Vargas, investigador del CSIC en el Real Jardín Botánico, en la presentación de resultados científicos de proyectos sobre especies amenazadas, como el lince ibérico, que ya habían contado con financiación del Banco Santander.

Por entonces la entidad bancaria mostraba gran interés en financiar proyectos científicos españoles y, durante aquella presentación, Pablo Vargas propuso a Emilio Botín la secuenciación completa del genoma del olivo empleando exactamente la misma metodología utilizada en la secuenciación del lince; es decir, la estrategia tecnológica más vanguardista en la consecución de un genoma de calidad.

Cinco meses después de aquel encuentro, se firmó el contrato para llevar a cabo la primera secuenciación completa del ADN del olivo, una investigación coordinada por Pablo Vargas y que ha durado tres años.

“En la secuenciación de un genoma hay tres fases: la primera, aislar todos los genes, que es algo que ya publicamos hace dos años. La

segunda, ensamblar el genoma, que es ordenar esos genes uno detrás de otro, como si concatenáramos frases sueltas de un libro. Y, finalmente, identificar todos los genes, es decir, montar el libro. Esas dos últimas fases son las que hemos realizado y presentado ahora”, explica este investigador.

#### Descifrar su historia evolutiva

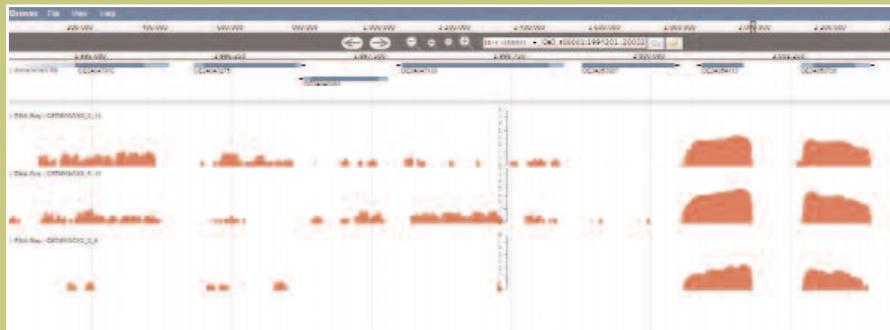
Además de la secuenciación completa del genoma del olivo, los investigadores también han comparado el ADN de este árbol milenario con otras variedades como el acebuche (olivo salvaje). Asimismo, han obtenido el transcriptoma, es decir, los genes que se expresan para valorar qué diferencias existen a nivel de expresión génica en hojas, raíces y frutos en diferentes estadios de maduración.

El siguiente paso será descifrar la historia evolutiva de este árbol, que forma parte de la vida de las poblaciones del Viejo Mundo desde que en la Edad de Bronce comenzara un proceso de domesticación a partir del acebuche en el este del Mediterráneo que resultó en los olivos actuales. Posteriores procesos de selección en distintos países del Mediterráneo dieron origen a las cerca de 1.000 variedades de cultivo que se conocen hoy día.

Asimismo, conocer la evolución de olivos de distintos países permitirá averiguar sus orígenes y desentrañar las claves que han posibilita-



Toni Gabaldón, jefe del grupo de Genómica Comparativa en el CRG y profesor de investigación ICREA.



Captura de imagen de la aplicación utilizada para visualizar el genoma del olivo. Se muestra una región que contiene siete genes (barras azules en la banda superior) y sus niveles de expresión (en rojo) en tres tejidos diferentes.

[Además de la secuenciación completa del genoma del olivo, los investigadores también han comparado el ADN de este árbol milenario con otras variedades como el acebuche (olivo salvaje) ]

do su adaptación a condiciones medioambientales muy diversas; además de obtener las claves de su extraordinaria longevidad, ya que pueden vivir hasta 3.000 y 4.000 años.

“De hecho, esa longevidad convierte al olivo que hemos secuenciado -de más de 1.300 años de antigüedad- casi en un monumento vivo”, señala Gabaldón. “Hasta el momento todos los individuos secuenciados, desde la mosca del vinagre (*Drosophila melanogaster*) hasta el primer ser humano analizado, han vivido un tiempo determinado, en función de la esperanza de vida de cada especie, pero luego han muerto o morirán. Ésta es la primera vez que se secuencia el ADN de un individuo de más de 1.000 años que puede seguir vivo tal vez otro milenio más”, explican Gabaldón y Vargas.

### Metodología y resultados

La estrategia utilizada que ha permitido conseguir un genoma de máxima calidad -no sólo para el olivo, sino también de referencia entre todas las plantas secuenciadas hasta la fecha- resulta innovadora en la ciencia de secuenciación de genomas. En concreto, la estrategia de secuenciación masiva

adoptada -combinación de Whole-Genome-Shotgun (WGS) y secuenciación de fragmentos clonados en librería de fósmidos- ha sido desarrollada por el CNAG-CRG y empleada con éxito en la secuenciación del lince ibérico, asegurando un ensamblado de máxima calidad. El tamaño del genoma ensamblado -alrededor de 1,31 GB de datos- representa el 95% del tamaño total estimado y contiene más del 98% de los genes del olivo. El proceso de anotación del genoma identificó la presencia de más de 56.000 genes codificantes de proteínas -significativamente más que los detectados en genomas secuenciados de plantas relacionadas y el doble que el genoma humano-, lo que sugiere la existencia de una duplicación completa del genoma del olivo respecto a otras plantas secuenciadas de la misma familia (Lamiales).

La disponibilidad de una secuencia de referencia para el genoma de la especie representa una oportunidad única para estudiar las diferencias genéticas entre las diferentes variedades cultivadas. La técnica de resecuenciación, que consiste en la secuenciación del material genético de un ejemplar y su comparación con un genoma de refe-

rencia, permite detectar diferencias genéticas (polimorfismos) entre el ejemplar secuenciado respecto a la referencia. Estas diferencias incluyen mutaciones puntuales y reordenamientos genómicos tales como inversiones, duplicaciones y pérdidas de material genético. La obtención de un catálogo de tales polimorfismos presentes en distintas variedades cuyas características agronómicas han sido caracterizadas permitirá identificar genes posiblemente implicados en rasgos de interés. Un paso necesario para acelerar los esfuerzos actuales de mejora genética y la obtención de nuevas variedades. Actualmente dichos esfuerzos están limitados por el lento crecimiento de esta especie y la necesidad de plantar grandes extensiones para observar rasgos en un número elevado de los descendientes de un determinado cruce. Los conocimientos derivados de un proyecto extenso de resecuenciación aplicado a un número suficiente de variedades posibilitarán el desarrollo de rápidos test genéticos que a su vez permitirán hacer la selección directamente a partir de semilla o de plántulas de pocos días, abaratando enormemente los costes y acelerando el proceso de mejora genética. 💧