

LA NUEVA ERA DEL ANÁLISIS CIENTÍFICO

Un alud de información satura el banco europeo de datos biológicos

► Los archivos sobre genes o moléculas están creciendo a pasos agigantados

► El gran laboratorio EMBL afronta el reto de que los repositorios se repartan por países

MICHELE CATANZARO
HEIDELBERG

La mayor base de datos biológicas de Europa, emplazada en el Reino Unido y perteneciente al Laboratorio Europeo de Biología Molecular (EMBL), debería teóricamente almacenar tres exabytes de datos en el año 2020. Eso es el equivalente a 3.000 millones de lápices USB (los de un gigabyte que se llevan en bolsos y llaveros), llenos de información sobre genes de personas y animales, estructuras de proteínas esenciales para la vida o compuestos químicos de los principales fármacos.

«Esta cantidad la estimamos a partir de la velocidad con la que se están generando nuevos datos», afirma Iain Mattaj, director general del EMBL desde el 2005, afincado en la sede central de Heidelberg (Alemania). El volumen de información se ha duplicado anualmente en los últimos 15 años. «Actualmente, nuestro centro dispone de una capacidad de almacenamiento mil veces menor (50 petabytes). Es impensable que lleguemos a tiempo para alcanzar la capacidad prevista para dentro de cinco años», explica.

Este es, según su director, el principal reto al que se enfrenta el laboratorio que fue creado hace 40 años con la ambición de convertirse en el «CERN de los genes». Es decir, para convertir a Europa en líder mundial en biología, tal y como lo ha hecho para la física el CERN (el centro de Ginebra que hospeda el gran acelerador de partículas).

21 PAÍSES // El EMBL -hoy una organización intergubernamental financiada por 21 países- celebra su aniversario en el clima de austeridad que impera en el continente. Sin embargo, el laboratorio ha conseguido que su presupuesto de los últimos tres años se haya mantenido más o menos estable con 206 millones de euros en el 2014. También sigue entre los 10 mejores centros del mundo en biología molecular. En total, 1.800 empleados trabajan en sus cinco centros en Heidelberg y Hamburgo, Grenoble (Francia), Hinxton (Reino Unido) y Monterotondo (Italia).

Sin embargo, el EMBL se está quedando corto de memoria. En 1980 empezó a acumular datos biológicos en el repositorio EMBL-Bank, em-



► Iain Mattaj, en la sede central del Laboratorio Europeo de Biología Molecular (EMBL), en Heidelberg.

Altibajos en la contribución española

► «España se ha beneficiado mucho del EMBL: muchos de sus mejores investigadores se han formado aquí», afirma Iain Mattaj. El barcelonés Centre de Regulació Genòmica (CRG), uno de los centros punteros en biología en España, se diseñó inspirándose en parte en el EMBL y con la participación de Mattaj, entre otros. Y su actual director, Luis Serrano, trabajó en el centro de Heidelberg. Sin embargo, a raíz de la crisis, en el EMBL se dudó de que España cumpliera con sus obligaciones financieras, admite Mattaj. «Finalmente, muy sabiamente lo hizo. Lo peor ha pasado», concluye. Actualmente, a España le corresponde aportar el 8% del presupuesto global del EMBL.

plazado en Hinxton, pero en los últimos años el coste y el tiempo necesarios para obtener información sobre el ADN y otras moléculas biológicas han bajado en picado. Gracias a ello, los científicos han podido hacer cosas como estudiar los componentes genéticos de muchas enfermedades o encontrar el parentesco evolutivo entre miles de animales. «Los biólogos están produciendo cada vez más datos, pero el espacio para almacenarlos y la potencia computacional para procesarlos están creciendo más lentamente», afirma Mattaj.

REPARTIR // «La solución no puede ser seguir concentrando los datos en un único repositorio», admite el científico. «Nuestra idea es crear una infraestructura que sume las capacidades de los institutos de bioinformática nacionales», explica. Desde el 2007, los expertos del EMBL han discutido esta idea y en diciembre del 2013 se formó la organización Elixir. «El proyecto se basa en una red de centros especializados. Por ejemplo,

los países escandinavos podrían consagrarse a datos de recursos marinos, por sus intereses pesqueros; Portugal, en datos de árboles, por su industria del corcho...», dice Mattaj.

Sin embargo, puede ser que el tsunami de información biológica sea aún más grande de lo previsto. «Por ejemplo, en todos los países se están generando enormes cantidades de otra clase de datos: las imágenes biológicas», observa Mattaj, en referencia a las gráficas producidas, por ejemplo, por las resonancias magnéticas. Por ello, en febrero, siete países del EMBL (España no está entre ellos) acordaron participar en el proyecto Euro-bioimaging, una red de centros para obtener, almacenar y analizar de forma coherente estas imágenes.

«En los próximos años, los métodos informáticos se aplicarán a todos los aspectos de la biología. Es un reto, pero gracias a ello podremos entender la complejidad de los organismos, especialmente el humano», concluye Mattaj. ≡

EL ADN de la semana

PERE
Puigdomènech



Africanos

Buena parte de la información sobre el ADN humano que hemos obtenido en los últimos años procede de personas que viven en Europa, Asia o América, pero no de donde más diferencias hay. Dos trabajos recientes nos recuerdan que es en África donde tenemos nuestros orígenes.

Uno de los trabajos, que ha analizado el genoma de más de 1.400 africanos, confirma que los estudios sobre genes y salud que se han hecho con poblaciones europeas pueden no ser válidos cuando se aplican a África. Por ejemplo, algunas poblaciones han acumulado variantes que les permiten resistir mejor la malaria y otras enfermedades. También se ha visto que una buena parte de los africanos actuales proceden de la llamada *expansión bantú* que se produjo a partir de Nigeria y Camerún hace 4.000 o 5.000 años.

La mayor diferencia entre humanos se produce en África. En los orígenes

Este trabajo y un segundo que ha estudiado el genoma de cinco bosquimanos del sudoeste de África describen detalles de la aventura inicial de los humanos. Según los autores, la especie se formó en algún lugar de África y se expandió a continuación por el resto del continente. Hace unos 150.000 años, seguramente por cambios en el clima, la población se redujo mucho en todas partes excepto donde viven los bosquimanos, que deberían de ser en ese momento los humanos más abundantes. Los restantes africanos, de los cuales proceden los bantús y aquellos que salieron de África para poblar el mundo, eran unos pocos miles.

La conclusión de estos estudios es que los bosquimanos son ahora quizá tantos como cuando comenzó la expansión de los humanos por el mundo, mientras que los pocos aventureros que salieron de África se han convertido en 7.000 millones. En su viaje encontraron otros homínidos, como los neandertales, y fueron cambiando en aspectos como el color de la piel y la reacción a determinadas comidas, además de fundar sociedades organizadas. Ahora el mismo estilo de vida de los bosquimanos está en peligro. Tenemos que decidir qué hacemos. ≡