

# Tendencias

Hacia una economía del conocimiento

JOSEP CORBELLA  
Barcelona

**B**arcelona, tenemos un problema: el volumen de datos biológicos que se está obteniendo en centros de investigación y hospitales está creciendo más rápido que la capacidad de almacenarlos y digerirlos. Llegará un punto, a medio plazo, en que la capacidad de los ordenadores no dará abasto para gestionar la avalancha de datos que generan médicos y biólogos.

Para una ciudad que se mira en el espejo de Boston para construir una economía del conocimiento, y que cuenta ya con algunos de los equipos de bioinformáticos más importantes de Europa, "resolver este problema es una cuestión estratégica", afirma Modesto Orozco, bioquímico de la Universitat de Barcelona (UB) e investigador del Institut de Recerca Biomèdica (IRB) y del Barcelona Supercomputing Center (BSC).

El problema, con el que no sólo se encuentra Barcelona sino todas las capitales de investigación biomédica, se explica por el drástico abaratamiento de secuenciar genomas que se ha registrado en los últimos cinco años. Si en el

## Barcelona, capital GENÓMICA

*La caída del coste de investigar el ADN provoca una avalancha de datos biológicos*

2007 el coste de secuenciar cada genoma humano era de diez millones de dólares, la llegada de una nueva tecnología de secuenciación ha hecho que el precio haya caído ahora por debajo de los 10.000 dólares. Este abaratamiento ha llevado a que un número creciente de investigadores y de médicos hayan empezado a secuenciar genomas en cuanto han podido permitirselo.

Hoy día se secuencian genomas de células tumorales, de variedades de melón, de fósiles de neandertales... En el caso del cán-

cer, se quieren comparar los genomas de células tumorales con los de células sanas del mismo paciente para ver cuáles son los genes que están alterados en el tumor. En un futuro próximo, se prevé poder comparar el genoma de células tumorales antes y después del tratamiento para evaluar la eficacia de los distintos fármacos... Y no sólo se almacenan datos de genomas. También se han empezado a estudiar epigenomas (que indican qué genes están activos y cuáles están silenciados en un momento dado en una

célula concreta); proteomas (que indican qué proteínas hay en una célula); transcriptomas (que indican qué ARN hay)... En fin, una cantidad enorme de datos que van camino de desbordar la capacidad de los superordenadores.

Según la famosa ley de Moore, la capacidad de almacenamiento de datos de los chips se duplica cada 18 meses. Pero según una estimación del Institut Europeu de Bioinformàtica, con sede en el Reino Unido, la cantidad de datos biológicos que se almacenan se duplica cada nueve meses.

"Si el volumen de datos aumenta más rápido que la capacidad de almacenarlos, nos encontramos con un cuello de botella que limitará la investigación y dificultará que estos datos se apliquen en beneficio de los pacientes", advierte Ivo Gut, director del Centro Nacional de Análisis Genómico (CNAG) en Barcelona.

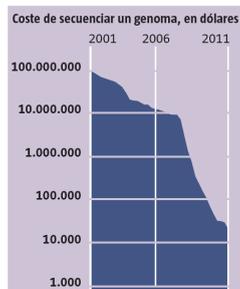
La solución, según Modesto Orozco, requiere gestionar los datos de la manera más eficiente posible. Si hasta ahora no se ha hecho, es porque la secuenciación de genomas es una tecnología joven con la que cada centro de investigación ha empezado a experimentar por su cuenta. Pero, llegados a la actual crisis de crecimiento de esta tecnología, "es mejor compartir recursos que duplicarlos", sostiene Orozco.

Compartir recursos significaría establecer un banco central de datos biológicos que guardara la información de los distintos hospitales y centros de investigación. "No tiene sentido -según Orozco- que cada institución gestione sus propios datos de manera aislada; sería mucho más eficiente que estuviéramos todos conectados".

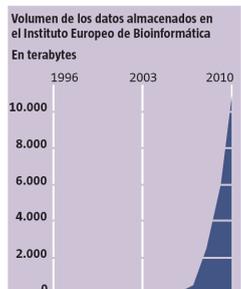
El Biodata Center -como lo llama Orozco de manera provisional- estaría vinculado al superordenador Mare Nostrum en el

*Los ordenadores no dan abasto para gestionar la avalancha de datos que aportan los biólogos*

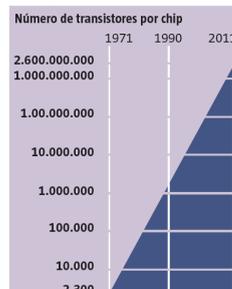
Como el precio de secuenciar genomas cae en picado desde el 2007...



... el volumen de datos biológicos se multiplica por dos cada nueve meses



Sin embargo, la capacidad de almacenar datos en un chip se duplica cada 18 meses



*Por lo tanto, el aumento de la capacidad de almacenamiento no será suficiente para gestionar todo el volumen de datos biológicos*

FUENTE: Elixir / BSC

LA VANGUARDIA

### Trece países crean un consorcio bioinformático europeo

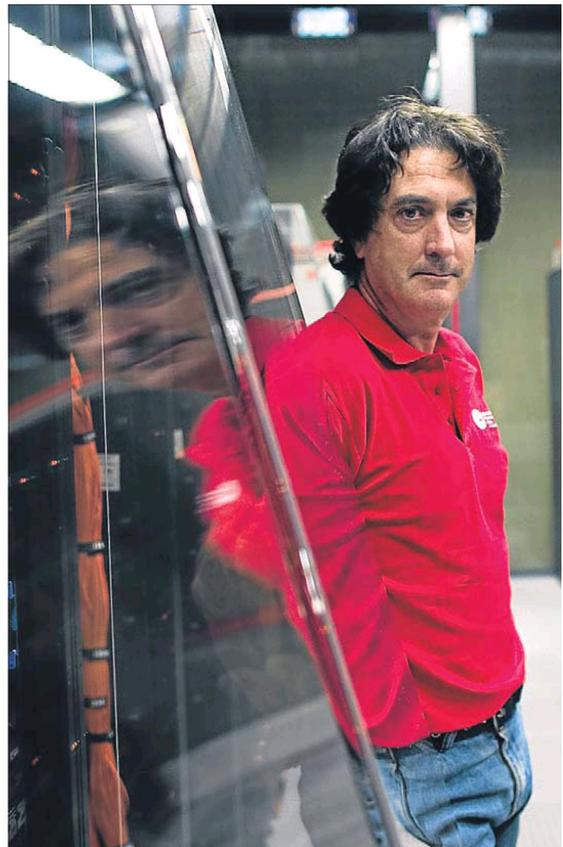
Trece países europeos -España entre ellos- han creado el consorcio Elixir para gestionar la avalancha de datos biológicos derivados de la revolución genómica. El plan de trabajo prevé que Elixir se construirá como una infraestructura con un *hub* principal situado en el Reino Unido -el Estado que más invierte en el proyecto y el que lidera la investigación genómica en Europa- y varios

nodos distribuidos por toda Europa. En España, están involucrados en el proyecto Elixir el Centro Nacional de Investigaciones Oncológicas (CNIO) y Barcelona Supercomputing Center.

"El reto que plantea la avalancha de datos biológicos es demasiado grande para que ningún país lo pueda afrontar por sí solo", declaró el 3 de septiembre Sorren Brunak, presidente de

la junta directiva de Elixir.

La fase preparatoria del proyecto se inició en el 2007 con financiación de la Comisión Europea y concluirá en diciembre de este año. El calendario de trabajo prevé que la fase de construcción de la infraestructura dure cuatro años y termine en diciembre del 2016. A partir del año 2017 se iniciará la fase permanente de explotación de la red Elixir.



Modesto Orozco, en el Barcelona Supercomputing Center

Printed and distributed by NewspaperDirect  
www.newspaperdirect.com US Call 1 877 662 8500 ISSN 1135-2545  
COPYRIGHT AND PROTECTED BY APPLICABLE LAW

