

# CIÈNCIA



TEXT\_\_MÓNICA L. FERRADO

LA INFORMACIÓ QUE CONTÉ EL GENOMA HUMÀ ocupa uns 3 gigabytes (GB). Un volum que equival a la informació que contenen els 11 volums de l'*Enciclopèdia Britànica*. Un espermatozou humà emmagatzema 37,5 megabytes (MB) d'informació i en una ejaculació s'alliberen 1.587 GB. La biologia humana ha entrat de ple dins del terreny del *big data*. Als laboratoris s'utilitzen tecnologies moleculars molt variades per obtenir el que es coneix com a *dades òmiques* (de la proteòmica, genòmica, exòmica...). Les noves tecnologies són capaces de mesurar els petits engranats moleculars que amaguen la clau de moltes malalties i que obren camí a la medicina personalitzada.

La biologia ha entrat indiscutiblement en el terreny del *big data*, l'anàlisi de multitud de dades de pacients o persones sanes, per tro-

bar les claus per lluitar contra malalties com el càncer, malalties neurodegeneratives i altres que avui són incurables. Catalunya participa en alguns dels projectes més rellevants de *big data* i biomedicina que s'estan desenvolupant a escala mundial.

#### BIOLOGIA I INFORMÀTICA

A Catalunya uns 800 científics treballen al voltant del sector de la bioinformàtica. Els bioinformàtics -els autèntics miners de les dades biomèdiques- han constituït l'associació Bioinformàtics Barcelona (BIB), que té 25 entitats membres entre universitats, centres de recerca, instituts de recerca hospitalària, grans infraestructures científiques, empreses i altres institucions. I n'hi ha 23 més d'interessades. "El BIB] Donarà resposta als reptes plantejats pel creixement de macrodades

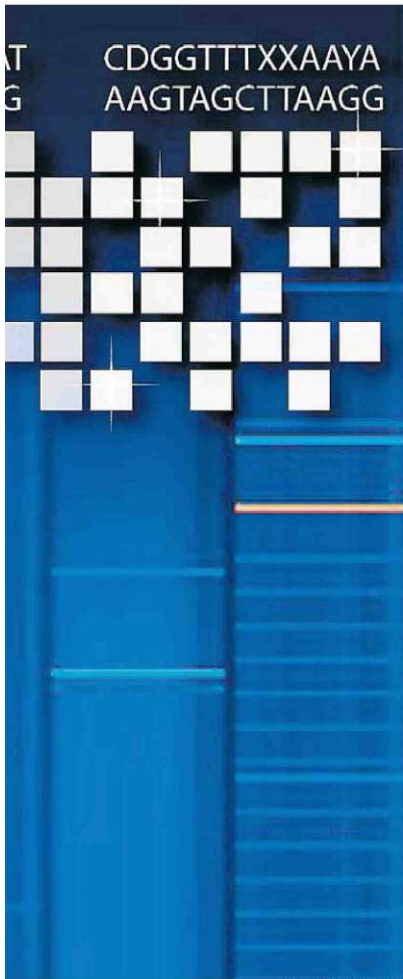
en l'àmbit de la salut i en l'agroalimentari, com per exemple la medicina personalitzada i l'agricultura de precisió", expliquen en un comunicat.

#### POTÈNCIA I EMMAGATZEMATGE

Les primeres bases de dades massives en biomedicina es van crear als anys 50, quan es van començar a emmagatzemar seqüències de proteïnes. Als anys 80, fent *mineria de dades* es va trobar el primer oncògen. Des dels orígens del *big data* fins ara tot ha canviat molt. Amb tot, avui l'ànim de superació continua centrant-se en el mateix aspecte: "La potència necessària per fer càlculs complexos i la capacitat d'emmagatzematge d'informació". Ho explica Roderic Guigó, coordinador del programa de bioinformàtica del Centre de Regulació Genòmica (CRG). Catalunya també té instal·la-

cions punteres, com el Centre de Supercomputació de Barcelona (BSC), amb gran potència per emmagatzemar i analitzar dades, i el Centre Nacional d'Anàlisi Genòmica (CNAG), amb grans màquines per seqüenciar.

Tots aquests ingredients han fet que l'Institut Europeu de Bioinformàtica (EBI) hagi confiat en la ciutat de Barcelona per establir-hi la seu de l'Arxíu Europeu del Genoma-Fenoma (EGA), encapçalat pel CRG i que recull tot un tresor: les dades genètiques de cent mil pacients de tot el continent, que han participat en més de set-cents estudis al voltant del càncer, la diabetis i altres malalties. Es guarden les dades genètiques però també el que es coneix com a fenoma, és a dir, dades que van des del color dels ulls de la persona fins a altres característiques fisiològiques i malalties. Hi poden accedir científics de tot arreu.



MAURO MANFREDINI...GETTY

El futur de la biomedicina és en el 'big data'. Obre camí a la medicina personalitzada i a desvelar les claus de malalties que ara són incurables, com és el cas del càncer i l'Alzheimer



GETTY IMAGES

El genoma humà ocupa el mateix espai d'informació que 11 volums de l'*Enciclopèdia Britànica*. Per estudiar el genoma del càncer s'analitzen prop de 1.000 genomes. En el projecte hi participen científics d'arreu del món, també de Catalunya.

Al món tan sols hi ha una altra base de dades similar, als Estats Units.

#### GRANS PROJECTES

Un dels projectes mundials de *big data* més ambiciosos és el Consorci Internacional del Genoma del Càncer. Elias Campo, director de recerca de l'Hospital Clínic, en codirigeix una de les parts, el Consorci Internacional per a l'Estudi del Genoma de la Leucèmia Limfàtica Crònica. "Hem fet dos tipus d'estudis: el genoma complet de 150 individus, i l'exoma de 400 pacients", explica Campo. L'exoma són les regions del genoma on hi ha les parts codificants dels gens, que formaran l'ARN missatger i que, en ser traduït per la maquinària cel·lular, donarà lloc a les proteïnes. És la part funcional més important del genoma perquè és el que finalment determina com és un organisme. Supo-

sa entre un 1,5% i un 3% del genoma humà. "Estem veient gens mutats fins ara no relacionats amb cap tumor humà, i alguns estan associats a formes més agressives de la malaltia", explica Campo. Al projecte també hi participen el CNAG i el BSC. A més, Catalunya també forma part del projecte Encode, un consorci per trobar la utilitat de les parts més fosques i desconegudes del genoma, més enllà dels gens.

#### DINS DEL CERVELL

Cada any es publiquen uns seixanta mil articles científics sobre el cervell. I tots són parcials, ja que se centren en algun aspecte concret. No obstant, el cervell és una xarxa integrada, i precisament a causa d'aquesta parcialitat hi ha molts aspectes que es resisteixen a la ciència. El programa Human Brain Project té com a objectiu integrar totes

aquestes dades i construir un gran cervell virtual, neurona a neurona. Només així es podrà determinar la complexitat amb què operen els circuits cerebrals, la funció exacta dels neurotransmissors i, en aquest entramat, què falla en moltes malalties psiquiàtriques i neurodegeneratives. També es podran desvelar grans misteris, com ara què passa al nostre cervell quan tenim una idea o quan prenem una decisió.

En el Human Brain Project hi participen 80 centres de tot el món. Entre ells, el BSC i l'Institut de Recerca Biomèdica de Barcelona (IRB). Treballen amb dades per fer simulacions a escala molt petita, entre neurones. A partir de les dades biològiques, els científics podran reproduir amb detall el procés de sinapsi (la unió especialitzada amb què les neurones s'envien senyals entre elles o a les cèl·lu-

les no neuronals). Amb models matemàtics simularan els potencials elèctrics entre neurona i neurona i el de les molècules que hi intervenen. Quan estigui fet, també es podrà simular l'acció de fàrmacs per potenciar o inhibir aquesta comunicació, una fita que seria aplicable a la millora dels tractaments per a malalties com la depressió i l'esquizofrènia.

#### DADES DE PACIENTS

La utilització de dades massives procedents d'històries clíniques de pacients també obre la porta a detectar causes de malalties que sense aquesta perspectiva global queden amagades. Fins a 39 investigadors catalans s'han adherit a un manifest a favor del projecte Visc+ de cessió de dades sanitàries per a la ciència, sempre amb estrictes mesures que garanteixin la confidencialitat. ■