

Genèticament bons

Considerat pel prestigiós Scimago Institutions Ranking com el sisè centre de recerca amb més qualitat del món en el sector de la salut, el CRG és un referent en la investigació genètica. Amb una clara voluntat internacional, més d'un 60% dels seus investigadors són estrangers que provenen de més de quaranta països. A més, el CRG col·labora en molts projectes europeus i gestiona l'Arxiu Europeu de Genoma-Fenoma, que emmagatzema de manera permanent i segura la gran quantitat de dades que genera el camp de la genètica.

Reportatge d'Àstrid Bierge

La vida és extremadament complicada. L'idioma genètic, format només per quatre lletres, codifica un extens i complex manual d'instruccions, l'ADN, que ens defineix, ens crea i ens regenera. Saber traduir-lo, llegir-lo i fins i tot modificar-lo és cabdal per poder arreglar els errors *tipogràfics* que provoquen malalties.

El Centre de Regulació Genòmica (CRG), creat l'any 2000 i ubicat al Parc de Recerca Biomèdica de Barcelona, busca comprendre la complexitat de la vida, des del genoma fins a la cèl·lula, els teixits, l'organisme complet i la interacció amb el seu ambient. Amb una trentena de grups de recerca i set unitats tecnològiques pròpies, el CRG duu a terme una investigació transversal per comprendre les malalties des d'un punt de vista genètic i d'una manera integrada.

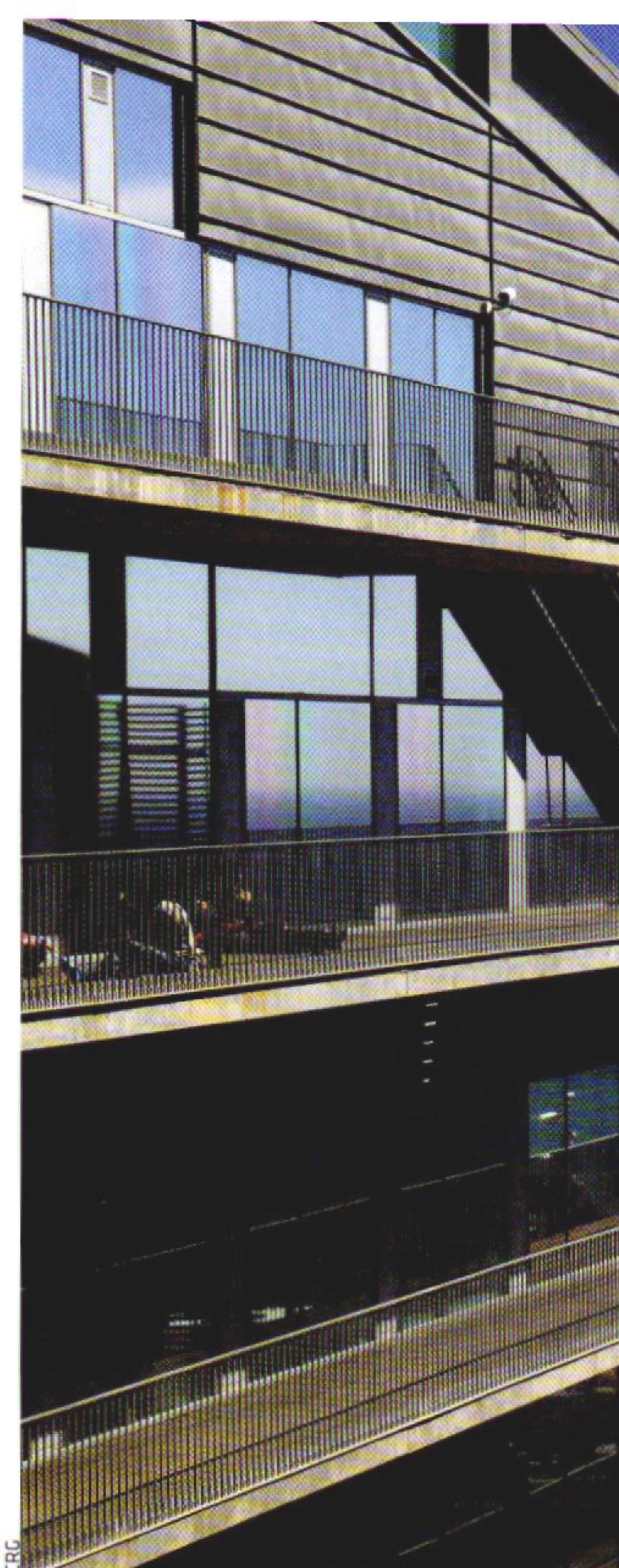
Regulació genètica, cèl·lules mare i càncer

Aquest és el nom d'un dels quatre programes del CRG. Els diferents grups de recerca que l'integren investiguen els mecanis-

mes que fan que els gens s'expressin. Què fa que les cèl·lules s'especialitzin i siguin diferents? Quines decisions prenen i com les prenen? L'estudi de les cèl·lules mare –que encara no estan especialitzades– és clau per entendre l'expressió genètica en el procés de diferenciació.

En aquest programa també s'investiga la reprogramació de cèl·lules. Per exemple, el grup d'en Thomas Graf ha aconseguit reprogramar limfòcits i convertir-los en macròfags. Ambdues són cèl·lules de la sang, però els limfòcits de vegades comencen a atacar cèl·lules que no toquen, de manera que poden provocar leucèmia i limfomes. En canvi, els macròfags sempre són inofensius. De moment aquestes transformacions només s'estan fent al laboratori, però algun dia aquesta tècnica podria servir per tractar el càncer, convertir les cèl·lules nocives en cèl·lules benignes. Un mètode molt menys intrusiu que la quimioteràpia, que ataca indiscriminadament.

El grup de Pia Cosma treballa en la reprogramació i la regeneració de cèl·lules, sobretot de neurones de la retina. Són



capaços de fer que un ratolí cec pugui veure-hi!

Quatre dels grups d'aquest programa estan implicats en un projecte de cinc anys que ha aconseguit un ajut de dotze milions d'euros del Consell Europeu de Recerca (ERC). Es tracta del 4DGenome, que investiga en la lectura del genoma, que està començant a canviar. Fins fa poc l'ADN es llegia d'una manera lineal, posant les lletres en fila com les paraules d'un llibre. Ara se sap que amb aquest mètode es passa per alt informació molt important. Dins del nucli d'una cèl·lula, l'ADN no està posat en fila. De fet, està entortolligat d'una manera molt concreta que varia segons les cèl·lules i que és tan complexa com significativa. Per exemple, quan fas una lectura lineal de l'ADN, no veus quins gens estaven físicament propers dins del nucli, i aquesta proximitat influeix en la relació que estableixen



entre ells. Cal tenir en compte que moltes de les nostres característiques fenotípiques, físiques, estan controlades per molts gens, i entendre aquestes connexions és un dels grans reptes de la genètica. També s'ha vist que, en l'empaquetat de l'ADN dins del nucli, els fragments que s'han de llegir més sovint en un tipus determinat de cèl·lula queden al descobert, mentre que les parts que s'expressen menys queden més recargolades, més amagades. La informació més important també queda especialment protegida pel sistema d'empaquetat, així s'evita que es faci malbé.

Les nostres fàbriques

El programa de biologia cel·lular i desenvolupament del CRG estudia el funcionament i els processos de les cèl·lules. És una aproximació més clàssica però en cap cas senzilla. Les cèl·lules són

fàbriques microscòpiques fascinants. Les proteïnes són les treballadores d'aquestes indústries, les que construeixen estructures i sistemes, fan de transportistes, reparteixen nutrients, fan noves proteïnes, etc. Les cèl·lules també tenen òrgans, que si fa no fa serien els seus òrgans. Com el complex de Golgi, que és una mena de centralita de correus. Rep les proteïnes, les empaqueta i les envia on cal. I com sap que la proteïna que s'acaba de fabricar és d'adhesió i ha d'anar a enganxar dues cèl·lules? El grup del CRG liderat per Vivek Malhotra estudia el funcionament d'aquest i altres processos cel·lulars.

Per reparar-se, les cèl·lules dupliquen el material genètic i després es divideixen en dos cèl·lules filles iguals. També en el programa de biologia cel·lular i desenvolupament, el grup de Manuel Mendoza va descobrir una proteïna que fa la funció d'un semàfor. Només es posa en verd

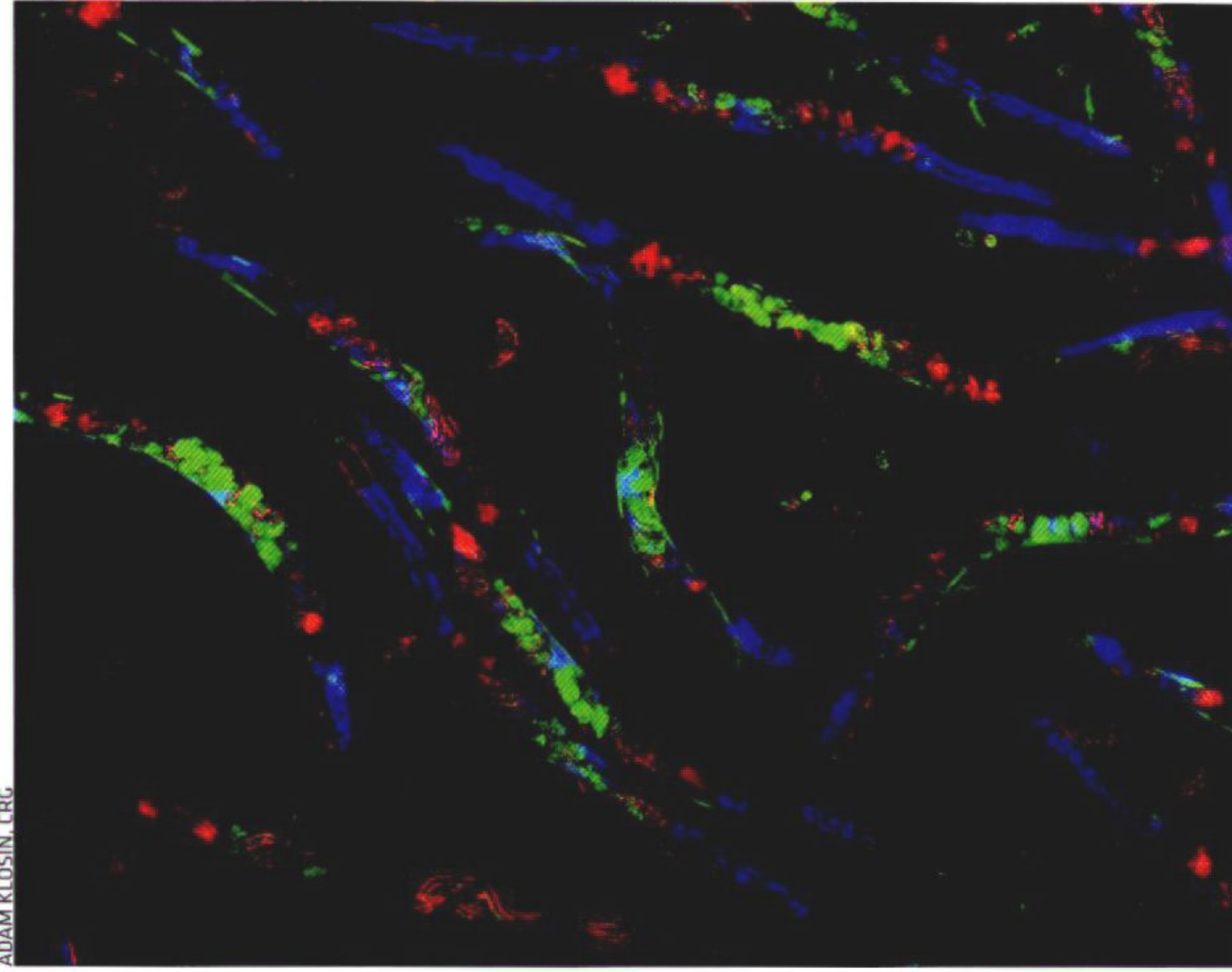
quan el material genètic ja ha quedat ben repartit i pot haver-hi la divisió final. Si se separen abans, malament.

El grup del Jérôme Solon fa una aproximació més biomecànica de la biologia cel·lular, estudiant l'embrió quan encara és una pilota de cèl·lules. Volen esbrinar si aquestes cèl·lules saben on anar o si es mouen influïdes per forces físiques.

Bibliotecaris dels gens

El programa de bioinformàtica i genòmica del CRG fa front als reptes que planteja la gran quantitat de dades generada per la tecnologia genòmica, tant pel que fa a l'emmagatzematge com a l'anàlisi. Aquests grups utilitzen la computació per estudiar i modelar fenòmens biològics, dissenyant també nous algorismes que ajudin a identificar patrons en les dades i facilitin aquest modelatge. També generen programari per estudiar els canvis →

Espècimens adults del cuc *Caenorhabditis elegans*. Dins porten diversos embrions, alguns dels quals han estat genèticament modificats perquè desenvolupin teixit muscular. Els embrions reprogramats es veuen en vermell i els que no ho estan, en verd.



ADAM KLOSIN, CRG

→ evolutius del genoma, amb models que permeten fer prediccions sobre el futur d'aquesta evolució.

A banda d'aquest programa, el CRG compta amb una unitat tecnològica de bioinformàtica disponible per a tots els investigadors del centre. El modelatge computacional permet experimentar i predir molts fenòmens, i ajuda els investigadors a triar amb molta més punteria els experiments que poden resultar interessants al laboratori.

El prestigi internacional del programa de bioinformàtica del CRG ha atret a Barcelona l'Arxiu Europeu del Genoma-Fenoma (EGA), un projecte que des del 2013 gestiona el CRG en col·laboració amb l'Institut Europeu de Bioinformàtica (EMBL-EBI).

Cada cop els hospitals, les universitats i centres de recerca tenen més capacitat per seqüenciar mostres d'ADN. Però un cop s'acaba un estudi concret, cal emmagatzemar les dades que s'han utilitzat perquè altres investigadors puguin tenir-hi accés. Alhora, cal garantir la privacitat de les persones a les quals corresponen les dades genètiques. L'EGA és un arxiu que centralitza aquesta informació perquè els investigadors puguin aprofitar les dades utilitzades en altres estudis i alhora garanteix la privacitat dels donants.

Físicament, aquestes dades -aproximadament un milió d'arxius provinents de tot el món- estan guardades al Barcelona Supercomputing Center, que és on hi ha el MareNostrum i altres superordinadors. Ocupen 3,5 petabytes, és a dir, 3,5 milions de gigabytes. Una xifra que creix al voltant d'un 50% cada any.

Un investigador que estigui estudiant, per exemple, la diabetis, pot descarregar de l'EGA seqüències genètiques de persones amb aquesta malaltia i de persones que no la tenen, així pot comparar i veure'n les diferències. Però no tots els arxius estan classificats per patologies, també n'hi ha de poblacionals. Per exemple, hi ha un arxiu que es diu "el genoma dels holandesos".

L'abast dels arxius de l'EGA varia. N'hi uns quants que contenen la seqüència del genoma complet d'una persona, però és més habitual trobar estudis d'exomes, en què hi ha seqüenciada només la part del genoma que codifica fer proteïnes. És a dir, al voltant d'un 1% del total. També guarda els anomenats xips d'ADN, que contenen una selecció de zones del genoma que tenen informació interessant.

El CRG gestiona aquest servei, organitzant i enriquint les dades i facilitant la feina dels usuaris. Per exemple, quan un investigador està interessat en un arxiu, el CRG li dona una certa visió de què podrà trobar-hi. Així pot saber si realment li serà útil abans de fer totes les gestions burocràtiques per poder-lo descarregar. A més, cal tenir en compte que alguns d'aquests arxius són molt grans i necessiten molt d'espai. L'investigador vol estar segur que aquesta informació li interessa.

L'EGA també està planejant un nou servei per actualitzar les dades dels arxius. Els mètodes de seqüenciació genètica avancen ràpidament, cada cop són capaços de llegir més informació i de comprendre-la millor. Actualment ens trobem en la versió 38 del genoma humà, però molts dels arxius de l'EGA han estat seqüenciats amb versions anteriors. El nou servei revisarà les dades

originals per anar-les actualitzant, així no haurà de fer-ho cada investigador que descarregui un arxiu concret.

Com hem dit, una de les funcions de l'EGA és garantir la privacitat dels donants d'ADN. Els arxius de les bases de dades genealògiques mai no porten un nom propi assignat, però, tot i així, sense estrictes mesures de seguretat, hi ha un risc que aquestes persones puguin ser identificades. A través de les seqüències genètiques es pot conèixer molta informació de la persona, com el gènere, l'edat aproximada, la raça i fins i tot els lligams parentius. Hi ha gent que amb el seu propi genoma ha accedit a bases de dades de genealogia i ha pogut localitzar parents desconeguts. L'EGA té un sistema d'accés a les dades que garanteix que la informació només pugui ser utilitzada per fer recerca i que mai no es trenqui el contracte de confidencialitat signat pels donants.

Jordi Rambla, que actualment dirigeix l'EGA, ens diu que aquest arxiu és "com una biblioteca de manuscrits". Si vols tocar un llibre original necessites un permís especial i estaràs sotmès a un control. Quan un grup de recerca vol descarregar un arxiu de l'EGA, s'ha de posar en contacte amb el grup que ha penjat aquella informació i demanar-li permís. Sovint, qui puja la informació designa un comitè ètic perquè s'encarregui de comprovar

El CRG compta amb set unitats tecnològiques disponibles per a tots els investigadors del centre. Aquí observem la unitat de seqüenciació. A més, dins la seva estructura hi ha el CRG-CNAG, la segona plataforma de seqüenciació més gran d'Europa.

LA SEQÜENCIACIÓ GENÈTICA JA ÉS HABITUAL ALS HOSPITALS, SOBRETOT PER DECIDIR EL TRACTAMENT DELS PACIENTS DE CÀNCER



CRG

que les investigacions dels sol·licitants no compleixin el consentiment signat per la persona donant.

Biologia de sistemes

En el marc del quart programa del CRG, els grups de recerca miren d'entendre processos biològics complexos a través de models computacionals. Per exemple, el grup de Luis Serrano, director del CRG, estudia el micoplasma –un bacteri– per saber com funciona a tots els nivells. Amb les dades obtingudes han creat un model computacional d'aquest bacteri que, com hem dit, permet avançar ràpidament en la recerca experimental. Què passaria si estigués en un ambient pobre de nitrogen? El programa informàtic és capaç de simular-ho. Altres grups utilitzen aquesta metodologia per entendre millor sistemes concrets dels organismes, com l'olfacte o la formació de les extremitats superiors humanes.

El programa de biologia de sistemes del CRG és una unitat compartida amb investigadors del Laboratori Europeu

de Biologia Molecular (EMBL). Aquesta sinergia ha fet possible que l'EMBL valori obrir la seva propera seu al Parc de Recerca Biomèdica de Barcelona. En té diverses, cadascuna especialitzada en una àrea concreta, però fa trenta anys que no se n'obre cap de nova. El consell de l'EMBL ja ha aprovat l'obertura de la nova seu a Barcelona i el projecte està esperant l'aprovació del Consell de Ministres espanyol.

Traduint-nos

El Centre Nacional d'Anàlisi Genòmica (CNAG-CRG) és una infraestructura d'àmbit estatal que des de l'any passat està integrada en el CRG i dona servei a tots els investigadors que ho necessitin. Amb setze màquines seqüenciadores i més de mil dos-cents processadors, és la segona plataforma de seqüenciació més gran d'Europa.

El famós Projecte Genoma Humà va necessitar milers de científics de tot el món i més de deu anys per seqüenciar els parells de bases del genoma humà.

Només tretze anys després, les màquines més avançades poden fer la feina en dues o tres hores. Luis Serrano ens explica que, en un món ideal, el procés encara hauria de ser més ràpid. Per exemple, seria molt útil que durant l'operació d'un tumor el cirurgià pogués agafar-ne una mostra i seqüenciar-ne l'ADN en uns minuts. Les cèl·lules cancerígenes tenen un ADN diferent del de les sanes, i conèixer-ne la mutació específica ajudaria a decidir si cal extirpar més o menys quantitat de teixit.

La seqüenciació, fins fa poc molt limitada a l'àrea de la investigació, cada cop tendeix a utilitzar-se més per fer diagnòstics i escollir tractaments. Actualment ja és una pràctica habitual als hospitals, sobretot amb els pacients de càncer, ja que ajuda a determinar quin tractament serà més efectiu. Segons Serrano, d'aquí a uns anys se seqüenciarà el genoma de tots els nens que neixin. Hi ha un cert tipus de malalties genètiques que es poden evitar si se segueixen uns hàbits determinats abans no sigui massa tard. Cal tenir en compte que l'estil de vida pot modificar cap a bé o cap a mal el nivell d'expressió de molts gens.

Fa uns anys, explica Serrano, el repte de la genètica era poder arribar a seqüenciar un genoma humà per mil euros. Ja quasi estem en aquest punt, en costa mil cinc-cents. Ara cal millorar en la comprensió de les dades. "L'objectiu és poder llegir el codi genètic d'una persona i saber predir correctament quines seran les seves debilitats i si poden o no modificar-se amb els hàbits de vida", diu el director del CRG. No és gens fàcil, encara som molt lluny d'aquesta quimera. Pel que fa a la comprensió de gens individuals, Serrano estima que hem fet entre un vint i un trenta per cent del camí. Però els gens no sols tenen un sentit individual sinó que creen complexes xarxes d'interaccions. A nivell global, Serrano estima que el nostre coneixement només està una mica per sobre del zero per cent. Vaja. Caldrà seguir investigant per poder traduir del tot el sofisticat idioma en què estan escrits tots els éssers vius. ●