

“La recerca en biologia s’aturaria en poques hores sense els càlculs computacionals”

Entrevista a **Modesto Orozco**, premiat amb una ERC Advanced Grant i un ICREA Acadèmia

FOTOS: JORDI PARETO

Modesto Orozco treballa a l'Institut de Recerca Biomèdica (IRB Barcelona) i al Departament de Ciències de la Vida del Barcelona Supercomputing Center, centrant-se en la simulació dels sistemes biològics, per representar, entendre i predir el comportament dels sistemes vius mitjançant algoritmes teòrics. Recentment, la seva recerca ha estat reconeguda amb un ICREA Acadèmia de la convocatòria de 2011 i una ERC Advanced Grant del Consell Europeu de Recerca pel projecte SimDNA, de simulació de l'ADN per entendre els mecanismes que controlen l'expressió dels gens. Ambdós reconeixements, que l'omplen d'orgull i el motiven, ajudaran a finançar els seus treballs que s'encaminaran en els propers anys cap a la simulació sistèmica de proteïnes i àcids nucleics per conèixer el funcionament de la natura i usar-lo en benefici de la societat.

Enguany ha obtingut una ERC Advanced Grant i un ICREA Acadèmia de la convocatòria de 2011. Què suposen per a vostè aquestes dues distincions?

Crec que aquests dos premis tenen dos vessants. El primer és el de reconeixement a la feina que has fet, que t'omple d'orgull i et motiva, i el segon, és el de finançament de la pròpia recerca. Aquests premis tenen darrera institucions molt prestigioses que confien en el que faràs, en aquest cas en els propers 5 anys, i volen finançar-ho perquè creuen en la teva recerca. En el cas del Consell Europeu de Recerca (ERC, en anglès), representa una injecció de diners molt gran que pràcticament garanteix durant cinc anys la recerca del grup. **En concret, l'ajut de l'ERC ha premiat el seu projecte Simulació de l'ADN a multiescala avançada: SimDNA. Quin és l'objectiu d'aquest projecte?** És un projecte que vam presentar en l'apartat de física i el que planteja és desenvolupar un conjunt d'eines per si-

mular l'ADN des de l'escala atòmica o subatòmica a l'escala del cromosoma sencer. Volem aplicar aquestes tècniques sobretot a temes d'estructuració de la cromatina i especialment a la relació amb la funció cel·lular.

És un projecte teòric, si bé té un component de validació experimental, que es divideix en tres grans parts. La primera és aquest desenvolupament de mètodes teòrics i algoritmes per mitjà de simulació computacional en diferents escales de resolució, la segona és la implementació de programes computacionals i, la darrera, la validació experimental de les prediccions.

Amb el projecte SimDNA es posaran a disposició de la comunitat internacional una sèrie d'eines computacionals que permetran entendre el comportament dels àcids nucleics. L'aplicació més directa es troba en el camp de la regulació de l'expressió de gens i en epigenètica, és a dir, les eines desenvolupades a SimDNA permetran entendre els mecanismes que controlen l'expressió dels gens.



MODESTO OROZCO

És catedràtic de Bioquímica i Biologia Molecular de la UB i director del Grup de Modelització Molecular i Bioinformàtica a l'IRB Barcelona. També dirigeix el Departament de Ciències de la Vida del BSC, és director del Programa Conjunt en Biologia Computacional (IRB Barcelona i BSC) i del node de bioinformàtica estructural de l'Institut Nacional de Bioinformàtica (INB). Destaca com un dels principals experts europeus en simulació de sistemes biològics i és un referent internacional en l'estudi teòric de sistemes macromoleculars, especialment dels àcids nucleics (DNA i RNA).

Al llarg de la seva carrera, ha publicat més de 330 articles científics i ha desenvolupat programes computacionals i algorismes dels quals es pot beneficiar avui tota la comunitat científica internacional. Orozco participa en els consells editorials de les publicacions científiques internacionals més prestigioses del seu camp de coneixement. També forma part de panels d'avaluació a Espanya, a Europa i als Estats Units i és consultor de diverses companyies farmacèutiques. El treball d'Orozco ha estat reconegut amb diferents premis nacionals i internacionals i des de fa quatre anys la seva recerca rep finançament de la Fundació Marcelino Botín.

Per què va decidir dedicar-se a la química i en concret a la química computacional?

Sempre havia volgut ser bioquímic. M'agradava també molt la física, la química i les matemàtiques. Vaig decidir llavors començar la carrera de química. En aquell moment, la carrera tenia una durada de cinc anys, tres eren comuns i els dos darrers anys d'especialitat. El tercer any de carrera vaig tenir un professor de física i química que era extraordinàriament bo, en Joan Bertran (vegeu *Teraflop 60*), que de fet em va fer veure tot el potencial que tenia la química teòrica i vaig estar dubtant moltíssim si dedicar-me a la química computacional o a la bioquímica. Tant en Javier Luque, que ha estat el meu soci durant molts anys, com jo mateix vam decidir finalment fer bioquímica. En acabar els estudis de bioquímica vam veure clar que el que volíem era dedicar-nos a aplicar la química computacional a l'estudi de sistemes biològics.

Com ha comentat, la seva recerca se centra en la simulació de sistemes biològics. En què consisteix exactament i quina finalitat té?

La simulació de sistemes biològics és un camp molt ampli que engloba des de la química a la biologia computacional i que, en definitiva, el que intenta és entendre els sistemes vius, que són molt més complexos que qualsevol sistema químic, però que es regeixen per les mateixes regles.

El que fem és, doncs, desenvolupar i adaptar metodologies per poder treballar en aquestes situacions de gran complexitat. Per exemple, pràcticament tota la química orgànica es desenvolupa en l'escala del nanòmetre, que són 10^{-9} metres, en canvi quan treballes amb sistemes biològics et mous des de l'escala del nanòmetre o fins i tot per sota, l'escala de angström (10^{-10} metres) o del picòmetre (10^{-12} metres), fins a escales de mil·límetres o fins i tot metres.

Com es pot veure, l'abast en l'estudi dels sistemes biològics és molt ampli. L'objectiu és el mateix que pot tenir un químic teòric, entendre la naturalesa d'un sistema a partir de primers principis, però el tipus d'eines que nosaltres hem de desenvolupar i aplicar per fer aquest estudi són molt més diverses perquè els sistemes biològics són complexos i tenen aquesta gran variabilitat d'escala que he esmentat.

Així, la nostra recerca se centra en

l'estudi dels processos de reconeixement cel·lular d'importància biològica, tant des del punt de vista metodològic com aplicacional. Els temes principals són les estructures no estàndard dels àcids nucleics, les propietats dinàmiques de les macromolècules i el reconeixement proteïna-ligand i proteïna-proteïna.

Al llarg de la seva carrera ha desenvolupat un conjunt de programes computacionals i algorismes amb els quals es beneficia tota la comunitat científica internacional. Quin paper juga avui dia la química computacional en la nostra societat?

Crec que té molta més importància que el que la gent es pensa, sobretot si entenem la química computacional en la seva concepció més àmplia. Per exemple, el

No hi ha cap fàrmac que surti al mercat que no hagi estat modelat en algun moment per un ordinador

cas més paradigmàtic són tots els estudis de desenvolupament de fàrmacs. No hi ha cap fàrmac que surti al mercat que no hagi estat modelat prèviament en algun moment per un ordinador. Un altre exemple són tots els estudis de nous materials que també es beneficien de la química computacional per ser desenvolupats. En el camp de la biologia el seu paper és inqüestionable. S'ha calculat que si les bases de dades que contenen informació derivada a partir de càlculs computacionals es tanquessin, la recerca en biologia i biomedicina a tot el món s'aturaria en unes poques hores. Ara mateix, té un paper molt troncal aquesta ciència que és la química computacional difusa, que va des de la física fins als temes biològics. Vint anys enrere la química computacional era una disciplina interessant, però que estava com al marge de la investigació troncal en biologia. Ara està intrínsecament lligada i al capdavant de la investigació actual.

És catedràtic de Bioquímica i Biologia Molecular a la UB i també desenvolupa diferents activitats de direcció a l'IRB Barcelona, al Barcelona Supercomputing Center (BSC) i al node de Bioinformàtica Estructural de l'Institut Nacional de Bioinformàtica

(INB). Com es complementen totes aquestes activitats entre elles?

Tots aquests càrrecs són honors. De fet, totes aquestes activitats porten al mateix fi. Jo em dedico a investigar i a fer possible que altres investigadors ho facin en millors condicions i això ho faig des de diferents paraigües institucionals que estan molt lligats. L'amo del meu temps és la UB, qui paga el meu salari. Estic a l'IRB i al BSC per un conveni de la UB amb l'Institut, i d'aquest amb el BSC que em permet fer recerca i dirigir un conjunt de gent, unes 120 persones entre informàtics, biòlegs i químics computacionals. Crec que aquesta col·laboració és un cas modèlic de com diferents institucions poden unir forces per un objectiu de recerca comú i no competint entre elles.

Pocs anys després de doctorar-se, va començar a utilitzar la infraestructura del CESCO per donar suport a la seva recerca. N'ha estat usuari més de 15 anys. Com valora aquesta col·laboració?

La valoro molt positivament. Al principi, tenir el CESCO va ser clau, sobretot per als investigadors que, com jo, depenem de l'anomenat *performance computing*. Durant els seus primers anys de vida, el

CESCO era un centre de referència al que un investigador podia enviar càlculs que d'altra manera no podia realitzar. També va ser pioner en implementar el concepte que ara està a tota Europa dels nivells de computació Tier-0, Tier-1, Tier-2... El CESCO va ser un dels primers que va introduir aquest concepte, que ve a ser que un investigador ha de tenir els seus propis recursos computacionals però que a

En els propers 10 anys estarem en situació de prendre el lideratge en entendre els sistemes biològics

més és necessari que disposi d'altres recursos aliens més potents, com els que ofereixen els centres de supercomputació, per poder ser competitiu en recerca computacional.

A més, el CESCO també ha estat molt important en prendre iniciatives per al desenvolupament de nous fàrmacs amb el seu Servei de Disseny de Fàrmacs. Crec que va ser una iniciativa molt bona

que ens ha ajudat moltíssim, sobretot als qui treballem en e-Ciència.

Quins reptes li agradaria assolir en la seva recerca en els propers 10 anys?

Quant a l'estudi teòric de sistemes biològics estem en un moment en què hem passat aquella etapa que el que podíem fer era explicar el que ja es coneixia, trobar una explicació racional amb les dades experimentals o bé llançar hipòtesis. Ara estem tenint un impacte directe sobre l'experimentació, de fet el nostre programa de recerca IRB Barcelona-BSC disposa d'un laboratori propi de validació experimental.

Crec que en aquests propers 10 anys estarem en situació de prendre el lideratge en entendre els sistemes biològics, ja no d'entendre cap a on van els estudis experimentals, sinó generar el tronc principal de la recerca. En el cas concret del nostre grup, les principals línies que portarem en els propers anys seran tot el que està relacionat amb les propietats de DNA, l'estructura i funcions de la cromatina, i el que es coneix com a dinàmica de proteïnes; camps en el que podem fer aportacions que ja no només valorin els químics teòrics sinó el conjunt de bioquímics, químics i biòlegs. ■

Seqüenciat el genoma del tomàquet

El Consorci del Genoma del Tomàquet ha seqüenciat i assemblet l'ADN genòmic d'aquesta espècie hortofrutícola, concretament el d'una varietat domesticada, la Heinz 1706. Científics de diferents centres europeus, entre els quals hi ha Modesto Orozco, David Torrents i Xavier Pastor, investigadors del programa conjunt del BSC-CNS i de l'IRB Barcelona, han combinat esforços per obtenir la seqüència d'aquesta planta i que ha publicat la revista *Nature*.

En aquest estudi, el genoma del tomàquet s'ha comparat amb diferents espècies properes, entre les quals hi ha una varietat salvatge de tomàquet, la patata i el raïm. Aquestes comparacions han posat al descobert que el genoma del tomàquet, igual que el d'altres plantes, ha patit duplicacions completes durant la seva evolució. Aquests fenòmens són molt rellevants des

d'un punt evolutiu ja que ofereixen un marc òptim perquè els gens afectats per aquestes duplicacions explorin noves funcions que contribueixin a l'evolució i millora de l'espècie.

En concret, en el tomàquet s'han observat evidències que suggereixen que el seu genoma ha patit dues rondes de triplicacions en els últims 120 milions d'anys de la seva evolució. Una d'elles va tenir lloc, probablement, abans que els raïms i els tomàquets evolucionessin de forma diferent. L'altra, la més recent, es remunta fins a un ancestre comú entre el tomàquet i la patata. La seqüenciació i publicació de la seqüència genòmica del tomàquet crea les bases moleculars per estudiar-lo i investigar formes de conreu efectives que permetin que resisteixi plagues i l'escassetat d'aigua en regions més pobres i àrides.



En aquest projecte, a banda de l'IRB Barcelona i el BSC, també hi han participat el CNAG, l'Institut Nacional de Bioinformàtica, el Centre de Regulació Genòmica, l'Institut de Biologia Molecular y Celular de Plantas (CSIC-UPV), l'Institut de Hortofruticultura Subtropical y Mediterrànea La Mayora (UMA-CSIC), ICREA i l'empresa Sistemes Genòmics del Parque Tecnológico de Valencia.