

GENOMES I EVOLUCIÓ: IMPACTE EN LA BIOLOGIA ACTUAL

MONTSERRAT AGUADÉ,¹ JAUME BERTRANPETIT² I JULI PERETÓ³

¹ *Departament de Genètica i Institut de Recerca de la Biodiversitat, Universitat de Barcelona*

² *Institut de Biologia Evolutiva (UPF-CSIC)*

³ *Institut Cavanilles de Biodiversitat i Biologia Evolutiva,
i Departament de Bioquímica i Biologia Molecular, Universitat de València*

Adreça per a la correspondència: Jaume Bertranpetit. Institut de Biologia Evolutiva,
Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra.
C. del Dr. Aiguader, 88. 08003 Barcelona. Tel.: 933 160 840.
Adreça electrònica: jaume.bertranpetit@upf.edu.

RESUM

L'estudi dels genomes és un tema central en la biologia del segle XXI. Tenim eines per fer-ne la descripció de la seqüència que avancen ràpidament i un repte important ara mateix és la interpretació de l'allau d'informació genòmica. La comprensió dels genomes comença amb la lectura comparada: les diferències entre els genomes dels diferents individus d'una població o una espècie i les diferències entre els genomes de diferents espècies són la clau per a la reconstrucció del procés evolutiu, per entendre els canvis funcionals i per a la integració de la informació genètica en complexes xarxes d'interacció funcional. Els genomes, doncs, recullen bona part de la informació dels éssers vius, i ens permeten entendre detalladament tant l'evolució de les espècies com les implicacions de l'adquisició de noves funcions i de l'emergència de noves adaptacions que, mitjançant la selecció natural, s'han anat esdevenint durant la història, la diversificació i el desplegament de la complexitat de la vida.

Paraules clau: genoma, evolució, selecció natural, canvis adaptatius, biologia de sistemes.

GENOMES AND EVOLUTION: IMPACT ON CURRENT BIOLOGY

SUMMARY

The study of genomes is a central theme in 21st century biology. We now have the tools that enable us to advance rapidly to obtain and describe the sequences. However, a major

challenge is to interpret the avalanche of genomic information. The understanding of genomes begins with a comparative reading: the differences between the genomes of different individuals in a population or in a species, and the differences between the genomes of different species, are the key to reconstruct the evolutionary process, to understand functional changes, and for the integration of genetic information into functional complex interaction networks. The genome thus includes much of the information of living beings, and it allows us to understand in detail both the evolution of species and the implications of the acquisition of new functions and the emergence of new adaptations, which, by natural selection, have been taking place along the history, the diversification and the unfolding of the complexity of life.

Key words: genome, evolution, natural selection, adaptive change, systems biology.

INTRODUCCIÓ

Vivim en l'època dels genomes i no solament en biologia, ja que formen també part d'un debat d'abast social. En els darrers anys la paraula *genoma* s'ha estès en la parla i ja forma part del vocabulari usual de molta gent. En l'imaginari social el genoma és la part més essencial dels éssers vius, la part fixa, permanent, de vegades fins i tot determinística, dels éssers vius, i alhora té una connotació física, orgànica, arrelada en la natura i que denota les bases químiques i estrictament biològiques dels éssers vius. El genoma, doncs, apareix com una explicació natural i comuna de tots els éssers vius, basada en els fonaments materials del DNA.

Tots podem entendre la importància de la informació continguda en el genoma si pensem en un exemple tan senzill que pot semblar banal: d'una dona embarassada sortirà un descendent que serà alhora, un ésser viu, un ésser humà i un humà concret. La informació per a tot això està en les dues còpies del genoma (una de la mare i una del pare) que conté l'òvul fecundat (zigot), atès que el nou ésser viu es desenvoluparà segons uns plans escrits en la tinta perdurable de les instruccions que hi ha en aquestes còpies concretes del genoma humà. De ben cert que un dels gran reptes

d'un sistema viu és poder desplegar i materialitzar en quatre dimensions, l'espai i el temps, el programari lineal escrit en el genoma, és a dir, la informació genètica. Per bé que actualment sabem moltes coses dels genomes, encara no tenim els coneixements que ens permetin generar un sistema informàtic que sigui capaç, *in silico*, de llegir les instruccions del genoma, entendre-les, executar-les i predir-ne les conseqüències últimes en un entorn concret. En altres paraules, no coneixem els detalls de com l'execució d'unes instruccions determinades escrites al genoma donarà lloc a un ésser viu (viable, que té vida pròpia i independent), el qual pertany a una espècie determinada (i, per tant, que té les instruccions concretes que canalitzen el procés d'execució d'aquella espècie concreta, sigui un llevat, una alzina o un humà), i que serà un individu concret, un individu ben únic, tant en les bases bioquímiques que el configuren com en la manera de ser i en el seu comportament. Tot i no saber encara com es pot interpretar la informació genètica en les dimensions temporal i espacial i en la seva interacció amb l'ambient, entenem que el genoma conté aquesta informació i que actualment la biologia s'obre pas en la immensitat de dades genòmiques per arribar a entendre el missatge que hi ha escrit.

En la seqüència del genoma hi podem, doncs, buscar moltes coses, perquè sabem que hi ha d'haver molta informació, i la biologia actual va descabdellant els secrets amagats en la lectura bioquímica del llibre de la vida. Perquè més enllà de la llarga llista de lletres que fem servir per representar la informació genètica, comprendre les implica el coneixement físic, orgànic, bioquímic i causal de la vida. La seqüència de nucleòtids (els graons bàsics de la informació genètica que representem per A, C, G, T) és el punt de partida bàsic de les instruccions contingudes en el genoma i moltes vegades (per raons històriques i pel biaix que ens dóna el coneixement actual) prenem aquesta seqüència com a base per entendre les funcions inscrites en el genoma en relació amb la síntesi de proteïnes. Però ara sabem que hi ha molt més que això i que la informació del genoma va més enllà. D'una banda s'amplia contínuament l'espectre estructural i funcional de l'RNA en multitud de formes, fins fa poc insospitades; de l'altra s'aprecien cada vegada més les modificacions químiques, temporals o permanents, dels nucleòtids o de les proteïnes que acompanyen el DNA, en el que s'anomenen *variacions epigenètiques*, algunes de les quals són claus per determinar les estructures que adopta el DNA en la compactació de la cromatina i en els cromosomes, fets que tenen implicacions funcionals extraordinàries.

En el genoma hi ha, doncs, la clau per entendre molts processos bàsics de la vida. Però també hi ha la clau per a preguntes lligades no ja als productes gènics (RNA o proteïnes) o a la seva regulació, sinó al significat de la diferència en biologia. No hi ha evolució sense variació i les diferències bàsiques entre els éssers vius tenen les seves bases en la informació continguda en el genoma que ha estat modelada en el transcurs del temps per les diverses forces evo-

lutives. Entendre les diferències entre genomes implica el reconeixement dels processos evolutius que les han generades. De fet, la diversitat de la vida en un moment determinat (normalment ens fixem en el present perquè és on tenim la pràctica totalitat de la informació) no és res més que la conseqüència del procés de diversificació que ha implicat l'evolució en el transcurs de les generacions: la diversitat en l'espai (biodiversitat) i en el temps (evolució) són dues cares de la mateixa moneda i l'estudi dels genomes ens il·lumina sobre totes dues. Podem determinar les diferències entre els éssers vius actuals i comprendre com aquestes diferències s'han incorporat durant un procés de canvi al llarg del temps en la informació continguda en els genomes, per la dinàmica mateixa del DNA com a molècula i per l'efecte de l'atzar però, sobretot, pel procés d'adaptació per selecció natural durant l'evolució. Els genomes que veiem avui són els supervivents d'un llarg procés de producció de canvis i un estricte filtratge de la selecció natural, que ha afavorit algunes variants (les quals esdevenen més freqüents i es fixen en les poblacions), o que n'ha eliminat d'altres si significaven una reducció de les possibilitats de sobreviure i reproduir-se dels individus que les han portades. Altra-ment en aquests genomes també s'hauran arribat a fixar, per l'efecte de l'atzar, alguns dels canvis produïts que no afecten directament aquestes possibilitats (mecanisme que anomenem *deriva genètica*).

Podem comparar genomes entre els individus d'una espècie o d'espècies diferents. En qualsevol cas, les lliçons que podem aprendre són moltes i fan referència als processos biològics bàsics (la regulació de les funcions codificades pel genoma, per exemple), a entendre la quantitat de diferències en regions del genoma sense implicacions funcionals (que ens permet recons-

truir els processos evolutius de poblacions i d'espècies), i a entendre les bases de l'adaptació a través del reconeixement de l'acció de la selecció natural. Però aquestes lectures comparades dels genomes no ens donen la imatge completa de la complexitat desplegada en els éssers vius. La comprensió profunda de la descodificació funcional d'un genoma requereix l'estudi de les xarxes d'interaccions moleculars dels genomes amb els seus productes (entre proteïnes, entre proteïnes i genoma, entre proteïnes i molècules no proteiques, etc.), interaccions moleculars que estan en la base de la vida i que han estat modelades pels processos evolutius. El repte de la biologia contemporània és que avui la limitació no està en els mètodes i les tècniques per obtenir dades, en grans quantitats en totes les escales dels sistemes vivents, sinó més aviat en els conceptes, les teories i els models que ens ajudin a treure'n l'entrellat de tot plegat. Reconstruir i entendre les xarxes d'interacció planteja grans reptes. D'una banda, l'estudi purament arquitectònic, estructural. De l'altra, la visió dinàmica, durant la vida d'un individu, en diverses circumstàncies o en el temps evolutiu, esbrinant com les xarxes creixen en complexitat o són demolides per la pèrdua de la informació genètica. Les nocions de modularitat, redundància, fragilitat o robustesa de les xarxes tenen èxit entre els investigadors de l'anomenada biologia de sistemes.

Una singularitat dels éssers vius, que els diferencia de manera fonamental de les màquines artificials, és que les causes que determinen les funcions cel·lulars són internes. Dit en termes bioquímics: el proteoma (el conjunt de proteïnes sintetitzades en un moment donat) genera el metaboloma (el conjunt total de metabòlits) i, alhora, n'és un producte d'aquest. Aquesta circularitat metabòlica, o recursivitat, és

una de les propietats més genuïnes de la matèria viva i cal tenir-la present tant per entendre el funcionament fisiològic o patològic, com per projectar modificacions tecnològiques d'un ésser viu o, fins i tot, dissenyar-ne la fabricació. Així doncs, interrogar els genomes i integrar la informació que n'obtenim ajuda a respondre moltes qüestions bàsiques en la biologia actual. La comparació de genomes o l'estudi de les xarxes dels seus productes ens permet investigar aspectes molt diferents de la biologia, i endinsar-nos en terrenys insospitats, nous i amb possibles respostes que havien romàs estancades dins la biologia, o ni tan sols sospitàvem els camins que ens podien portar a il·luminar-los. De fet, actualment la comparació de seqüències de material genètic forma part del treball quotidià de la recerca en la majoria de les branques de la biologia. Així «fer un BLAST» o fins i tot «blastejar», és una expressió comuna que consisteix a llançar una seqüència contra les bases de dades existents i obtenir les seqüències més semblants a la nostra amb una estimació estadística del grau de semblança; si fer això té sentit i poder analític és perquè hi ha un procés evolutiu subjacent. Dins de totes aquestes possibilitats ens centrarem en tres aspectes, tots nous i amb gran impacte en la biologia actual:

- a) Els genomes com a base de la reconstrucció de l'evolució.
- b) L'adaptació i la selecció natural en la gènesi i evolució dels genomes.
- c) La integració de la informació de genomes sencers.

Finalment, el coneixement de la dinàmica dels processos vius i de la informació bàsica que els dirigeix és una de les grans noves eines de la biologia actual que, partint dels processos moleculars, ens permet integrar la informació fins a altres nivells com el del fenotip dels individus (en la dis-

secció de les bases genètiques dels caràcters complexos) o de poblacions, i ens mostra com processos adaptatius generals (de població, d'espècie o d'altres nivells d'organització) tenen lloc modelant les seves bases moleculars.

LA RECONSTRUCCIÓ DE L'EVOLUCIÓ

Un dels usos més estesos i fàcils d'entendre de l'estudi dels genomes és reconstruir l'evolució de les formes vives que hi ha a la Terra; és a dir, fer servir el procés evolutiu del genoma com a marcador i com a rellotge (tal com veurem més endavant). Cal, per a aquest propòsit, entendre la dinàmica de les regions neutres del genoma, és a dir, l'evolució del material genètic independentment de la selecció natural. Aquesta aproximació ens permet fer servir el canvi genètic com a eina per reconstruir l'evolució, tant en l'aspecte més ampli de la diversitat de la vida (l'arbre de la vida), com en grups d'éssers vius concrets o, dins d'una espècie, entendre la diversitat genètica com a resultat d'un procés demogràfic i històric.

La dinàmica del genoma

En la reconstrucció dels processos històrics (siguin del tipus que siguin) podem fer servir informació de dos tipus: servir-nos de restes del passat, que ens vagin mostrant les diferents formes al llarg del temps (els fòssils en els éssers vius) o bé estudiar la diversitat actual i servir-nos del coneixement que podem tenir de la dinàmica del canvi (la velocitat del canvi biològic en aquest cas). Tradicionalment s'han emprat els fòssils en la reconstrucció de la història de la vida, però cada vegada està tenint

més impacte l'ús de la informació de les formes vives actuals. Però no en l'aspecte extern tal com ha estat exclusivament descrit durant segles: la morfologia no canvia en el transcurs del temps d'una manera contínua, sinó que la velocitat de canvi depèn fortament de les pressions de la selecció natural. Així, espècies en hàbitats molt constants s'han mantingut amb aspecte molt semblant durant molts milions d'anys i d'altres, sotmeses a processos adaptatius constants, han anat canviant a fort ritme. Així, tot i que es poden reconèixer innovacions evolutives que ens marquen un inici únic, no podem mesurar la velocitat del canvi morfològic per tenir una reconstrucció temporal acurada. Però sí que ho permet l'estudi del material genètic.

La mutació, com a canvi en la informació genètica que es transmet d'una generació a la següent, és el veritable motor de l'evolució, el procés que aporta la matèria primera inicial de la variació i sobre el qual altres processos (selecció natural i deriva genètica sobretot) en decidiran el futur, envers la desaparició o expansió. Si ens fixem en regions dels genomes que no portin informació rellevant per a l'organisme (i en eucariotes és la major part del genoma), el procés de canvi per mutació estarà determinat fonamentalment per les propietats fisicoquímiques del material genètic i l'atzar, sense que hi actuï la selecció natural, i el procés es pot fer servir com a rellotge molecular de l'evolució, propietat que implicaria un procés constant de canvi i que, per tant, marcaria la quantitat de diferència com a funció del temps. De fet, tot i que hi ha certa incertesa en les mesures exactes, podem dir que coneixem força bé les taxes de canvi genètic i, per tant, si observem un cert nombre de diferències entre les seqüències de la mateixa regió de dues espècies, podem estimar el temps de separació dels avantpassats dels dos llinatges. La compa-

ració es pot fer entre conjunts amplis d'espècies i la mesura del nombre de diferències ens permet estimar la topologia del procés evolutiu (en quin ordre es produí la separació dels diferents grups d'ancestres) i una estimació del temps del procés. Tot això no és, en el detall, senzill, per bé que els principis generals són fàcils d'entendre i aplicar.

L'estudi de la diversitat de la vida a partir de la seqüència del genoma ens porta grans transformacions en la nostra concepció de l'estructura de la biodiversitat, tant si volem entendre el conjunt de la vida com si anem a grups concrets, molts dels quals han vist refetes les classificacions que els naturalistes havien fet i revisat contínuament durant segles. I en els casos en què la morfologia era de poca ajuda, especialment en éssers vius senzills com procariotes o eucariotes unicel·lulars (protists), ha permès reconèixer els grans grups d'éssers vius en una visió estrictament evolutiva: les classificacions dels éssers vius poden ara reflectir les semblances i diferències de manera objectiva i respondre al procés evolutiu que les ha produïdes.

Entendre la biodiversitat dins del procés d'evolució

Tot i que ja és conegut de fa anys, cal esmentar que la reconstrucció dels nodes més basals de la diversitat de la vida (vegeu la figura 1; López-Garcia i Moreira, 2008) ens porta a reconèixer tres grans grups: els bacteris (a vegades anomenats *eubacteris*), els arqueus i els eucariotes. Dins dels eucariotes és interessant constatar la gran proximitat entre tres grups independents i ben coneguts: fongs, plantes i animals. Sens dubte aquesta estructura de la diversitat ens emmarca de manera molt diferent la percepció de la diversitat de la

vida i quins són els grans plans d'organització; és per això que actualment la genòmica comparada està esmerçant molt esforç no ja a reconstruir quin fou el procés evolutiu, sinó quines innovacions moleculars permeteren les grans novetats evolutives, com l'aparició del nucli cel·lular o la pluricel·lularitat. I anant a grups més concrets s'està avançant molt a comprendre les bases genètiques de les diferències fonamentals entre espècies o fins i tot entre individus.

La feina de reconstruir l'evolució no està mancada de dificultats ni polèmiques. Algunes dificultats són ben intrínseques als mètodes i inclouen aspectes com la quantitat de seqüència que cal comparar; o com se sap que dues seqüències que presenten un parentiu evolutiu (i que anomenem homòlogues) i amb un alt grau de semblança en la seva seqüència en espècies diferents, vénen de la mateixa seqüència en un avantpassat comú (que és el que anomenem *seqüències ortòlogues* o *gens ortòlegs*) o són producte de duplicacions dins del genoma prèvies a l'especiació (que anomenem *seqüències paràlogues* o *gens paràlegs*); o si la transferència horitzontal de gens (entre espècies de vegades molt llunyanes) ha estat un fenomen abundant o no; o sobre els mètodes numèrics en l'establiment de la mesura de la diferència. Totes aquestes polèmiques tenen una dinàmica, perden virulència i finalment es van acceptant els resultats que l'anàlisi comparativa del genoma proporciona.

Un dels casos que més controvèrsia va produir en el seu moment va ser la proposta de la posició dels humans a la natura i molt especialment la gran similitud genètica amb els ximpanzés (vegeu la figura 2). Actualment ja no ens estranya gaire, i no ens sobta gens, veure un arbre filogenètic dels primats en què la branca humana és ben curta enmig de les d'altres espècies, fet

indicatiu que la singularitat humana s'as-solí en un temps relativament breu (entorn de 6 milions d'anys). Aquest fet ben establert ha obert altres interrogants, ben interessants: quines són les adaptacions específiques dels humans que han donat lloc a la singularitat de la nostra espècie? O en altres termes més interessants, preguntar-nos què podem trobar en el genoma que ens ajudi a respondre la pregunta de què fa humans els humans. Entendre això implica ja estendre la comprensió de l'evolució a l'adaptació i la selecció natural actuant sobre els genomes.

ADAPTACIÓ I SELECCIÓ NATURAL

Els éssers vius presenten moltes característiques que poden ser considerades adaptatives. La biologia evolutiva sempre ha estat interessada a identificar aquelles característiques que poden considerar-se veritables adaptacions, ja que han sorgit per canvi adaptatiu. Això vol dir que en un moment determinat els canvis genètics (o mutacions) subjacents a la nova característica van conferir un avantatge selectiu als seus portadors respecte als no-portadors. Aquest avantatge va conduir a l'increment de freqüència en la població dels individus

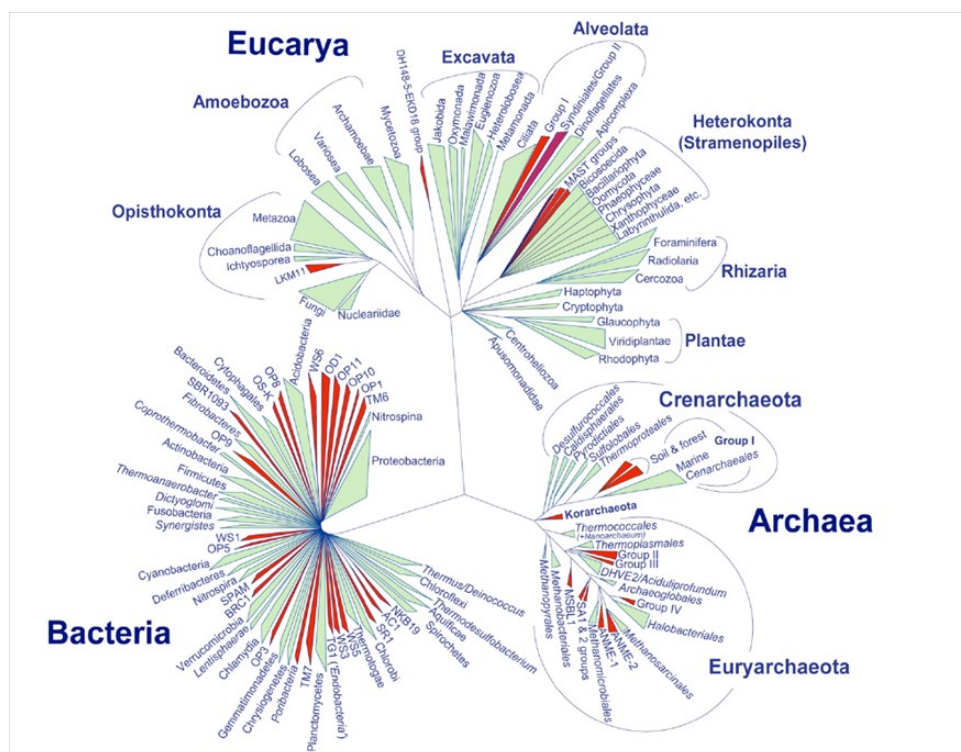


FIGURA 1 Arbre de tots els organismes vius, en què es veuen tres grans grups: bacteris, arqueus i eucariotes; aquests darrers comprenen, en branques molt properes, els fongs, i els metazous. L'arbre és molt simplificat, amb només una part dels grups d'eucariotes representats. Els triangles verds representen grups en què alguna espècie s'ha cultivat o descrit acuradament; els triangles vermells, en canvi, són llinatges molt divergents poc coneguts. Extret de López-García i Moreira, 2008.

amb la característica i en molts casos es va produir, ja sigui en una població o en el conjunt d'una espècie, el reemplaçament de la característica anterior per la nova.

La selecció natural és la força evolutiva que hi ha darrere de les adaptacions. De fet, la variabilitat genètica que sorgeix per mutació i recombinació és indispensable perquè pugui haver-hi canvi al llarg del temps. En aquest apartat, que està centrat en les adaptacions, parlarem dels canvis adaptatius, és a dir, dels que donen lloc a noves característiques adaptatives i que són promoguts per la selecció natural positiva (o adaptativa). És important indicar que la selecció positiva pot conduir a la fixació de la nova característica en la població (selecció direccional positiva), però que

també pot conduir-la a una freqüència intermèdia (selecció equilibradora) (vegeu la figura 3). A més, parlarem de la selecció purificadora que manté les adaptacions eliminant les variants deletèries (és a dir, les que disminueixen l'eficàcia biològica dels seus portadors) sorgides per mutació.

La comparació de la variabilitat present a les seqüències de DNA permet identificar l'acció de la selecció natural en l'evolució de les poblacions i espècies perquè hi deixa empremtes que podem reconèixer i destruir del rerefons de la variació neutra. Aquestes empremtes que l'acció de la selecció natural deixa en els genomes s'utilitzen sovint per inferir la funció de determinades parts del genoma. En aquest sentit, durant les darreres dècades s'ha fet servir

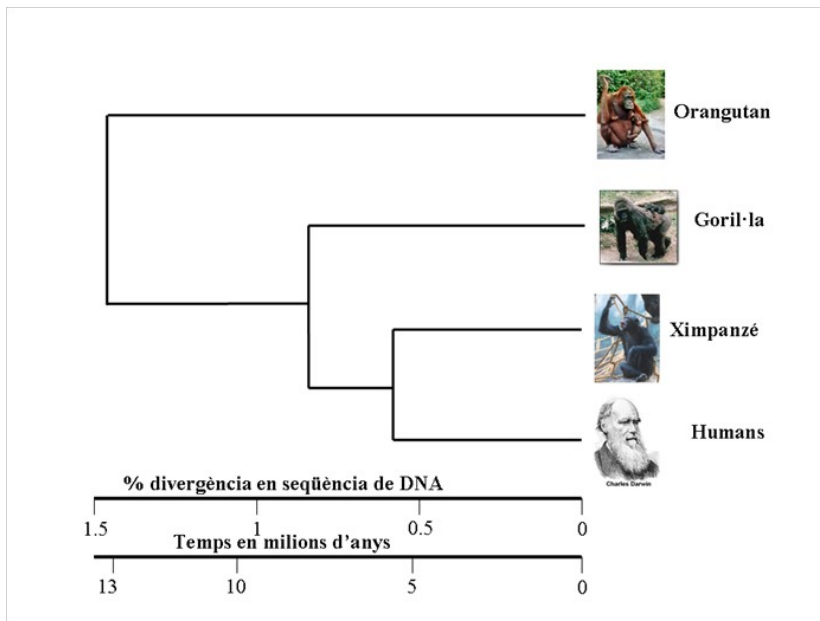


FIGURA 2 Filogenia dels humans i altres espècies properes obtinguda a partir de comparació de seqüències de DNA no codificant que han tingut una evolució neutra. En les abscisses la llargada de les branques indica el percentatge de diferència entre els diferents genomes (per exemple, entre humà i ximpanzé, 0,6 per cada branca, és a dir, un 1,20 % de diferència entre les seqüències de nucleòtids). La similitud genètica, per temps no gaire llargs, pot convertir-se en temps de divergència.

el grau de conservació entre espècies de les seqüències intergèniques per identificar elements funcionals en la regulació de l'expressió gènica (elements reguladors). En aquests estudis l'èmfasi s'ha posat en allò que està conservat per acció de la selecció purificadora i, per tant, en les regions del genoma que no canvien o ho fan poc. Però ~~en els darrers anys~~ s'hi ha incorporat una altra visió: si el que es pretén és saber com s'ha generat la biodiversitat i entendre quina importància funcional han tingut determinats gens i regions genòmiques, l'èmfasi s'ha de desplaçar envers la informació genètica que canvia i és, per tant, diferent entre espècies o entre individus d'una mateixa espècie.

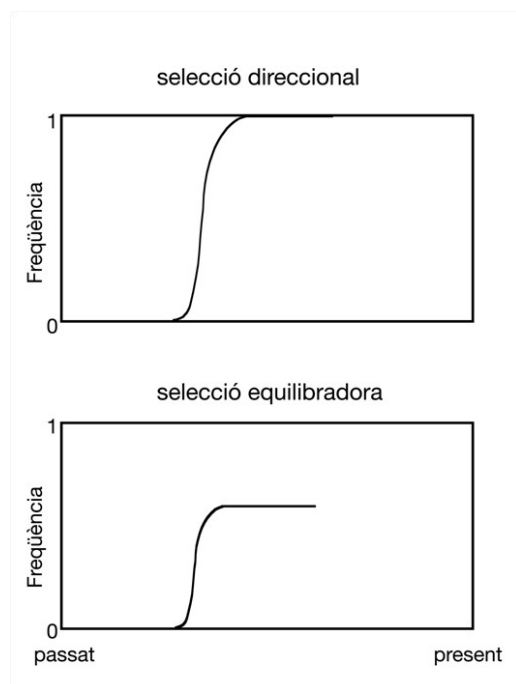


FIGURA 3 Representació esquemàtica de la trajectòria d'una mutació avantatjosa conduïda a la fixació (freqüència = 1) per la selecció direccional o a una freqüència intermèdia per la selecció equilibradora.

L'aproximació inicialment utilitzada per tal de desvetllar l'acció de la selecció natural en l'evolució de les molècules i dels organismes és l'anomenada *aproximació de gens candidats*, és a dir, de gens que per la seva implicació en una característica considerada adaptativa poden ser candidats a haver evolucionat per acció de la selecció positiva. Un cop identificats els gens candidats a partir dels resultats d'estudis genètics (i no específicament evolutius) es procedeix a seqüenciar-los en diverses espècies o en diversos individus de la mateixa espècie per tal de detectar l'empremta que la selecció natural deixa en les seqüències de DNA. ~~En cas~~ que el gen candidat presenti aquesta empremta, permet establir el caràcter adaptatiu del canvi que es va produir en el passat. L'obtenció de les primeres seqüències genòmiques (i no solament de gens o regions curtes) al començament del segle XXI ha propiciat el canvi d'aproximació que s'ha donat durant aquesta dècada en els estudis orientats a detectar canvis adaptatius i a identificar els gens implicats. De fet, la disponibilitat del genoma sencer d'una espècie i de tècniques d'alt rendiment per a l'estudi de la variabilitat nucleotídica (genotipatge i seqüenciació) ha possibilitat estendre aquests estudis a múltiples regions del genoma o a tot el genoma sense enfocar específicament en gens o regions de funció coneguda.

Empremtes selectives i escala temporal

Canvis adaptatius recents. Tal com ja s'ha indicat en la introducció, únicament una fracció de totes les mutacions que es produeixen tenen èxit evolutiu, en el sentit que assoleixen una certa freqüència en la població. Es pot entendre fàcilment que la majoria de mutacions deletèries (un cas ho són les causants de malalties) es perden

poc després d'haver-se originat, és a dir, que la seva freqüència passa a ser nul·la per acció de la selecció purificadora. Tanmateix, les mutacions neutres poden assolir una certa freqüència en el seu camí envers la fixació o la pèrdua per efecte de l'atzar (o deriva genètica). L'atzar té també un paper important en el destí de les mutacions avantatjoses, en el sentit que pot conduir a perdre-les quan, poc després del seu origen, es troben encara a molt baixa freqüència. Les mutacions avantatjoses que perduren són aquelles que s'escapen d'aquesta pèrdua per atzar, i la selecció pot conduir-les ràpidament a la fixació (selecció direccional positiva) o a una freqüència intermèdia (selecció equilibradora). Així, podem dir, de manera genèrica, que les mutacions que trobem en les poblacions, ja sigui a freqüències intermèdies o fixades, són les que han perdurat per l'efecte de

l'atzar i de la selecció positiva (vegeu la figura 4).

La trajectòria de les mutacions avantatjoses envers la fixació o freqüències intermèdies difereix de la de les mutacions neutres, que és molt més ràpida, per la qual cosa pot assolir freqüències elevades molt més de pressa (vegeu la figura 4). La manca de recombinació fa que l'increment de la freqüència d'una mutació avantatjosa en una població condueixi a l'increment de freqüència de les mutacions neutres adjacents, és a dir de tot el conjunt de variants properes i situades en el mateix cromosoma, que s'anomena *haplotip*. Aquest arrossegament selectiu dóna lloc a l'anomenat *efecte autoestopista* (*hitchhiking effect*, Maynard-Smith i Haigh, 1974; revisat a Aguadé, 2009) que pot afectar tota la molècula de DNA en el cas que no recombini, com seria en el DNA mitocondrial. En cromosomes que presenten recombinació, la zona

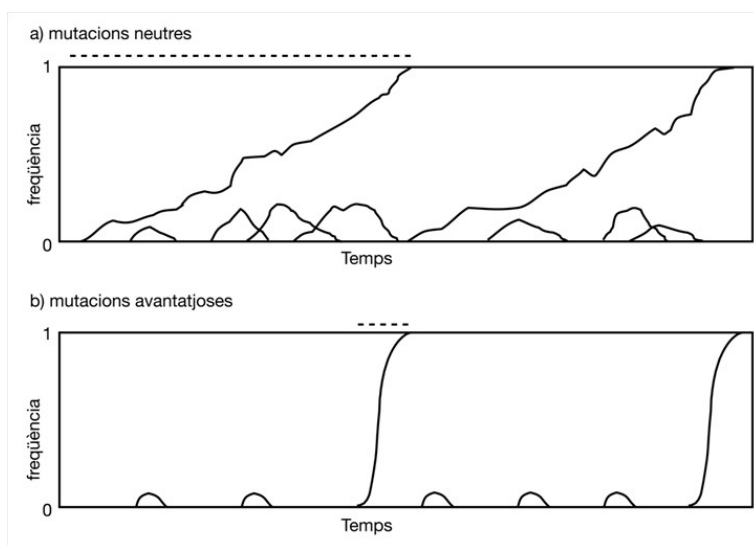


FIGURA 4 Representació esquemàtica de la trajectòria envers la fixació (freqüència = 1) i la pèrdua (freqüència = 0) a) de mutacions neutres i b) de mutacions avantatjoses. Les línies discontinües representen el temps **condicionat** de fixació, que és molt superior per a les mutacions neutres que per a les mutacions avantatjoses conduïdes a la fixació per la selecció direccional (vegeu la figura 3).

afectada per l'arrossegament selectiu varia en extensió segons quina sigui la taxa de recombinació en la regió i la força de la selecció.

L'arrossegament selectiu deixa una empremta en la variació de la zona afectada, empremta que varia en funció del temps transcorregut des que es va produir l'arrossegament selectiu fins que finalment es va perdent per acció de la mutació i la recombinació. L'empremta és, per tant, transitòria, i se'n detecten diferents versions que difereixen en la profunditat del temps en què són detectables, ja que després ho deixen de ser. Això ha conduït al desenvolupament d'una sèrie de tests estadístics que intenten detectar una o diverses versions de l'empremta que deixa un canvi adaptatiu recent.

En el cas de selecció direccional, l'empremta inicial que pot detectar-se en estudiar una sola població consisteix en una davallada del nivell de variació o polimorfisme en la regió afectada per l'arrossegament selectiu; és a dir, es tracta de reconèixer una davallada de variació en el DNA que és localitzada i que no es pot explicar exclusivament per l'efecte de l'atzar i de la història demogràfica de la població. En aquesta fase inicial també es pot detectar un excés de variants derivades (és a dir, de la variant genètica recentment produïda) a elevada freqüència i també per un increment en el nivell de desequilibri del lligament a prop de la diana de la selecció. En passar el temps, es va recuperant la variació en la zona afectada, fet que provoca un excés de variants derivades a baixa freqüència perquè són mutacions molt recents. Aquestes empremtes es van erosionant ràpidament en passar el temps. Altrament, hi ha una empremta que únicament es pot detectar si s'estudien diverses poblacions en les quals s'hagin donat canvis adaptatius diferencials en regions con-

cretes del genoma. Així, la diferenciació genètica entre poblacions (mesurada per exemple amb l'estadístic F_{st}) serà més gran, en comparació de la resta del genoma, en una regió que hagi sofert un canvi adaptatiu únicament en una de les poblacions.

En el cas de selecció equilibradora o de selecció direccional en marxa (vegeu la figura 3), és a dir, en el cas d'un arrossegament selectiu parcial (Aguadé, 2009), l'haplotip amb la mutació avantatjosa haurà assolit ràpidament una freqüència intermèdia. La variació en aquest haplotip presentarà les mateixes empremtes que s'han esmentat en el cas d'un arrossegament selectiu total, però es conservaran encara en la població altres haplotips sense l'empremta de la selecció. Altrament, aquest haplotip seleccionat, en ser un haplotip jove, presentarà, a diferència de la resta d'haplotips, una zona molt extensa d'identitat o quasi identitat.

Canvis adaptatius llunyans. En parlar del relloige molecular s'ha comentat que els canvis provocats per mutacions neutres s'acumulen en les seqüències de DNA amb una taxa constant. Aquesta taxa pot diferir, per raons funcionals, entre regions del genoma i també en un mateix gen segons sigui la limitació per variar del producte gènic, que depèn de la fracció de mutacions amb efecte deleteri que hagin estat eliminades per acció de la selecció purificadora. Enfront d'aquesta constància, la taxa amb què es donen els canvis adaptatius pot accelerar-se en determinats moments de la història d'una espècie o d'un llinatge, com per exemple en ocupar nous hàbitats, en iniciar-se el procés d'especiació, etc. Aquesta acceleració es reflectirà en un increment en la divergència entre espècies, de tal manera que en les regions del genoma que hagin sofert canvis adaptatius no es detectarà la proporcionalitat esperada en la variació neutra entre els nivells de polimorfisme (o

diversitat dins de l'espècie) i de divergència (entre espècies). Quan la diana de la selecció han estat les proteïnes, els canvis adaptatius llunyans poden donar lloc a un increment en la taxa de substitució en posicions no sinònimes (amb canvi d'aminoàcid) sense que es vegi afectada la taxa de substitució en posicions sinònimes (sense canvi d'aminoàcid). Aquest fet ha conduït al desenvolupament d'una sèrie de tests que intenten detectar aquestes empremtes que deixen els canvis adaptatius llunyans (Hudson, Kreitman i Aguadé, 1987; McDonald i Kreitman, 1991; Hughes i Nei, 1988), empremtes que, a diferència de les deixades pels canvis adaptatius recents, no són transitòries.

Nous habitats, nous reptes i adaptacions recents

L'ampliació de l'àrea de distribució d'una espècie comporta reptes importants per a la supervivència i reproducció. Quan aquesta ampliació ha estat recent, com és el cas de les espècies *Homo sapiens* i *Drosophila melanogaster* —que van ser de les primeres espècies eucariotes a tenir el seu genoma seqüenciat—, la variació en els descendents de les poblacions colonitzadores encara conté les empremtes dels canvis adaptatius que aquestes han sofert enfront dels diversos reptes que els han suposat, entre d'altres, els canvis d'hàbitat. És fàcil pensar que en aquesta expansió a través d'Euràsia els individus amb variants més resistents a les noves condicions climàtiques haurien tingut més probabilitat de sobreviure i reproduir-se, la qual cosa comportaria l'increment de freqüència d'aquestes variants a la població.

Aquestes situacions associades a nous reptes han propiciat estudis genèticopoblacionals per detectar l'acció de la selecció

natural positiva que s'han beneficiat, d'una banda, de la disponibilitat de la seqüència del genoma humà i del genoma de *D. melanogaster*, i de l'altra, dels diversos estudis a escala genòmica de la variació nucleotídica (SNP, abreviatura del terme anglès *single nucleotide polymorphism*).

Rastreig genòmic i canvis adaptatius recents en Drosophila i els humans. La possibilitat d'estudiar la variació a escala genòmica, és a dir, en múltiples regions del genoma o al llarg de tot el genoma, ens permet reconèixer l'acció de la selecció natural sense que calgui saber quina és la funció concreta del gen o de la regió del genoma analitzada. Aquesta aproximació permet una detecció «a cegues» de gens i regions genòmiques implicats en una determinada característica i que a més han tingut un paper important en el seu canvi adaptatiu recent. La validació d'aquesta aproximació genòmica l'ha proporcionada el fet que a més de permetre detectar l'empremta de l'acció recent de la selecció positiva en múltiples regions anònimes (és a dir, sense funció coneguda) l'hagi detectat també en gens candidats.

Aquesta aproximació, també denominada *agnòstica* (Hancock *et al.*, 2010), va sovint unida a l'estudi de poblacions candidates, és a dir, poblacions que han estat sotmeses recentment a reptes concrets durant la seva expansió a nous habitats. Així, en *D. melanogaster* s'han dut a terme estudis multilocus en poblacions europees (Glinka *et al.*, 2003; Ometto *et al.*, 2005; Orenge i Aguadé, 2004) que han permès identificar l'empremta de canvis adaptatius recents en diverses regions del cromosoma X. En aprofundir l'estudi d'algunes d'aquestes regions, tenint en compte la dificultat afegida de destriar els efectes de la selecció positiva dels del coll d'ampolla associat a la sortida d'Àfrica de l'espècie, s'ha confirmat que aquestes regions reflecteixen canvis adaptatius recents i a més s'ha po-

gut localitzar el gen diana de la selecció. Els gens identificats en dos dels casos estan implicats en l'adaptació al fred (Svetec *et al.*, 2009, 2011), mentre que en el tercer codifica un enzim de la via de síntesi d'ecdisteroides (Orengo i Aguadé, 2007).

Un estudi recent de la variació en l'exoma humà (Yi *et al.*, 2010) —és a dir, la fracció codificant del genoma— ha comparat la variació en una població candidata d'haver patit adaptacions a una altitud molt elevada (població del Tibet) i altres dues poblacions que viuen a altituds baixes (xinesa i danesa). La inclusió en l'estudi de dues poblacions amb diferent parentiu amb la població tibetana ha permès establir el model demogràfic d'aquestes poblacions, estimar la diferenciació genètica entre poblacions (F_{st}) i a més estimar els canvis de freqüència en la branca específica dels tibetans (vegeu la figura 5). Entre les empremtes més fortes detectades en aquest estudi es troben diversos gens implicats en el transport i la regulació de l'oxigen, també anomenats *gens de resposta a la hipòxia*. L'enriquiment en aquests tipus de gens concordaria amb l'adaptació a nivells baixos d'oxigen entre els tibetans. El senyal més fort, que continua sent significatiu quan d'una banda es té en compte la història demogràfica de la població tibetana i de l'altra es corregeix amb múltiples tests, correspon al gen *EPAS1*, també anomenat *Hif-2a*, que codifica el factor de transcripció HIF-2a o factor 2a induïble per hipòxia. L'al·lel present en els tibetans a elevada freqüència, o al·lel «tibetà», confereix molt probablement adaptació a l'ambient hipòxic en altituds elevades.

Comparació de genomes i canvis adaptatius llunyans. La disponibilitat de genomes complets de múltiples espècies amb un grau variable de parentiu n'ha possibilitat la comparació per tal de detectar regions codificadores que han estat sotmeses a can-

vis adaptatius en el seu passat llunyà. La comparació d'una única còpia del genoma de cada espècie permet obtenir estimacions de la divergència genètica entre espècies i també en les diferents branques de la filogenia, incloent-hi branques internes. Els models emprats consideren que la variació en posicions sinònimes és neutra, per tal com les mutacions no sinònimes poden ser avantatjoses i deletèries a més de neutres. L'estadístic generalment utilitzat en el contrast d'hipòtesis és el quocient entre la divergència no sinònima i la divergència sinònima, és a dir omega: $\omega = K_a / K_s$. En absència de selecció, s'espera que ω sigui igual a 1, mentre que s'espera que sigui inferior a 1 quan la selecció purificadora ha conduït a la pèrdua d'una fracció de les mutacions no sinònimes, i superior a 1 úni-

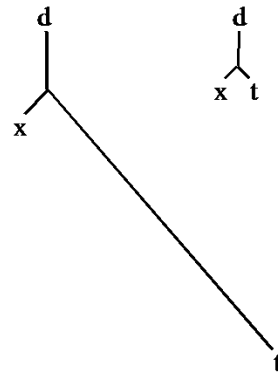


FIGURA 5 Representació gràfica de la diferenciació entre les poblacions estudiades (d, danesa; x, xinesa; t, tibetana) en el gen *EPAS1* (esquerra) i a la resta del genoma (dreta). Les branques que condueixen als xinesos i tibetans són de la mateixa longitud a escala genòmica, mentre que per al gen *EPAS1* la longitud més gran de la branca que condueix als tibetans reflecteix la diferenciació produïda en aquest gen en aquest llinatge.

cament quan s'han donat canvis adaptatius, és a dir, quan algunes mutacions no sinònimes han estat conduïdes a la fixació per la selecció positiva. Com l'acció de la selecció purificadora és força ubiqua pel que fa a ~~la proteïna~~, és molt difícil que quan es considera tot un gen el valor de ω sigui superior a 1. El desenvolupament de models que consideren que ω pot variar entre codons i també entre branques de la filogènia ha incrementat molt la potència dels tests emprats per detectar l'acció de la selecció positiva. S'ha d'esmentar que quan es detecta un increment de K_a , aquest pot reflectir tant l'acció de la selecció positiva en llocs no sinònims com la relaxació de la selecció purificadora.

Des que se seqüenciaren els genomes humà i del ximpanzé hi ha hagut molt d'interès a identificar gens que han evolucionat de manera adaptativa en cadascun d'aquests llinatges en un esforç per entendre què és el que ens va fer humans. Entre els diversos estudis amb aquest objectiu, parlarem de l'estudi basat en la comparació del genoma humà amb els del ximpanzé, gos, rata i ratolí (Arbiza *et al.*, 2006). Aquest estudi ha utilitzat tests especialment dissenyats per diferenciar l'acció de la selecció positiva de la relaxació de la selecció purificadora. D'un total de prop de 10.000 gens ortòlegs estudiats, es pot dir que uns 100 han evolucionat de manera adaptativa en el llinatge humà i uns 600 en el llinatge del ximpanzé. Aquests valors són molt conservatius, atès que l'esforç per disminuir el nombre de falsos positius condueix inevitablement a incrementar el nombre de veritables positius que són estadísticament rebutjats en el contrast d'hipòtesis. Entre les classes funcionals a les quals pertanyen aquests gens es poden destacar les de percepció sensorial, resposta immunitària, metabolisme proteic cel·lular, etc. No sembla que les categories funci-

onals enriquides en gens que han sofert canvis adaptatius siguin diferents entre el llinatge humà i el del ximpanzé. El que sí que són diferents són els gens específics de cada llinatge que han sofert canvis adaptatius, per la qual cosa aquests gens poden de ben cert estar involucrats en les diferències biològiques que han conduït, respectivament, als humans i als ximpanzés. L'estudi d'aquests gens ens aproparà a respondre la pregunta que molta gent es fa sobre què ens fa humans, i també a la pregunta complementària sobre què els fa ximpanzés.

Molta de la informació generada pels estudis de selecció ens porta a reconèixer el paper rellevant de determinades variants en gens concrets que, finalment, tindran unes implicacions funcionals que hem d'entendre dins d'un context ampli de l'organisme. Anar, però, del gen concret i el seu producte a les implicacions funcionals, no és un pas senzill, i cal la integració de la nova funció dins del context del conjunt del genoma i els seus productes.

LA INTEGRACIÓ DE LA INFORMACIÓ DE GENOMES SENCERS

Tractem ara el desplegament de la informació dels genomes, en l'espai i el temps, a través dels productes de la seva expressió. La comprensió del fenotip requereix la integració dels diferents nivells de complexitat de l'expressió genètica: la transcripció, la traducció, la catàlisi enzimàtica i la interacció mútua de tots aquests components moleculars en els circuits de regulació (RNA petits, interacció entre proteïnes o entre proteïnes i seqüències reguladores del genoma, la regulació dels fluxos metabòlics, etc.). Des d'una perspectiva conceptual, convé remarcar quelcom que pot sem-

blar a primera vista trivial però que té conseqüències notables en la comprensió, manipulació i fabricació eventual dels sistemes biològics. Ens referim al concepte de recursivitat o circularitat metabòlica que, de manera sintètica, significa que el metaboloma és un producte del proteoma i el proteoma és un producte del metaboloma. En la formulació del biòleg teòric Robert Rosen, els sistemes biològics són tancats respecte a les seves causes eficients, cosa que els fa radicalment i fonamentalment diferents de les màquines de fabricació humana. Cap màquina artificial no es repara i regenera des de dins, com ho fa la cèl·lula (Hofmeyr, 2007). Malgrat aquesta constatació, molts llibres de text i molts articles científics insisteixen en la comparació de màquines (o computadors o fàbriques) amb cèl·lules i un dels corrents actuals de la biologia sintètica, amb les arrels i la inspiració en el camp de l'enginyeria, insisteix en aquesta visió purament cartesiana, menystenint les propietats emergents i la certa imprevisibilitat que caracteritzen la matèria vivent. En aquest context contrasta l'arrogància amb la qual ens promet el control, el domini i la síntesi d'una cèl·lula l'escola d'enginyers del MIT, i el realisme amb el qual es reconeix la manca de comprensió de com funciona una de les cèl·lules més senzilles, un micoplasma, malgrat el detallisme de la descripció que n'ha fet el grup de Luis Serrano, al Centre de Regulació Genòmica.

Estem lluny d'entendre l'allau de dades que es deriven de les tecnologies òmiques, tot i els sofisticats mètodes computacionals. Molts autors, centres de recerca i administradors de la política científica empenen el terme *biologia de sistemes* per referir-se als esforços combinats d'experimentalistes i teòrics d'arribar a disposar de models biològics computacionals amb capacitat de predicció. Però no tothom dóna el mateix

significat a aquest terme i, com ha passat sovint al llarg de la història de la ciència, hi ha també molt d'oportunisme a l'hora de revestir el treball de sempre amb un nom de moda per aconseguir fons i projectes. Com ha suggerit l'epistemòleg Michel Morange, la biologia de sistemes contemporània no ha vingut a contraposar-se al reduccionisme de la biologia molecular o a superar-lo (Morange, 2011). Aquest és un mite filosòfic i històric. El reduccionisme metodològic, l'estudi exhaustiu dels gens o les proteïnes individuals, ha donat i segueix donant uns fruits intel·lectuals fabulosos. Cap científic mínimament conscient admetria que conèixer fins al més mínim detall cada component del sistema ens és suficient per comprendre el comportament global, del conjunt de tots aquests components.

Els organismes vivents són sistemes complexos, tant des del punt de vista estructural com del dinàmic. Per arribar a comprendre'ls Morange (2011) reivindica la necessitat d'articular tots els esquemes explicatius que s'han emprat en les ciències biològiques, a saber: els principis d'intel·ligibilitat evolutiva, mecanicista i purament física. És a dir, la confluència del pensament filogenètic, de l'escrutini molecular i de la perspectiva de les xarxes de molts components interaccionant. Les tradicions de la teoria evolutiva, de la bioquímica i la genètica molecular i dels estudis sistèmics (en ecologia o en metabolisme) han de confluir en la biologia de sistemes. A mesura que augmenta la grandària del sistema, el nombre de components i les interaccions entre ells, s'incrementa la necessitat de la modelització computacional per tal d'explicar el funcionament correcte o patològic del sistema. Morange (2011) també assenyalava que reconèixer la necessitat d'una articulació explicativa de les diverses tradicions biològiques no ens for-

neix amb la recepta per implementar-la. Un desconeixement mutu dels camps d'estudi restarà credibilitat a les aproximacions. No ajuda gens, per exemple, que el primer article publicat per un grup de físics sobre xarxes metabòliques «descobris» que el principal node de connectivitat (*hub*) del metabolisme cel·lular és l'aigua (Jeong *et al.*, 2000). Pot semblar que tot aquest esforç de reconstruir la xarxa metabòlica a partir de la informació d'un genoma complet no ens ensenya res de nou. La lliçó per aprendre és que el treball pluridisciplinari requereix un coneixement dels diversos esquemes explicatius per generar sinergies originals i valuoses, alhora que cal acceptar que potser els fenòmens biològics, derivats d'una llarga història evolutiva, despleguen una complexitat que no encaixa en la cerca de lleis universals, característica de les ciències físiques (Keller, 2007).

Una perspectiva nova sobre conceptes vells

Durant els estudis de les xarxes d'interacció física entre proteïnes (l'interactoma) i les xarxes metabòliques, s'ha apreciat l'extensió de comportaments poc coneguts o menystinguts fins ara, alhora que s'ha estimulat la investigació de propietats sistèmiques. Per un cantó s'ha vist que la coexistència de diverses funcions en una mateixa proteïna és més la regla que no pas l'excepció. Tradicionalment ens referíem a les cristallines com un dels exemples de bricolatge evolutiu a escala molecular (Jacob, 1977). Certs enzims metabòlics han estat reclutats per fer de moduladors de l'índex de refracció de les lents dels ulls. Ara tot sovint descobrim que molts enzims també són factors de transcripció que estableixen tot un complex entramat de circuits de regulació metabòlica i transcripcional.

Aquest caràcter polifacètic dels enzims (en anglès, *moonlighting*) ajuda a resoldre, en part, la paradoxa del nombre aparentment petit de gens en organismes complexos. Però la pluriocupació també ha estat invocada com una de les característiques que serien crucials en els primers moments de l'evolució metabòlica. L'ambigüitat enzimàtica, l'acció d'un enzim sobre més d'un substrat amb generació de productes diferents, seria necessària per fer possible una diversitat de transformacions químiques en una època en la qual la longitud dels genomes i el nombre de gens serien més aviat petits (Peretó, 2011). Lluny de la idea estereotipada dels enzims com a estructures rígides i extraordinàriament específiques en el reconeixement del substrat, ens convé pensar en termes d'estructures flexibles i amb capacitat de fer reaccionar molècules amb configuracions diferents (James i Twafik, 2003). Aquesta fluïdesa estructural és una dreuera de l'evolució metabòlica. A través de duplicacions gèniques, divergència i assumptió de noves funcions, la selecció natural ha explorat l'espai de la química orgànica i ha innovat el metabolisme dels diferents llinatges amb l'adquisició de noves rutes divergents a partir d'una xarxa central, antiga, quasi universal (Peretó, 2011).

Pel que fa a les propietats sistèmiques, l'estudi de les xarxes ha permès endinsar-se en aspectes com la robustesa o la fragilitat de la seva arquitectura (Solé, 2009). De la mateixa manera que els enginyers tenen un gran interès a caracteritzar la robustesa de la distribució elèctrica o de la xarxa d'Internet (el trencament d'un sol cable pot deixar sense electricitat una gran ciutat o sense Internet tot un país), els biòlegs cerquen els punts centrals, de gran interconnexió, en les xarxes funcionals, o investiguen les redundàncies i camins alternatius afavoridors de la tolerància envers els

danyos o les mutacions. La noció de robustesa —entesa com la capacitat de mantenir l'estabilitat fenotípica davant les pertorbacions i la incertesa del medi— té una llarga tradició en les ciències biològiques (Wagner, 2005), amb conceptes emparentats com el de dominància en genètica (Mendel, Fisher), la canalització ontogènica (Waddington) o l'homeòstasi fisiològica (Bernard). Els sistemes biològics i els tecnològics s'assemblen en la manera com combinen la robustesa amb les situacions de fragilitat. L'interès a conèixer i caracteritzar aquests extrems (*robust yet fragile*) de les xarxes biològiques molt evolucionades es troba en la convicció que és un camí eficaç per entendre, prevenir o combatre els estats patològics (Csete i Doyle, 2004).

Un altre concepte que igualment ha fet fortuna en l'estudi de les xarxes és el de modularitat, també amb una llarga tradició en biologia (Solé, 2009). Totes les xarxes biològiques complexes mostren una estructura jerarquitzada i estan compostes per mòduls. S'entén per mòdul una subxarxa en la qual les interconnexions internes són més abundoses que les connexions amb la resta de la xarxa. La modularitat ofereix avantatges de disseny arquitectònic, com ara la inèrcia en la propagació general d'una pertorbació produïda dins d'un mòdul, cosa que confereix robustesa al conjunt (Wagner, 2005). El que és més interessant és que els models d'evolució, emprant regles simples com la duplicació de nodes i la connexió amb altres elements de la xarxa evitant les redundàncies, generen xarxes modulars (Solé, 2009). És a dir, la modularitat i les seves conseqüències funcionals podrien ser un resultat inevitable del mode evolutiu de les xarxes biològiques.

La simbiosi com a model per a la biologia de sistemes

La naturalesa és promíscua. Només cal reparar en l'extensió fabulosa de la simbiosi, és a dir, la interacció estable entre individus de diferents espècies amb el resultat de l'aparició d'innovacions estructurals, metabòliques i de comportament, absents en les espècies abans de l'associació (Moya *et al.*, 2008; Moya i Peretó, 2011). Un dels exemples de simbiosi més ben conegut és el dels bacteris endosimbionts d'insectes. Aquests bacteris s'instal·len en cèl·lules i òrgans especialitzats, poden heretar-se verticalment i tenen unes funcions metabòliques relacionades amb la dieta de l'hoste. El fet que la gran majoria d'aquests bacteris no siguin cultivables fora de la cèl·lula que els hostatja ha limitat seriosament l'abast del seu estudi. Amb les tecnologies òmiques s'poden analitzar amb detall molts aspectes d'aquestes simbiosis (Moya *et al.*, 2008). Per exemple, podem seqüenciar el genoma complet dels bacteris endosimbionts, inferir la xarxa metabòlica que s'hi codifica i postular les interconnexions amb el metabolisme de l'hoste, fer estudis de transcripció, tant del simbiote com de l'hoste, conèixer la regulació mútua dels dos subsistemes, modelitzar la coevolució amb l'hoste, etc. Durant la convivència de l'insecte i l'endosimbiote pot introduir-se un segon convidat bacterià que complementaria o, fins i tot, substituirà el primer endosimbiote. En el primer cas, s'estableix un consorci bacterià de manera que el fenotip final és el resultat de les interaccions entre els genomes de l'hoste (nuclear i mitocondrial) i els genomes bacterians instal·lats més recentment (Moya i Peretó, 2011).

La dinàmica evolutiva d'aquests endosimbionts presenta també peculiaritats remarcables. En general, s'observa que l'evolució d'aquests bacteris intracel·lulars

comporta una reducció del nombre de gens que conté el seu genoma. És evident que aquelles funcions biològiques irrelevantes o supèrflues (pel canvi de mode de vida lliure al de vida intracel·lular o per la redundància amb funcions exercides amb més eficàcia per l'hoste) són ràpidament eliminades, per la via de la inserció d'elements transposables o per la pseudogenització. Els fenòmens de minimització de genomes i de xarxes metabòliques es poden simular *in silico* amb l'ajut de models d'evolució molecular. Així, és possible estudiar el paper de la contingència històrica, de l'atzar, de la redundància funcional o dels canvis en la robustesa de la xarxa metabò-

lica durant l'evolució dels endosimbionts (vegeu la figura 6). Observant l'evolució metabòlica en els endosimbionts som testimonis d'un fenomen contrari al que degué ocórrer en les primeres etapes de la història de la vida. La reducció del nombre d'elements en la xarxa afavoreix que alguns enzims despertin la seva capacitat d'actuar sobre més d'un substrat, i en relaxin així l'especificitat. L'evolució reductiva dels genomes durant l'endosimbiosi és també un bon model per a l'origen dels orgànuls eucariòtics, mitocondri i cloroplast, i serveix d'inspiració en els estudis sobre els components mínims d'una cèl·lula en el context

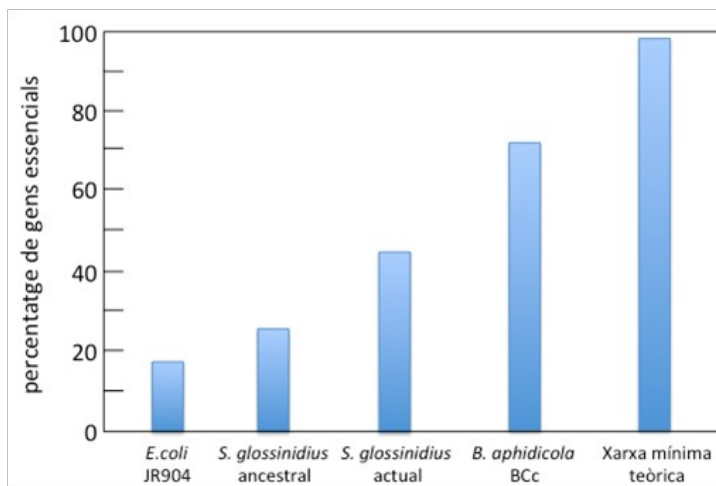


FIGURA 6 Anàlisi de la robustesa de diverses xarxes metabòliques reconstruïdes. Les xarxes sotmeses a anàlisi de balanç de flux (FBA, Orth *et al.*, 2010) foren reconstruïdes a partir del genoma complet corresponent: *Escherichia coli* K12 (Reed *et al.*, 2003), *Sodalis glossinidius* (xarxes ancestral i actual), endosimbiont facultatiu de la mosca tse-tse, i *Buchnera aphidicola* BCc, endosimbiont obligat del púg del cedre *Cinara cedri* (Belda *et al.*, en revisió). Aquestes xarxes constaven de 931, 741, 560 i 89 reaccions, respectivament. En la xarxa ancestral de *S. glossinidius* es va considerar que tots els pseudogens (148) i gens deletats (41) del genoma actual eren actius. La xarxa metabòlica mínima teòrica tenia 50 reaccions, es va deduir a partir del genoma mínim postulat per Gil *et al.* (2004) i l'anàlisi de robustesa es descrigué a Gabaldón *et al.* (2007). Es representa la fracció de gens essencials definits durant les simulacions de genoinul·lació. Un gen essencial és aquell que en ser eliminat provoca una reducció de la producció de biomassa, determinada per FBA, de més del 99 %.

de la biologia sintètica (Moya i Peretó, 2011).

Un graó addicional de complexitat en simbiosi l'assolim quan estudiem les comunitats bacterianes que interaccionen entre si i amb un organisme pluricel·lular sobre el qual estan instal·lades, ja sigui en la superfície externa o en les cavitats internes. És el cas dels microbiomes intestinals dels animals. Hi ha molta informació acumulada sobre la importància d'aquestes comunitats en el funcionament normal o patològic dels animals, del seu paper durant la digestió i el processament dels materials de la dieta. Tanmateix, l'adveniment de les tecnologies òmiques ha permès de fer els estudis més sistèmics. Avui, per exemple, podem catalogar, caracteritzar o descobrir noves activitats enzimàtiques implicades en el processament del material lignocel·lulòsic que ingereixen els herbívors. L'estudi del microbioma dels remugants o dels tèrmits pot adreçar-nos al desenvolupament d'estratègies noves en biocombustibles. La investigació del microbioma intestinal humà ha suggerit la connexió entre la seva composició i l'obesitat o d'altres estats patològics.

Però descriure no és comprendre. Una vegada més, ens calen eines conceptuais i teòriques més avançades per poder integrar, fer intel·ligibles i interpretables totes les dades que es deriven dels estudis genòmics i d'altres incursions en la globalitat dels sistemes. El gran repte de la biologia del segle XXI és assolir la maduresa d'una ciència que pugui gestionar un volum immens d'informació i treure'n un millor coneixement i una millor comprensió de la complexitat biològica en totes les seves dimensions: l'amplitud de la biodiversitat i la profunditat de la història evolutiva.

EL CAMÍ QUE ENS ESPERA

En els darrers anys s'ha produït una autèntica revolució en la biologia gràcies a l'estudi dels genomes, especialment per les possibilitats d'obtenir seqüències de DNA i per les potencialitats informàtiques d'analitzar-les. En els propers anys això encara augmentarà molt més, i l'estudi dels genomes i l'evolució tindrà, sens dubte, un impacte creixent en la biologia. Una biologia que haurà de tenir cura de quantitats immenses de dades i de complexes anàlisis. Una biologia que haurà de fer de la complexitat no un repte, sinó una eina. La vida està mostrant una complexitat molt gran i la seva comprensió no pot evitar haver de tractar grans quantitats d'informació. L'estudi dels genomes, inaugurant el tracte elegant de gran quantitat d'informació, està ja actualment obrint pas a aquesta nova biologia, amb preguntes velles i noves, però amb unes eines molt més poderoses que mai no ha tingut la biologia.

BIBLIOGRAFIA

- AGUADÉ, M. (2009). «Polimorfisme nucleotídic i canvis adaptatius». *Treb. Soc. Cat. Biol.*, 60: 93-102.
- ARBIZA, L.; DOPAZO, J.; DOPAZO, H. (2006). «Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome». *PLoS Comput. Biol.*, 2: e38.
- CSETE, M.; DOYLE, J. C. (2004). «Bow ties, metabolism and disease». *Trends Biotechnol.*, 22: 446-450.
- GABALDÓN, T.; PERETÓ, J.; MONTERO, F.; GIL, R.; LATORDA, A.; MOYA, A. (2007). «Structural analyses of a hypothetical minimal metabolism». *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 362: 1751-1762.
- GIL, R.; SILVA, F. J.; PERETÓ, J.; MOYA, A. (2004). «Determination of the core of a minimal bacterial gene set». *Microbiol. Mol. Biol. Rev.*, 68: 518-537.
- GLINKA, S.; OMETTO, L.; MOUSSET, S.; STEPHAN, W.; DE LORENZO, D. (2003). «Demography and natural selection have shaped genetic variation in *Drosophila melanogaster*: a multi-locus approach». *Genetics*, 165: 1269-1278.

- HANCOCK, A. M.; ALKORTA-ARANBURU, G.; WITONSKY, D. B.; RIENZO, A. DI (2010). «Adaptations to new environments in humans: the role of subtle allele frequency shifts». *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, 365: 2459-2468.
- HOFMEYER, J.-H. S. (2007). «The biochemical factory that autonomously fabricates itself: A systems biological view of the living cell». A: BOOGERD, F. C.; BRUGGEMAN, F. J.; HOFMEYER, J.-H. S.; WESTERHOFF, H.-V. (2007). *Systems biology. Philosophical foundations*. Amsterdam: Elsevier, 217-242.
- HUDSON, R. R.; KREITMAN, M.; AGUADÉ, M. (1987). «A test of neutral molecular evolution based on nucleotide data». *Genetics*, 116: 153-159.
- HUGHES, A. L.; NEI, M. (1988). «Maintenance of MHC polymorphism». *Nature*, 335: 402-403.
- JACOB, F. (1977). «Evolution and tinkering». *Science*, 196: 1161-1166.
- JAMES, L. C.; TWAFIK, D. S. (2003). «Conformational diversity and protein evolution –a 60-year-old hypothesis revisited». *Trends Biochem. Sci.*, 28: 361-368.
- JEONG, H.; TOMBOR, B.; ALBERT, R.; OLTVAI, Z. N.; BARABÁSI, A.-L. (2000). «The large-scale organization of metabolic networks». *Nature*, 407: 651-654.
- KELLER, E. F. (2007). «A clash of two cultures». *Nature*, 445: 603.
- LÓPEZ-GARCÍA, P.; MOREIRA D. (2008). «Tracking microbial biodiversity through molecular and genomic ecology». *Research in Microbiology*, 159: 67-73.
- MAYNARD-SMITH, J.; HAIGH, J. (1974). «The hitch-hiking effect of a favorable gene». *Genet. Res.*, 23: 23-35.
- MCDONALD, J. H.; KREITMAN, M. (1991). «Adaptive protein evolution at the Adh locus in *Drosophila*». *Nature*, 354: 114-116.
- MORANGE, M. (2011). *Els secrets del que és viu: contra el pensament únic de la biologia* (col·l. «Pensament científic», 1). Barcelona: IEC; València: PUV.
- MOYA, A.; PERETÓ, J. (2011). *Simbiosis. Seres que evolucionan juntos*. Madrid: Síntesis.
- MOYA, A.; PERETÓ, J.; GIL, R.; LATORRE, A. (2008). «Learning how to live together. Genomic insights into prokaryote-animal symbioses». *Nat. Rev. Genet.*, 9: 218-229.
- NIELSEN, R. (2005). «Molecular signatures of natural selection». *Annu. Rev. Genet.*, 39: 197-218.
- OMETTO, L.; GLINKA, S.; LORENZO, D. DE; STEPHAN, W. (2005). «Inferring the effects of demography and selection on *Drosophila melanogaster* populations from a chromosome-wide scan of DNA variation». *Mol. Biol. Evol.*, 22: 2119-2130.
- ORENGO, D. J.; AGUADÉ, M. (2004). «Detecting the footprint of positive selection in a european population of *Drosophila melanogaster*: multilocus pattern of variation and distance to coding regions». *Genetics*, 167: 1759-1766.
- (2007). «Genome scans of variation and adaptive change: extended analysis of a candidate locus close to the phantom gene region in *Drosophila melanogaster*». *Mol. Biol. Evol.*, 24: 1122-1129.
- ORTH, J. D.; THIELE, I.; PALSSON, B. O. (2010). «What is flux balance analysis?». *Nat. Biotechnol.*, 28: 245-248.
- PERETÓ, J. (2011). «Origin and evolution of metabolisms». A: GARGAUD, M.; LÓPEZ-GARCÍA, P.; MARTIN, H. [ed.]. *Origins and evolution of life. An astrobiological perspective*. Cambridge: Cambridge University Press, 270-288.
- REED, J. L.; VO, T. D.; SCHILLING, C. H.; PALSSON, B. O. (2003). «An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR)». *Genome Biol.*, 4: R54.
- SOLÉ, R. (2009). *Xarxes complexes. Del genoma a Internet*. Barcelona: Empúries.
- SVETEC, N.; PAVLIDIS, P.; STEPHAN, W. (2009). «Recent strong positive selection on *Drosophila melanogaster* HDAC6, a gene encoding a stress surveillance factor, as revealed by population genomic analysis». *Mol. Biol. Evol.*, 26: 1549-1556.
- SVETEC, N.; WERZNER, A.; WILCHES, R.; PAVLIDIS, P.; ALVAREZ-CASTRO, J. M.; BROMAN, K. W.; METZLER, D.; STEPHAN, W. (2011). «Identification of X-linked quantitative trait loci affecting cold tolerance in *Drosophila melanogaster* and fine mapping by selective sweep analysis». *Mol. Ecol.*, 20: 530-544.
- WAGNER, A. (2005). *Robustness and evolvability in living systems*. Princeton: Princeton University Press.
- YI, X. [et al.] (2010). «Sequencing of 50 human exomes reveals adaptation to high altitude». *Science*, 329: 75-78.

SOBRE ELS AUTORS

Montserrat Aguadé i Porres (Barcelona, 1949). Catedràtica de genètica de la Universitat de Barcelona, membre de l'Institut de Recerca de la Biodiversitat de la Universitat de Barcelona, i investigadora principal del grup de recerca de genètica molecular evolutiva.  seu interès investigador se

centra en l'anàlisi de la variabilitat en les seqüències de DNA per establir el paper que ha tingut la selecció natural durant l'evolució. Ha publicat més de 80 articles científics amb repercussió internacional. Sota la seva direcció s'han format més d'una desena de doctors i diversos investigadors postdoctorals. Ha estat presidenta de la Society of Molecular Biology and Evolution (2005-2007), tresorera de la Societat Catalana de Biologia (1999-2005) i és actualment presidenta de la Sociedad Española de Genética. Ha estat, a més, editora de diverses revistes científiques internacionals com *Genetics* (des de 2000) i *Heredity* (1999-2004), i ha format part de diverses comissions nacionals i internacionals d'avaluació científica. Ha estat guardonada per la Generalitat de Catalunya amb la Distinció per a la Promoció de la Recerca Universitària (2001-2007), per la Universitat de Barcelona amb un ajut del Programa d'intensificació de l'activitat investigadora (2007-2010) i ha rebut el Premio Nacional de Genética 2010 en la modalitat de recerca bàsica.

Jaume Bertranpetit i Busquets (Camprodon, 1952). Catedràtic de biologia a la Universitat Pompeu Fabra (Barcelona). Fundador de la Unitat de Biologia Evolutiva en el Departament de Ciències Experimentals i de la Salut d'aquesta universitat. Antic professor de la Universitat de Barcelona i deixeble de Luca Cavalli-Sforza a Stanford, Califòrnia. Els seus camps de recerca són entorn de l'estudi i la comprensió de la diversitat del genoma humà: genètica de les poblacions humanes, evolució molecular, genòmica comparada i la interacció entre biologia evolutiva humana i altres camps, incloent-hi la genètica dels caràcters i malalties complexos, genètica estadística i història de les poblacions huma-

nes. Les seves publicacions més recents tracten del desequilibri de lligament i selecció natural en el genoma humà, tot buscant les bases genètiques de les adaptacions humanes o de poblacions concretes. Actualment treballa en la comprensió evolutiva de xarxes funcionals en el que es pot anomenar *biologia evolutiva de sistemes*. Té nombroses publicacions. Membre numerari de la Secció de Ciències Biològiques de l'IEC, directiu de diverses organitzacions científiques (ICHN, SCB, ICA), ha estat degà de la Facultat de Ciències de la Salut i la Vida de la UPF (1998-99), vicerector de política científica de la UPF (1999-2001), director del Centro Nacional de Genotipado (1994-2011) i director d'ICREA (Institució Catalana per a la Recerca i Estudis Avançats) des de 2007.

Juli Peretó i Magraner (Alzira, 1958). Professor de bioquímica i biologia molecular de la Universitat de València, membre del Grup de Genètica Evolutiva de l'Institut Cavanilles de Biodiversitat i Biologia Evolutiva i membre numerari de la Secció de Ciències Biològiques de l'IEC. El seu interès investigador se centra en l'origen i evolució de les xarxes metabòliques, la noció de genoma mínim i la història de les idees sobre l'origen natural i la síntesi artificial de vida. Ha introduït en la docència de la Universitat de València l'estudi de l'origen i evolució primitiva de la vida. Ha esmerçat molts esforços en la difusió del pensament evolutiu a través de conferències, jornades, cursos, publicacions, exposicions i documentals, adreçats especialment a estudiants i professorat preuniversitari. És membre de la SCB, ha estat vicerector de la Universitat de València (1994-2002) i secretari (2005-2008) i vicepresident segon (2011-2014) de la International Society for the Study of the Origin of Life.