

Lliçó inaugural

Genomes i proteomes

JOSEP M. SERRAT I JURADO

Una lliçó com la que avui tinc el gust de llegir amb motiu de l'acte inaugural del curs 2000-2001 ha de tenir forçosament unes característiques una mica diferents de les lliçons que impartim al llarg del curs a les aules de la nostra universitat. Els assistents a aquest acte proveniu de camps del coneixement ben diversos i no voldria pas cansar l'audiència menys familiaritzada amb els conceptes de les ciències naturals amb un tema d'especialització bioquímica. La lliçó d'avui vol tenir un caràcter general, encara que això comporti haver de deixar alguns temes oberts i que d'altres no es puguin aprofundir com es faria en una classe convencional. A l'hora de preparar aquesta lliçó, la intenció per part meua ha estat d'explicar d'una manera suggerent i prou àmplia el món de la informació genètica, el protagonista del qual és el DNA, l'àcid desoxiribonucleic.

Moltes de les propietats que fan del DNA una molècula tan especial es poden explicar sense que calgui fer prèviament una descripció a fons de la seva estructura i del seu comportament químic; de fet, algunes de les característiques funcionals del DNA s'han descobert sense la intervenció de les eines experimentals de la bioquímica¹. Sabem que el DNA es pot representar mitjançant una seqüència lineal de quatre lletres, la C, la G, la T i la A, que en el vocabulari especialitzat s'anomenen abreviadament bases o nucleòtids. Cadascuna d'aquestes lletres simbolitza un compost químic diferent. El DNA està format, doncs, per una cadena de molècules; en termes químics, es tracta d'un polímer. La informació genètica d'un determinat individu és el «text» que es llegiria en el seu DNA complet, el seu genoma. En el cas dels éssers humans aquest text té aproximadament tres mil milions de caràcters. El famós Projecte Genoma Humà, del qual tornarem a parlar més endavant, té per objectiu saber què diu el DNA dels humans; des de la primera fins a la tres mil milionena lletra.

Però, què hem d'entendre per «informació genètica»? De què informa, el genoma? Per poder-ho explicar, deixeu-me parlar abans d'un altre polímer biològic d'una importància cabdal per al funcionament dels éssers vius: les proteïnes.

Tots estem familiaritzats amb el terme «proteïna» i sovint l'associem amb els aliments que consumim cada dia. No obstant això, la funció d'aquestes molècules en l'organisme ultrapassa de

1. Robert Holliday ho expressava de la manera següent: «els geneticistes [dels anys 60] tendien a pensar en termes conceptuals més que no pas en termes bioquímics o moleculars. El DNA, per a ells, era una estructura física que podia mutar i recombinar-se i no calia preocupar-se en aquell estadi de la bioquímica subjacent» (Holliday, 1990).

molt la mera funció nutritiva. De fet, la major part d'éssers vius no en consumeixen mai. Per copsar la importància de les proteïnes potser ens ajudarà presentar-ne aquí una petita mostra: les estructures macroscòpiques dels animals, com ara l'esquelet, la musculatura, les plomes de les aus, la pell, el pèl i les ungles, són proteïnes; els enzims, que catalitzen totes les reaccions bioquímiques, també són proteïnes; els anticossos, l'hemoglobina de la sang, la insulina, la clorofil·la de les plantes, la càpside dels virus, els cilis i flagels dels microorganismes, els antígens responsables del rebuig d'òrgans en els transplants, les molècules que regulen el transport de substàncies a través de la membrana cel·lular i els agents patològics que causen la malaltia de les vaques bojes, els prions, són també proteïnes. Al llarg de la vida d'un ésser humà hi actuen aproximadament cent mil proteïnes diferents². No seria agosarat de dir que les proteïnes ens fan allò que som.

Quina és, doncs, la relació entre el DNA i les proteïnes? Les proteïnes són també un polímer lineal, en aquest cas, format per vint tipus diferents de molècules, els anomenats aminoàcids, que també es poden representar mitjançant una lletra³. Doncs bé, la seqüència d'aminoàcids d'una proteïna es pot traduir mitjançant un codi, el «codi genètic», en una seqüència de l'alfabet de quatre lletres del DNA. La seqüència que en resulta forma part de la informació genètica de l'individu i es troba inserida en algun lloc o altre del seu genoma. Aquest segment s'anomena el «gen» de la proteïna⁴. El que s'esdevé en la realitat, però, és el procés invers. La cèl·lula disposa d'un mecanisme químic per poder traduir la informació del DNA i sintetitzar totes les proteïnes que necessita per al seu metabolisme.

Per tant, quan parlem d'informació genètica volem dir informació per poder sintetitzar proteïnes. El DNA porta la informació i les proteïnes realitzen les funcions necessàries per al manteniment en vida d'un ésser viu. Ara bé, com de seguida veurem, el missatge genòmic conté encara altres tipus d'informació.

Els éssers vius estem formats per cèl·lules. Cada cèl·lula disposa d'una quantitat fixa de material genètic, la mateixa quantitat i la mateixa seqüència de bases en totes les cèl·lules d'un mateix individu, tant si es tracta de les cèl·lules que formen el teixit adipós com si es tracta de les neurones de l'escorça cerebral. El que diferencia els teixits a nivell molecular, és el conjunt de gens que poden estar actius en un determinat teixit. Hi ha gens del genoma que són actius en tots els teixits i

2. Aquesta quantitat és l'estimació que es feia abans de la seqüenciació completa dels cromosomes humans 21 i 22. L'anàlisi del DNA d'aquests cromosomes indica que el nombre de gens d'un ésser humà podria ser menor del que s'havia estimat (Dunham *et al.*, 1999; Hattori *et al.*, 2000). No obstant, encara no es pot dir res. Els cromosomes 21 i 22 són els més petits dels humans, sense comptar els cromosomes sexuals. Per poder conèixer amb més exactitud aquesta dada haurem d'esperar la seqüenciació de la resta de cromosomes. Cal tenir en compte, a més, que la relació entre el nombre de proteïnes i el nombre de cromosomes, no és pas sempre 1 a 1 (vegeu la nota 4).

3. Els aminoàcids es poden simbolitzar amb un codi de tres lletres o bé amb un codi d'una lletra (per ex. la lisina, Lys o bé K; la leucina, Leu o bé L, etc.)

4. La identitat un gen, una proteïna, és una simplificació. De vegades, com en el cas dels anticossos, famílies de gens codifiquen per famílies de proteïnes, combinant la informació dels gens es tradueixen diferents proteïnes. En bacteris, es pot donar el cas que una mateixa seqüència gènica es comenci a transcriure a partir de diferents punts de la seqüència del gen, de manera que s'acaben traduint proteïnes diferents.

n'hi ha d'específics per a cada teixit. El conjunt de proteïnes que se sintetitzen en una cèl·lula s'anomena el «proteoma». El proteoma dependrà de la naturalesa del teixit al qual pertanyi la cèl·lula i de l'estat fisiològic en què es trobi la cèl·lula en aquell moment.

El procés d'activació i desactivació dels gens del genoma comença amb el desenvolupament embrionari. El procés biològic que porta des del zigot (la cèl·lula que resulta de la fusió de l'òvul amb l'espermatozoide) fins a l'individu adult s'explica en bona mesura per l'acció de les proteïnes que es va sintetitzant de manera coordinada en diferents regions de l'embrió. Una vegada les cèl·lules s'han diferenciat en un teixit determinat perden per sempre més la facultat de produir unes determinades proteïnes. En aquest fet, hi ha la base de la polèmica sobre el clonatge d'embrions per a fins mèdics, perquè és només en les primeres fases del desenvolupament embrionari que les cèl·lules mantenen encara la seva totipotència, és a dir, la possibilitat de produir qualsevol de les proteïnes codificades en el genoma. Les cèl·lules dels embrions joves tenen el potencial de transformar-se en qualsevol teixit. Aquest fet fa dels embrions un material molt valuós per a la gènesi *in vitro* de teixits i per al desenvolupament de nous agents terapèutics.

Entre el genoma dels humans i el dels ximpanzés hi ha una similitud inquietant. Tan sols un 1% dels caràcters del missatge genètic són diferents (Calafell, 2000). Per tant, immediatament pensem que quan sapiguem quines proteïnes codifica la part del missatge genètic que ens fa diferents del ximpanzé haurem fet un pas molt important per conèixer l'essència molecular humana; no obstant això, hi ha estudis que indiquen que les diferències obvies entre un ximpanzé i un individu de la nostra espècie poden ser més degudes a l'ordre amb el qual s'activen els gens durant al desenvolupament embrionari que no pas a les diferències de la seqüència genòmica (Jacob, 1998).

De manera que la capacitat d'adaptar la cèl·lula als possibles canvis fisiològics i el «pla» d'activació i desactivació de gens durant el desenvolupament embrionari s'ha de comptar també com a part de la informació genètica del DNA.

La bellesa del DNA rau en un fet tridimensional: és la clau per a l'estudi de l'organització del metabolisme, per a l'estudi del desenvolupament ontogenètic i també per a l'estudi filogenètic dels éssers vius. Acabem de veure a què es deguda la variabilitat de les cèl·lules d'un individu. Al final d'aquesta lliçó parlarem de les causes de la variabilitat entre els individus d'una mateixa espècie. Deixeu-me explicar ara un altre tipus d'informació que conté el missatge genètic.

Entre les diferents espècies, la informació genètica es diferencia per la llargada de la seqüència de bases i pel text que conté la seqüència. Els genomes més petits són els genomes dels virus, tenen entre tres mil i cent mil bases⁵; ben lluny dels tres mil milions de bases dels humans. Si es consideren totes les espècies existents, es pot observar que la mida del genoma és més gran segons s'avança en la complexitat de l'espècie. Les formes de vida unicel·lulars, com els bacteris, que ja disposen d'una maquinària cel·lular per multiplicar-se autònomament, tenen genomes que van de mig milió a deu milions de bases. A continuació, amb genomes entre deu i cinquanta milions de bases, hi trobem les formes de vida pluricel·lulars, amb poca comunicació intercel·lular, com ara els

5. L'any 1977 es va seqüenciar completament el genoma del bacteriòfag φX174, que té 5386 bases (F. Sanger *et al.*, 1977)

fongs. Finalment hi ha les plantes i els animals, que són formes de vida pluricel·lulars, amb funcions cel·lulars ben diferenciades, és a dir, amb teixits i molts missatgers moleculars. Les espècies que pertanyen a aquest darrer grup tenen genomes que van dels cinquanta als milers de milions de bases. No obstant això, tret del cas dels mamífers, ocells i rèptils, dins de cada grup hi ha una gran variació en la mida del genoma, fins al punt que algunes espècies sobrepassen a d'altres que gosem qualificar de més complexes. És el cas d'alguns amfibis que tenen un 60% més de DNA que l'espècie humana. Ja veurem, no obstant, que més DNA no vol dir necessàriament més gens, i per tant, tampoc més proteïnes.

A part de la mida, l'estructura física del genoma pot ser molt diversa. Les cèl·lules de les espècies més evolucionades, com ara les dels humans, tenen el material genètic duplicat⁶; el genoma d'aquestes cèl·lules es troba en forma de fragments lineals i confinat en un orgànel cel·lular, el nucli. Cadascun d'aquests fragments s'anomena cromosoma. L'espècie humana té 23 parelles de cromosomes; en total, doncs, 46. Els ximpanzés en tenen 48.

Malgrat les diferències que acabem d'esmentar, hi ha moltes parts de la seqüència gènica que es conserven amb més o menys similitud en totes les espècies, cosa que només s'explica bo i acceptant un determinat grau de parentiu entre tots els éssers vius.

Dues de les característiques més notables del DNA sembla que es contraposen. Per una banda, la molècula de DNA es pot replicar i transferir a la descendència de forma pràcticament inalterada al llarg de milers de generacions; de manera que els humans anem tenim fills humans i els micos, per posar un exemple, van tenint fills micos. Per altra banda, el DNA pot experimentar mutacions de manera natural, és a dir, el text que conté la seqüència d'informació genètica canvia al llarg del temps. En el cas més senzill, una mutació consisteix en la substitució d'una base de la seqüència per una altra base, no obstant, hi ha mutacions que afecten segments més o menys llargs del genoma. Segons el que hem dit més amunt, si un gen experimenta un canvi també l'haurà d'experimentar la proteïna que codifica i això pot afectar la funció d'aquesta proteïna. Val a dir, però, que el codi genètic té unes característiques que permeten atenuar en bona mesura l'efecte de moltes mutacions⁷.

La substitució d'un aminoàcid per un altre en una proteïna com a conseqüència d'una mutació genètica pot afectar la funció de la proteïna en diferent grau. La magnitud de les repercussions sobre l'organisme dependrà de quina sigui la funció d'aquesta proteïna. Per exemple, l'anomenada anèmia cel·lular falciforme és una malaltia deguda a una mutació en únic nucleòtid del gen de l'hemoglobina. Aquesta mutació provoca que els glòbuls rojos dels individus afectats tinguin forma de falç i que es puguin ocasionar problemes circulatoris (Vella, 1980). En canvi, una mutació en una de les proteïnes que intervenen en la formació dels cartílags de la rata té un efecte múlti-

6. Excepte en el cas de les cèl·lules germinals que són haploides, és a dir, amb una única còpia del material genètic.

7. Tots els aminoàcids, excepte la metionina i el triptòfan, estan codificats per més d'un codó (triplet de bases). Amb les quatre bases A, T, C i G es poden fer 64 combinacions de tres bases; aquestes combinacions es reparteixen entre els 20 aminoàcids. El codi genètic ha evolucionat de tal manera que una mutació en una de les bases comporta la traducció del mateix aminoàcid o d'un altre aminoàcid amb propietats fisicoquímiques semblants.

ple en l'anatomia dels animals afectats: engruixudiment de les costelles, constricció del conducte traquial, pèrdua d'elasticitat dels pulmons, hipertròfia del cor, narius tancats i morro rom (Strickberger, 1986). En molts casos, no obstant, les conseqüències d'una mutació són simplement letals.

És precisament aquesta capacitat d'experimentar canvis que té el DNA el que permet comprendre, en termes generals, la teoria de l'evolució de les espècies de Darwin. Com ja sabem, aquesta teoria explica com els canvis aleatoris que es produeixen en un únic individu poden arribar a canviar, al cap d'unes quantes generacions, el missatge genètic d'una espècie.

Un dels descobriments més útils, i a la vegada més sorprenents, per a l'estudi de l'evolució de les espècies ha estat el que s'anomena el «rellotge molecular» (Cavalli-Sforza, 1997). Els gens, com acabem de dir, experimenten mutacions i aquestes mutacions es poden veure reflectides en les proteïnes que codifiquen. Doncs bé, s'ha pogut observar que en espècies per a les quals es coneix el temps de separació en l'arbre filogenètic gràcies a mètodes paleontològics, el nombre d'aminoàcids diferents en una mateixa proteïna de les dues espècies es pot correlacionar molt bé amb el temps de separació. Això vol dir que la taxa de canvis per a una proteïna determinada és constant. De manera que comptabilitzant els canvis d'aquesta proteïna en dues espècies actuals es pot estimar quant temps fa que van començar a divergir, deixant enrere el seu ancestre comú. Utilitzant aquest mètode es va poder establir que la línia de l'arbre filogenètic que porta fins als humans es va separar de la línia que porta al nostre parent més proper, el ximpanzé, fa cinc milions anys⁸ (Ruvolo, 1997).

La nostra espècie, l'*Homo sapiens*, té una història que va començar molt possiblement a l'Àfrica fa aproximadament cent cinquanta mil anys. Els estudis de variabilitat genètica que van donar suport a l'origen africà dels humans es van portar a terme amb un material genètic una mica especial: el DNA mitocondrial. Els mitocondris són uns orgànuls cel·lulars que disposen d'una dotació de DNA diferent de la del nucli cel·lular. Es creu que aquests orgànuls provenen d'una antiga simbiosi entre els bacteris i les cèl·lules dels organismes superiors. Això s'hauria produït fa un milió d'anys. De fet, entre el DNA dels mitocondris i el DNA d'alguns bacteris actuals hi ha moltes similituds. A cada cèl·lula hi ha milers de mitocondris. La funció d'aquests orgànuls és produir energia química útil per a les reaccions metabòliques. En la fecundació, els mitocondris de l'espermatozoide no s'incorporen al zigot, es perden; de manera que els mitocondris del nou ésser són descendents dels mitocondris presents en l'òvul matern. El DNA d'aquest orgànuls es transfereix a la descendència per via materna. Aquesta característica⁹ del DNA mitocondrial facilita la construcció de genealogies, ja que no és la barreja del material genètic dels progenitors. Amb l'anàlisi del DNA

8. La formació d'una nova espècie a partir d'una espècie existent requereix com a primera condició l'aïllament sexual d'una població de l'espècie respecte a la resta d'individus, per exemple a causa d'una migració. Quan les dues poblacions han acumulat independentment un nombre crític de canvis en el genoma, a causa de mutacions, recombinació genètica i selecció natural, es perd la capacitat de reproducció entre els individus d'aquestes dues poblacions. No es tracta d'un impediment morfològic sinó d'un bloqueig de la reproducció a nivell molecular. Mentre l'encreuament entre els individus d'una població sigui possible, els canvis individuals s'aniran repartint aleatòriament al llarg de les generacions sense que es puguin establir dues línies evolutives diferents a dins de la població. Certament, aquesta situació seria materialment possible, però la probabilitat que es doni és molt petita.

9. A més, en el DNA mitocondrial no hi ha recombinació genètica i la taxa de mutació és més gran.

mitocondrial complet d'un japonés, un europeu i un africà i seguint el raonament que més variació en el DNA significa més temps de separació, es va poder comprovar que el DNA mitocondrial africà era el més antic (Horal *et al.* 1995). D'aquí ve que la revista de moda Vogue, sol·licités una entrevista a L.L. Cavalli-Sforza (el prestigiós geneticista de poblacions) perquè els expliqués quant temps feia que havia nascut l'*Eva africana* (Cavalli-Sforza, 1997). La construcció acurada de l'arbre filogenètic de les poblacions humanes actuals requereix, no obstant, l'estudi a fons de la variabilitat del DNA dels cromosomes del nucli. Aquests exemples il·lustren com el missatge genètic porta inclosa també informació sobre la història evolutiva dels organismes i com l'anàlisi de les diferències entre les espècies a nivell molecular pot permetre reconstruir l'arbre genealògic dels éssers vius.

Fem ara un petit càlcul¹⁰. Els éssers humans tenim aproximadament cent mil proteïnes. Segons el codi genètic, cada aminoàcid s'identifica en la seqüència del DNA mitjançant tres lletres de l'alfabet C, T, G i A¹¹. El nombre d'aminoàcids d'una proteïna típica és a l'entorn de quatre-cents. Una proteïna d'aquesta mida requereix, doncs, una seqüència de DNA de mil dues-centes bases. Per cada proteïna cal afegir-hi cinc-centes bases més que correspondrien a seqüències reguladores i als anomenats «introns»¹². Totes cent mil proteïnes requereixen, per tant, uns dos-cents setanta milions de caràcters del text genètic. Com que el genoma humà té una mida de 3340 milions de bases, hi ha d'haver aproximadament un 95% del DNA sense cap gen¹³. Les mutacions de què hem estat parlant no només es donen en els gens, de manera que entre els individus d'una mateixa espècie també hi haurà diferències genètiques en les zones no codificadores del seu DNA.

Per polimorfisme hem d'entendre qualsevol diferència que es pugui detectar en la seqüència gènica dels individus d'una mateixa espècie. Al llarg del genoma humà, per exemple, hi ha uns cent mil segments de DNA constituïts per repeticions de la parella de bases CA. Aquests blocs s'anomenen microsatèl·lits. Els microsatèl·lits poden tenir d'una a quaranta repeticions seguides de la parella de bases CA i es transmeten a la descendència conservant la mida del bloc. El polimorfisme consisteix en canvis de la mida d'un determinat microsatèl·lit en diferents individus de la població. Altres polimorfismes importants, que es troben tant en els gens com en la resta del DNA, són els anomenats «polimorfismes de dianes de restricció» (RFPL), que no comentarem, i els «snips», entre altres. «Snip» és la pronúncia anglesa de les sigles SNP, que traduïdes signifiquen «polimorfisme nuclear simple». Els SNP són diferències en una única base, en alguna posició del missatge genètic, dels individus de la població. Actualment s'està elaborant un base de dades en la qual es van introduint els SNP que es van descobrir (<http://snp.cshl.org>). Es preveu que a la primavera del 2001 s'haurà construït un mapa del genoma humà amb la posició i naturalesa d'aproximadament set cents mil polimorfismes nuclears simples (Roses, 2000). Un dels avantatges dels SNP és que s'està

10. Aquest càlcul només pretén ser il·lustratiu.

11. Vegeu la nota 7.

12. Els introns són seqüències del DNA d'un gen que no es troben en el RNA missatger madur.

13. Aquest DNA que no codifica cap proteïna es considera DNA estructural. Sembla ser que és necessari perquè els gens es puguin transcriure a més velocitat (Baldi i Brunak, 1998).

desenvolupant la tecnologia, els anomenats «xips de DNA»¹⁴, que ha de permetre portar a terme l'anàlisi simultània de milers de SNP, automàticament.

Hem vist més amunt que les cèl·lules dels humans tenen una dotació doble de material genètic; per tant, cada individu pot tenir com a màxim dues versions d'un determinat polimorfisme, una versió situada en cadascun dels dos cromosomes homòlegs. Les diferents versions d'un gen s'anomenen *clàssicament* al·lels, aquest concepte s'ha fet ara extensiu a tots els polimorfismes, es trobin o no en una seqüència codificadora. Un exemple familiar per a tothom són els al·lels que determinen el grup sanguini dels humans.

En el procés que porta a la formació dels gàmetes masculí i femení, els quals només tenen una còpia de material genètic, és a dir, que els cromosomes no hi estan duplicats, es produeixen dos fenòmens que són la causa principal de la diversitat dels individus d'una població. Per una banda, els 23 cromosomes d'un òvul o un espermatozoide poden provenir de qualsevol dels cromosomes de la parella de l'individu adult. Això fa que es puguin produir més de vuit milions (2^{23}) de tipus diferents de gàmetes en relació a la seva dotació cromosòmica. Per altra banda, al llarg del procés de formació dels gàmetes, mitjançant un fenomen anomenat «recombinació genètica», els cromosomes homòlegs s'intercanvien material genètic, de manera que els diferents al·lels es redistribueixen al llarg del cromosoma, cosa que fa augmentar encara més el nombre possible de gàmetes genèticament diferents. Aquesta és la causa per la qual els germans d'una família són diferents; els fills reben una combinació aleatòria dels trets dels pares.

Gràcies a l'impuls del Projecte Genoma Humà¹⁵, a la base de dades GenBank¹⁶, als EEUU, s'hi entren 1.700.000 noves bases per dia pertanyents a seqüències genètiques de representants de tots els grups d'espècies (Baldi i Brunak, 1998). Aquestes seqüències són a l'abast de tothom a través d'Internet de forma gratuïta. En aquesta mateixa base de dades es pot trobar la seqüència completa de diversos organismes, entre els quals hi ha *Escherichia coli* (un bacteri), *Saccharomyces cerevisiae* (un llevat), *Arabidopsis thaliana* (una planta), una bona part de la seqüència de *Drosophila melanogaster* (un animal), els cromosomes 21 i 22 humans i milers de fragments del missatge genètic de centenars d'espècies. És clar que l'estudi de la filogenètica dels éssers vius mitjançant l'anàlisi de la diversitat genètica de les espècies actuals té un gran futur. I aquí, cal afegir-hi, que la tecnologia per analitzar el DNA d'organismes fòssils aviat serà una pràctica a l'abast de tots els laboratoris, cosa que permetrà afinar, i qui sap si fins i tot canviar, el coneixement actual sobre la història de l'evolució.

La possibilitat de determinar amb mètodes fiables i ràpids (com en el cas dels SNP) la diversitat del genoma de cada individu, ultra l'interès purament científic que com hem vist suposa, ha de permetre diagnosticar i tractar amb més eficàcia moltes de les malalties que per ara s'escapen

14. Es tracta d'uns suports, d'aproximadament un centímetre quadrat, en els quals s'immobilitzen les sondes de DNA adequades per poder detectar un determinat conjunt de SNP.

15. Amb la tecnologia de l'any 1993 es podien seqüenciar 90 milions de bases per any, amb un cost de 100 pts per base; actualment es seqüencien 500 milions de bases per any, a un preu de 50 pts per base.

16. <http://www.ncbi.nlm.nih.gov>

als tractaments farmacològics. La farmacogenètica és l'estudi de com les diferències genètiques condicionen la resposta del pacients a un determinat medicament. Virtualment, tota malaltia de caràcter no infeccios té al darrere la disfunció d'una proteïna. Sovint, els fàrmacs es dissenyen per actuar sobre aquestes proteïnes. Ara bé, petites diferències en una proteïna degudes a la variabilitat genètica de l'individu poden atenuar i fins i tot invalidar els efectes d'un fàrmac. O, mirat des d'un altre punt de vista, poden provocar uns efectes secundaris més perniciosos en uns individus que en altres (Roses, 2000). De manera que és possible de concebre (i de fet les companyies farmacèutiques ja ho estan explorant) fàrmacs per a una determinada afecció, específics per al perfil genètic del malalt.

Però a part de tot el que acabem de dir, la ingent quantitat de dades generades pel Projecte Genoma Humà comencen a canviar la manera d'estudiar el fenomen vital.

Els organismes són xarxes de gens que fan xarxes de proteïnes que regulen els gens¹⁷. Per tant, veiem que una altra de les funcions de les proteïnes que no hem amononat al principi de la lliçó és precisament interactuar físicament amb el DNA per portar el control de la síntesi de proteïnes. És un peix que es mossega la cua. Qui controla, doncs, el sistema de control? La resposta és que el sistema s'autocontrola. Aquesta conclusió no sorprendrà gens els enginyers, que estan acostumats a tractar amb mecanismes d'aquest tipus. Ara bé, si es vol tractar el funcionament molecular d'un ésser viu com un sistema dinàmic, cal conèixer prèviament quins són els elements de què està format i com estan relacionats entre ells. El Projecte Genoma Humà obre la possibilitat d'emprendre aquest estudi. Es tracta de passar d'un enfocament reduccionista de la biologia a un enfocament holístic. Per això cal desenvolupar les eines conceptuals, matemàtiques i informàtiques que han de permetre detectar els gens en el desert que representa la seqüència genòmica, deduir a partir de la seqüència del gen la funció de la proteïna que codifica (d'això se'n comença a ocupar un nou camp de la ciència anomenat «genòmica funcional»), detectar llavors les proteïnes que interactuen per portar a terme una determinada funció i establir finalment la dinàmica d'aquesta xarxa. Feina de sobres per al segle XXI.

Arribats a aquest punt de la lliçó, la pregunta que ens podem fer és si el coneixement complet del missatge que conté el DNA ens permetrà descobrir alguna cosa essencial de la naturalesa humana. Jo no gosaria pas pronunciar-me. No obstant, m'agradaria posar dos exemples de

17. La informació continguda en el DNA no es tradueix directament a una seqüència d'aminoàcids sinó que ho fa a través de la síntesi d'un intermediari molt semblant en tots els aspectes al DNA, l'anomenat RNA, àcid ribonucleic. De manera que la informació per a la síntesi de proteïnes flueix en el sentit DNA, RNA, proteïnes. El procés invers no es coneix. Si bé alguns virus, els anomenats retrovirus (entre els quals hi ha el virus de la SIDA) poden produir una molècula de DNA a partir de RNA, no es coneix cap procés bioquímic capaç de produir una molècula d'un àcid nucleic a partir de la informació continguda en una proteïna. Quan s'intenta comprendre l'origen de la vida conforme a una evolució química de tipus darwiniana es topa amb una paradoxa del tipus «què fou abans, l'ou o la gallina?»: la informació per a la síntesi de les proteïnes es troba continguda en el DNA i el DNA no pot sintetitzar proteïnes sense la intervenció, precisament, de les proteïnes. Una possible solució de la paradoxa passa pel descobriment dels ribozimes: molècules de RNA amb activitat enzimàtica (Schuster, 1995).

proteïnes relacionades amb fenòmens mentals. En la mosca del vinagre, *Drosophila melanogaster*, s'han pogut aïllar dues proteïnes que intervenen en la memòria a llarg termini, l'una activa aquest tipus de memòria i l'altra l'inhibeix. Les mosques transgèniques dissenyades perquè sintetitzin una quantitat més gran de la proteïna activadora, només necessiten una sessió d'entrenament per aprendre a associar una determinada olor amb un corrent elèctric; les mosques normals requereixen dues sessions (Connolly i Tully, 1996; Tully, 1998)¹⁸. El segon exemple està relacionat amb la consciència. En la transmissió de senyals al llarg del sistema nerviós hi intervenen unes estructures proteiques, els microtúbuls, que formen l'anomenat esquelet cel·lular o citoesquelet. La pèrdua de consciència causada per anestèsics com ara l'èter o el cloroform és deguda a un efecte sobre la tubulina, la proteïna que forma els microtúbuls. La tubulina es pot trobar en dues formes tridimensionals diferents, que si volem podem simbolitzar amb un 0 i un 1. Tot i que no es disposa de proves definitives, sembla ser que l'anestèsic bloquejaria la «commutació» entre les dues formes; la tubulina no podria passar de 0 a 1 o de 1 a 0. Des d'un punt de vista matemàtic, els microtúbuls tenen una estructura molt interessant: són cilindres buits constituïts per 13 columnes de molècules de tubulina. Si ens imaginem el cilindre desplegat en un pla, les molècules de tubulina s'ordenen en línies inclinades que arriben del centre als extrems oposats amb un desplaçament de 5 o 8 llocs. El 5, el 8 i el 13 són tres nombres consecutius de la sèrie de Fibonacci. La disposició de les molècules de tubulina conforme a nombres de Fibonacci pot afavorir, segons diuen els matemàtics, la funció del microtúbul com a «processador d'informació». Els microtúbuls i la tubulina podrien portar a terme computacions a l'estil d'un «autòmata cel·lular» i transportar els missatges al llarg del sistema nerviós (R. Penrose, 1996).

M'agradaria acabar aquesta lliçó amb unes paraules d'Umberto Eco, extretes del seu llibre *Sis passejades pels boscos de la ficció*. Diu el següent: «Hi ha una regla àurea per a tot criptoanalista o descriptador de codis secrets, i és que tot missatge pot ser descriptat sempre que se sàpiga que es tracta d'un missatge. El problema amb el món real és que ens estem demanant, des de fa mil·lennis, si hi ha un missatge i si aquest missatge té un sentit».

Moltes gràcies.

18. Exemple extret de «La comprensió de la conducta: aportacions de la genètica molecular» de Rosa M. Escoriuela, dins *La biologia a l'alba d'un nou mil·lenni*. Treballs de la Societat Catalana de Biologia, Volum 50. Barcelona: Jaume Bertranpetit Editor, 2000.