

L'EVOLUCIÓ VISTA A TRAVÉS DE L'ESTUDI DE LA COMPOSICIÓ DE LES PROTEÏNES. HIPÒTESIS EVOLUTIVES DE PROTAMINES I D'HISTONES

per J. PALAU

INTRODUCCIÓ

Actualment, hom coneix força bé determinades característiques estructurals de les proteïnes bàsiques que acompanyen l'àcid desoxiribonucleic dins el complex cromosomàtic dels mamífers i d'altres espècies animals i vegetals. Així, hom ha aïllat cinc components d'histones ben caracteritzades per l'anàlisi electroforètica i d'aminoàcids. Similarment, les protamines típiques presents en l'esperma d'alguns peixos també han estat fraccionades i caracteritzades com a components polipeptídics individuals. Al llarg dels darrers dos o tres anys, diversos laboratoris han fet un esforç considerable per a determinar l'estructura primària de les histones i de les protamines, esforç que ha permès d'obrir les portes a l'estudi estructural, funcional i evolutiu d'aquestes proteïnes. La present comunicació representa un intent d'aprofundir en el significat evolutiu dels coneixements actuals sobre seqüència d'aminoàcids de les protamines i de les histones, tenint en compte les limitacions i ambigüitats que, encara actualment, presenta l'estudi de l'evolució de les proteïnes en general.

Nota. — El codi d'una sola lletra emprat en aquesta comunicació es: A, Alanina; B, Acid Aspàrtic o Asparagina (sense decidir); C, Cisteïna; D, Acid Aspàrtic; E, Acid Glutàmic; F, Fenilalanina; G, Glicina; H, Histidina; I, Isoleucina; K, Lisina; L, Leucina; M, Metionina; N, Asparagina; O, Tirosina; P, Prolina; Q, Glutamina; R, Arginina; S, Serina; T, Treonina; V, Valina; W, Triptòfan; Z, Acid glutàmic o glutamina (sense decidir).

L'EVOLUCIÓ DE LES PROTEÏNES I EL SEU SIGNIFICAT DINS EL CONTEXT
GENERAL DE L'EVOLUCIÓ BIOLÒGICA

Bé que des de fa molt de temps han estat estudiades determinades variacions genètiques a nivell molecular, com és ara les dels pigments de certes plantes, només des de fa uns 20 anys hom ha començat a obtenir dades importants sobre el desenvolupament de l'evolució molecular. En el sistema biomolecular ADN→ARN→Proteïna hi ha, evidentment, un emmagatzematge i una transferència d'informació bioquímica. Qualsevol canvi, al llarg de l'evolució, del contingut d'informació en l'ADN transcripible, serà reflectit en un canvi de contingut d'informació en l'ARN i en les proteïnes del sistema. És a dir, una variació en el fenotip proteïna reflecteix una variació en el genotip ADN. Tenim clarament, doncs, una correspondència a escala molecular anàloga a la clàssica de fenotip-genotip postulada per a les espècies biològiques.

D'altra banda, el conjunt de fenotips moleculars i llurs variacions han de correspondre naturalment al conjunt de fenotips morfològics i funcionals observables en les espècies vivents, així com a llurs variacions. Per tant, un estudi de l'evolució dels fenotips moleculars, en principi ha de donar informació tan representativa de la història dels éssers vivents com pugui donar-la l'estudi de l'evolució d'altres fenotips més clàssics.

La impossibilitat actual d'avaluació global de variacions de fenotips moleculars és òbvia, i la raó més senzilla d'això és perquè hom desconeix moltes dades particulars, estructurals i funcionals de la major part de proteïnes. Per tant, tota especulació evolutiva muntada sobre el coneixement de variacions estructurals i funcionals d'una proteïna cal limitar-la, en principi, a una especulació únicament per a aquella proteïna. L'extrapolació al fet evolutiu general pot ésser greument perillosa, en especial si els resultats obtinguts no es corresponen amb els acumulats fins al present sobre conjunts fenotípics més amplis.

Fent ja una limitació a l'estudi de l'evolució de les proteïnes com a tals, podríem dur a terme una simplificació important si consideràvem que és vàlid el postulat que la seqüència d'aminoàcids d'una proteïna dicta les seves estructures secundària, terciària i quaternària, és a dir dicta la seva conformació. La simplificació seria de considerar la seqüència d'aminoàcids, per ella mateixa, l'única peça important per a l'estudi de l'evolució de les proteïnes. Malauradament aquesta simplificació seria errònia.

Seguint el procés evolutiu general, l'evolució d'una proteïna és el resultat del doble mecanisme mutació-selecció natural. La selecció natural resulta d'un enfrontament entre conformació funcional i medi ambient.

Ara bé, conformacions proteiques no funcionals en certs medis de pH, de força iònica, natura d'ions, etc. poden resultar funcionals en altres medis. També cal considerar com a modificadors de la conformació, i per tant de la funció, tota mena d'interaccions allostèriques i de cofactors necessaris per a l'expressió de la dita funció.

Finalment, una proteïna, per canvi d'un aminoàcid, pot passar a inactiva, i per canvi posterior d'un altre aminoàcid, de posició i natura diferents de les del primer, pot revertir a activa. Això pot donar lloc a variacions de seqüència ambigüament interpretades com a producte de mutacions tolerades, quan són mutacions prohibides. Com a confirmació del que hem expressat anteriorment, podem esmentar el fet que es troben proteïnes de seqüència molt diferent que tenen funcions i propietats molt semblants i, viceversa, proteïnes de seqüència molt semblant que tenen funcions i propietats molt diferents.

Per tant, per a conèixer inequívocament relacions evolutives de les proteïnes, cal combinar les dades de seqüència, d'estudis de difracció de raigs X i de representació dels centres actius, químicament i enzimàticament. En aquest nivell podem dir que en l'actualitat hom té alguna informació d'unes poques proteïnes: hemoglobina, mioglobina, ribonucleasa, quimotripsina (i quimotripsinogen), citocrom C i lisozim. El panorama, de moment, és poc encoratjador. No obstant això, aprofundint en l'estudi del conjunt de seqüències d'aminoàcids de diverses proteïnes extretes d'espècies pertanyents a diferents tipus d'animals i de vegetals, hom pot trobar certes relacions de significat evolutiu que val la pena d'assenyalar:

A. Per a una mateixa proteïna

1. Homologies internes de seqüència. Són interpretades com un resultat de la duplicació de gens ancestrals que inicialment codificaven oligopèptids. Per duplicacions successives, parcials o totals, i fusió dels duplicats es produí un gen que codificava un prototipus de proteïna ancestral de la qual provenen les actuals.

2. Constància de residus d'aminoàcids dins la seqüència. Això és altament indicatiu que la funció biològica és determinada almenys per alguns d'aquests residus.

3. Substitució direccional, al llarg de l'escala evolutiva, d'uns aminoàcids per uns altres. És indicatiu de millora funcional de la proteïna i, per tant, indicatiu de mutacions afavorides.

4. Substitució indiferent, al llarg de l'escala evolutiva, d'uns aminoàcids per uns altres. Pot indicar invariabilitat funcional i, per tant, assenyalar la presència de mutacions neutres.

B. *Per a grups de proteïnes*

Podem considerar les següents possibilitats:

1. Diverses proteïnes tenen una seqüència i una funció biològica similars i hom sospita que els gens que les codifiquen provenen, probablement per duplicació i posterior evolució, d'un gen ancestral comú. El procés evolutiu és conegut com a procés de divergència i les proteïnes són considerades proteïnes homòlogues. Hom diu que presenten isologia de seqüència.

2. Dues proteïnes o més tenen una seqüència i una funció similars i hom sospita que els gens que les codifiquen provenen de gens ancestrals diferents. El procés evolutiu és conegut com a procés de convergència, i les proteïnes són considerades anàlogues amb isologia de seqüència.

3. Dues proteïnes o més tenen una seqüència diferent i una funció similar. Hi ha una gran probabilitat que els gens que les codifiquen provinquin per convergència de gens ancestrals diferents. Les proteïnes són considerades anàlogues sense isologia de seqüència.

Realment és un problema difícil de distingir a la pràctica entre convergència i divergència. Normalment seran considerades proteïnes homòlogues quan tindran propietats funcionals similars i moltes propietats químico-físiques comunes, i quan semblarà probable un origen genètic comú. La demostració rigorosa d'homologia o d'analogia requereix un coneixement profund de seqüència d'aminoàcids, de funció, fonts del material i altres informacions de tipus biològic.

L'EVOLUCIÓ DE LES PROTAMINES

En general les protamines es caracteritzen pel fet de tenir al voltant d'uns trenta residus, dels quals les dues terceres parts aproximadament són arginina. Hom coneix actualment l'estructura primària de set protamines, i, de la comparació entre aquestes estructures, hom pot treure informació valuosa, que pot ésser resumida de la manera següent:

1. Les espècies biològiques les protamines de les quals han estat estudiades, contenen 3 o 4 components d'aquestes proteïnes.

2. La seqüència de totes les protamines estudiades és molt similar.

3. Hom troba l'arginina dins la protamina en forma d'agrupaments de 4 a 6 residus de distribució irregular.

4. Els residus altres que l'arginina són pocs (A, G, V, I, T, P i S) i hom els troba també agrupats.

Fa uns pocs anys, BLACK i DIXON³ establiren una hipòtesi sobre l'evo-

lució de la protamina clupeïna Z, estudi que pot ésser estès a les altres protamines de seqüència coneguda, fent-hi algunes adaptacions: Inicialment hauria existit el pèptid ancestral AR₄, probablement codificat per un ADN ancestral. Per duplicacions successives amb fusió dels duplicats s'hauria transformat en AR₄AR₄AR₄AR₄. A continuació, per duplicacions parcials (insercions) haurien arribat a una forma de protamina ancestral comuna AR₄A₂RAAR₄ARAAR₄AR₄. Més tard, per duplicacions redundants del gen ancestral, seguides de duplicacions parcials (insercions) específiques, de delecció (en un sol cas), i de mutacions (normalment d'una base única, i en un cas de dues bases), haurien arribat a les formes actuals de les protamines (fig. 1). En la figura 2 hem col·locat les seqüències conegudes de les protamines de la forma més adient per a descriure l'evolució probable d'aquestes proteïnes, d'acord amb la hipòtesi establerta ⁷.

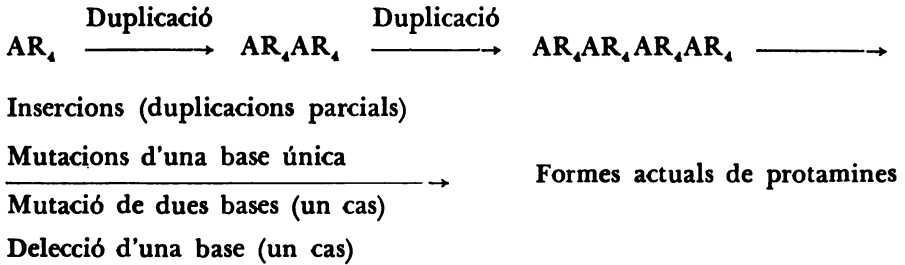


FIG. 1. — Evolució probable de protamines, generalització de la hipòtesi de Black i Dixon (1967)

Clupeïna Y-I	ARRRR	S	S	SR	P	IRRRR
Clupeïna Y-II	PRRR	TRR ..	A	SR	P	VRRRR
Clupeïna Z	ARRRR	S RR ..	A	SR	P	VRRRR
Salmina A-I	PRRRR	S	S	SR	P	VRRRR
Iridina Ia	PRRRR	S	S	SR	P	VRRRR
Iridina Ib	PRRRR	.. RR	S	S	SR	P	IRRRR
Iridina II	PRRRR	S	S	SR	P	VRRRR
	PRRR .	T	TRRRR	A	GRRR
	PRR .	V	SRRRR	ARRR
	PRR .	V	SRRRR	ARRR
 R	PR ..	V	SRRRR RR	G	GRRR
 R	PRR .	V	SRRRR RR	G	GRRR
	PRR .	V	SRRRR R	G	GRRR
	ARR ..	V	SRRRR RR	G	GRRR

FIG. 2. — Estructura primària d'algunes protamines, adaptació de l'esquema de Phillips (1971)

La figura 3 mostra les mutacions probables dels aminoàcids de les protamines, que corroboren la hipòtesi.

Finalment, presentem en la figura 4 l'evolució probable dels gens ancestrals de les protamines.

Cal destacar la relació d'aquest gen ancestral amb l'ADN repetitiu de tipus G-C. Potser tenim una pista per a atribuir en general a aquest tipus

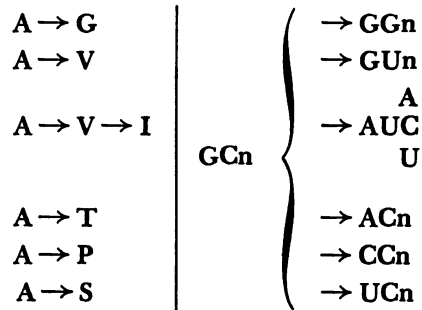


FIG. 3. — Mutacions probables per a les protamines

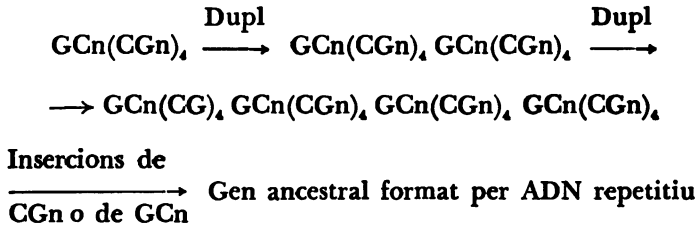


FIG. 4. — Evolució probable dels gens ancestrals de les protamines

d'ADN, conegut també per ADN satèl·lit, una procedència directa del material genètic ancestral. Si fos així, la permissibilitat evolutiva hauria estat mínima, i el cromosoma necessitaria regions de superprotecció on les possibilitats de mutació del material genètic fossin mínimes. Com que l'ADN repetitiu sembla estar localitzat en el centròmer, aquesta podria ésser una regió de superprotecció. La hipòtesi, a primera vista, pot semblar heterodoxa. Hem de tenir en compte, però, que els cromosomes sexuals X també semblen tenir superprotecció del material genètic. D'altra banda, el fet particular que una quantitat mínima de material genètic no seguis necessàriament les lleis de l'atzar i de la necessitat, no invalidaria el fet general

del procés de selecció natural, i, per contra, el complementaria donant-li les seves dimensions correctes. Més endavant, per a les histones, tornarem a insistir sobre la hipòtesi de superprotecció.

ESPECULACIONS EVOLUTIVES SOBRE LA HISTONA F_{2a1} O HISTONA IV

Dins l'àrea de recerca sobre evolució de proteïnes causà un gran impacte el resultat obtingut pel prof. SMITH i col·laboradors l'any 1969 en establir les seqüències de la histona F_{2a1} (IV) de dues espècies vivents tan diferents com són el vedell i el pèsol^{4,5}. De 102 residus d'aminoàcids només trobem dos canvis. El residu 60 és valina per al timus de vedell, i isoleucina per a l'embrió de pèsol; i el residu 77 és lisina per al primer, i arginina per al segon. Aquestes variacions es refereixen a aminoàcids molt similars i, per tant, no poden fer variar greument l'estructura de la proteïna. D'altra banda, les mutacions que els corresponen són de tipus conservatiu, és a dir, impliquen el canvi d'una sola base de l'ADN.

El fet de la invariabilitat quasi perfecta és espectacular i no es dona en cap altre tipus de proteïna de seqüència estudiada. Així, per exemple, el nombre de residus invariables establerts l'any 1967 per al citocrom C tipus mamífer, aïllat de diferents espècies, era de 35 sobre 104⁶.

Els processos de convergència i de divergència assenyalen una característica de funcionalitat plàstica per a les proteïnes al llarg de l'evolució. Llavors ens hem de preguntar: ¿quines són les raons biològiques que fan que la histona F_{2a1} (IV) no hagi pogut canviar al llarg de mil o dos mil milions d'anys? ¿És que aquesta proteïna té una funció tan precisa que no admet el canvi de quasi cap aminoàcid? ¿Havia evolucionat abans d'adquirir l'estructura primària actual? Tractarem de trobar respostes especulatives a aquestes qüestions.

Dins la concepció actual, la histona F_{2a1} (IV) hauria de tenir una funció extraordinàriament precisa que impossibilités qualsevol mutació al gen que la codifica. La seva conformació, per ella mateixa o per estar associada a altres proteïnes o a l'ADN, ha de restar perfectament definida de manera que tota la molècula participi en la funció biològica. Pràcticament qualsevol substitució d'un aminoàcid per un altre ha d'ésser letal per als individus de les espècies vivents.

Una funció atribuïble a les histones és la d'ésser reguladores de l'activitat de transcripció de l'ADN. Particularitzant per a la histona F_{2a1} (IV), aquesta regulació podria venir establerta, segons la hipòtesi d'ALLFREY i col·laboradors, per l'acetilació i metilació d'uns determinats residus de la molècula¹. Fa poc, a Barcelona, el doctor GERSHEY, del grup de la Universitat Rockefeller, ens explicava que aquesta histona podia actuar com

un enzim del tipus acetiltransferasa, per a acetilar-se ella mateixa. Aquesta proteïna seria al mateix temps enzim i substrat! Les acetilacions de residus bàsics farien disminuir la interacció d'aquesta histona amb l'ADN, la qual cosa ajudaria a iniciar el procés de transcripció.

Si la hipòtesi abans esmentada podia ésser demostrada inequívocament, hauríem donat un pas gegantí dins l'estudi de la regulació gènica dels organismes eucariotes. Sembla convenient, doncs, d'estudiar altres possibilitats, i establir altres hipòtesis alternatives que puguin explicar també el fet d'invariabilitat evolutiva. Si aquestes hipòtesis podien ésser rebutjades d'una manera raonable, sortiria enfortida la concepció de funcionalitat precisa de la histona F_{2a1} (IV).

Resultats obtinguts conjuntament amb els doctors SUBIRANA i RUIZ-CARRILLO indiquen que probablement les histones de l'esperma d'equinoderms són codificades per uns altres gens diferents dels que codifiquen les histones somàtiques⁶. La possibilitat que això sigui veritat ens ha emmenat a estudiar junt amb el doctor PARELLÓ la seqüència de la histona F_{2a1} (IV) de l'esperma d'eriçó de mar per tal de veure si trobàvem més abundor de variacions. Els resultats actuals ens demostren que la regió C-terminal de 19 aminoàcids és idèntica, si no tenim en compte un eventual canvi d'àcid glutàmic per glutamina. Serà molt interessant d'intentar de localitzar la posició de l'aminoàcid cisteïna en aquesta fracció, detectat pel doctor SUBIRANA⁶, puix que una variació com aquesta pot ésser d'una importància evolutiva extraordinària. Els treballs es troben encara en una fase experimental i no en podem treure altres conclusions sinó que hi ha una gran similitud de composició i de seqüència (parcial) en comparar la histona F_{2a1} (IV) de l'esperma d'eriçó de mar amb la fracció somàtica.

D'un altra punt de vista, si observem la seqüència de la histona

Posició	Seqüència A	Seqüència B
1-9	SGRGKGG - K - - -	
10-19	- GLGKGGAKRHR	
20-27		K - V - LRDN - IQ -
28-36	- GITKP - AIR - R	
37-45	- - LARGGVKR - R	
47-49	SGL	
54-48	- - - TRGVL - - - -	
59-68		K - VFL - ENVIRD
69-76		- AVTYTE - - - HA
77-78	- - - - - KR - -	
79-90		KTVTAMDVVYAL
99-102	- - - GFGG	

FIG. 5. — Homologies internes de la histona F_{2a1} (IV), segons K. Bauer (1971)

F_{2a1} (IV) podem veure-hi homologies internes (fig. 5). Aquestes homologies semblen indicar l'existència ancestral d'un oligonucleòtid o més que codificaven oligopeptids precursors de l'actual histona². Per duplicacions totals i parcials, per deleccions i per mutacions haurien arribat a un gen ancestral, que és pràcticament el gen actual. Immediatament apareix una situació interessant: durant el temps més ancestral, l'evolució ha construït una molècula seguint un procés de selecció natural. En un moment donat, l'evolució ha restat congelada i la molècula ja no ha evolucionat més. Voldríem aquí introduir una altra vegada la hipòtesi de superprotecció de l'ADN, ja esmentada per al cas de l'ADN repetitiu. El gen codificador de la histona F_{2a1} (IV) hauria passat a una àrea de superprotecció de l'ADN en un moment donat de la història del material genètic, i des de llavors aquest gen s'hauria mantingut invariable.

La funció actual de la histona F_{2a1} (IV) seria la que tenia la histona en el moment de congelar-se l'evolució, però la dita funció no seria necessàriament d'una precisió extraordinària, bé que podria ésser mantinguda com a funció essencial per al funcionament normal de la cèl·lula.

CONSIDERACIONS FINALS

Creiem que controvèrsies com les expressades en aquest treball són estimulants sobretot en uns moments d'extraordinària expectació pels descobriments fets dins el camp de recerca de les histones i de les protamines. Per això, hem cregut adient de portar a la discussió d'aquesta taula rodona algunes de les noves troballes, així com les diverses especulacions que hi poden ésser fetes.

BIBLIOGRAFIA

1. ALLFREY, V. G. — «Histones and Nucleohistones», Phillips, D. M. P. (Ed.), Plenum Press, 260 i 264, Londres (1971).
2. BAUER, K. — «Internat. J. Prot. Res.», 3, 165 (1971).
3. BLACK, J. A. i DIXON, G. H. — «Nature», 216, 152 (1967).
4. DELANGE, R. J., FAMBROUGH, D. M., SMITH, E. L. i BONNER, J. — «J. Biol. Chem.», 244, 319 (1969).
5. — «J. Biol. Chem.», 244, 5669 (1969).
6. PALAU, J., RUIZ-CARRILLO, A. i SUBIRANA, J. A. — «Eur. J. Biochem.», 7, 209 (1969).
7. PHILLIPS, D. M. P. — «Histones and Nucleohistones», Phillips, D. M. P. (Ed.), Plenum Press, 65, Londres (1971).
8. STEVENS, F. G., GLAZER, A. N. i SMITH, E. L. — «J. Biol. Chem.», 242, 2764 (1967).
9. SUBIRANA, J. A. — «FEBS Lett.», 16, 133 (1971).

DISCUSSIÓ

GADEA

¿Quin paper representen per a la filogènia —tan íntimament relacionada amb l'evolució— l'homologia, l'analogia, la isologia i la convergència de les proteïnes? Evidentment, l'homologia és sempre, en aquest aspecte, una dada positiva, mentre que no ho són ni l'analogia ni la convergència; la qüestió és més difícil en el cas de la isologia, puix que aquesta pot ésser d'origen homòleg o anàleg. ¿Hi ha possibilitat de conèixer quan hom tracta d'un tipus o de l'altre?

PALAU

Al meu entendre, caldria parlar d'homologia i d'analogia, no únicament de seqüència sinó també de funcionalitat, que és un terme de significat més ampli. Sense considerar els aspectes d'estructura terciària, és difícil de saber inequívocament quan una proteïna ha d'ésser considerada homòloga o bé anàloga.

PREVOSTI

Crec que el factor determinant de l'homologia és la seqüència i que poden existir molècules homòlogues amb funció diferent. Avui en dia hi ha algunes dades sobre homologia o analogia de les proteïnes. Hi ha una base forta per a considerar que són homòlogues si l'estructura primària és semblant.

PALAU

En casos clars tenim com a pista la seqüència d'aminoàcids. Això és ben cert, però a la pràctica és difícil de discutir per a moltes proteïnes si es tracta de convergència o de divergència.

Rigorosament, perquè dues proteïnes siguin considerades homòlogues cal que hi hagi: 1) Seqüència similar. 2) Propietats funcionals similars. 3) Propietats físico-químiques comunes. 4) Probable origen genètic comú.

PARELLÓ

Segons DICKERSON pot existir un gran canvi de seqüència però amb conservació de l'estructura terciària. La selecció podria actuar en el camp de la funcionalitat.

PREVOSTI

La selecció natural pot seleccionar la mateixa forma o funció amb punts de partida diferents. Com més ens allunyem del nivell primari, més possibilitats hi ha d'analogia.

PARELLÓ

Realment ens falta informació sobre les seqüències de l'ADN.

PREVOSTI

Però hom coneix el codi dels nucleòtids, la qual cosa permet de traslladar el coneixement de la seqüència al camp de l'ADN.

PARELLÓ

Hem de tenir en compte, no obstant això, la degeneració del codi.

PREVOSTI

Sí, bé que la degeneració afecta fonamentalment una base del triplet.

PARELLÓ

Sembla important l'estudi d'espècies intermèdies. Pot donar pistes sobre un origen comú o no.

PARÉS

En els bacteris no hi ha protamines ni histones. En relació amb la hipòtesi de superprotecció de l'ADN, ¿hi ha algun comentari a fer des del punt de vista comparatiu?

PALAU

Realment, per als eucariotes estem en el camp de la hipòtesi. Crec adient de llegir un paràgraf d'un article aparegut a «Science» el 17 de desembre de 1971, sobre *Heterocromatina, ADN satèl·lit i funció cel·lular* (YUNIS, J. J. i YASMINEH, W. G., «Science», 174, 1200 [1971]). Tradueixo de l'anglès: «L'heterocromatina constitutiva pot ésser considerada un tipus especial de cromatina que conté la major part de l'ADN satèl·lit, o sigui que conté les seqüències d'ADN altament repetitiu del genoma que no són transcrites a ARN per la síntesi de proteïna. Blocs d'aquestes seqüències estan generalment localitzats a les regions dels organitzadors nucleolars, centròmers i telòmers, o s'intercalen algunes vegades dins altres regions dels cromosomes. Una avaluació del coneixement sobre heterocromatina constitutiva i ADN satèl·lit suggereix que els papers de l'heterocromatina constitutiva són de natura estructural: 1) Pot protegir àrees vitals del genoma de forces externes disruptives i de canvis evolutius. 2) Pot atraure cromosomes homòlegs per una alineació inicial durant la miosi, i atraure cromosomes no homòlegs per establir proximitat entre cromosomes o regions de cromosomes que funcionalment estiguin relacionats. 3) Pot establir *barreres de fertilització* que proporcionin mitjans per a la diversitat evolutiva i per a l'especiació.»