

## AMPLIFICACIO, SEQUENCIACIO, EVOLUCIO I FUNCIO DE REGIONS PROMOTORES DELS GENS DE PROTAMINA P1 DE MAMIFERS.

Rosa Queralt i Rafael Oliva

Grup de Genètica Molecular, Facultat de Medicina, Universitat de Barcelona

### Resum.

Fent servir oligonucleòtids complementaris de regions conservades de la regió 5' del gen de la protamina humana, de brau i de ratolí, hem aconseguit amplificar el promotor corresponent de 6 espècies addicionals de mamífers (*Cercopithecus ascanius*, *Gorilla gorilla*, *Papio doguera*, *Pongo pygmaeis*, *Cavia porcellus* i *Rattus norvegicus*). La seqüència nucleotídica del promotor d'aquestes espècies demostra una marcada divergència entre la seqüència del conill porquí i la seqüència de la rata i del ratolí la qual cosa podria contribuir a clarificar la polèmica sobre si el conill porquí és o no és un rosegador (Graur et al., 1991). Entre les regions conservades destaquen la regió de la TATA-box i la regió corresponent a l'element Tet-1 proposat recentment com a responsable de l'expressió específica testicular d'aquests gens.

### Introducció.

Els gens de protamina proveeixen un dels exemples més clars de gens específicament expressats a la línia germinal (Oliva and Dixon 1991). Malgrat això no es coneix en detall els mecanismes que determinen la seva especificitat de teixit. Per tal d'esbrinar quines són les regions promotores responsables d'aquesta especificitat d'expressió hem adoptat l'estratègia de seqüenciar i comparar les regions homòlogues corresponents a diverses espècies (Queralt and Oliva 1991; Adroer et al. 1992).

D'aquesta manera és possible detectar regions conservades en l'evolució i per tant candidates a desenvolupar una funció important potencialment reguladora.

En aquest treball demostrem que oligonucleòtids dissenyats a partir de regions conservades predites a partir d'un nombre molt petit d'espècies ha permès l'amplificació i seqüenciació de la regió homòloga corresponent a moltes altres espècies. Aquests resultats aporten un nou consens de la regió promotora del gens de protamina P1 de mamífers.

### Materials i mètodes.

S'han amplificat els ADN amb les condicions de PCR següents 200 ng d'ADN genòmic, 50 mM KCl, 10mM Tris-HCl (pH 8.3), 1.5 mM MgCl<sub>2</sub> i 200µM de cada dNTP. 5 pmols de cada "primer" OP19 o OP11 i OP14 o OP112 (Queralt i Oliva, manuscrit en preparació) i una unitat de Taq I DNA polymerasa (Perkin- Elmer), una desnaturalització inicial a 94 °C durant 10 minuts, dos minuts a 94 °C, "Touchdown" de 60 °C a 45 °C durant dos minuts, tres minuts a 72 °C i 10 cicles a 45 °C. Els productes de PCR es varen



clonar en el vector Bluescript KS+ (Statagene) tallat amb Eco RV, utilitzant el mètode del "double Geneclean" (Bio 101) i *E. coli* JM 109. Es varen aïllar ssDNA i seqüenciar pel mètode de la Sequenase (USB).

Resultats.

L'amplificació de la regió promotora dels gens de protamina P1 de *Cercopithecus ascanius*, *Gorilla gorilla*, *Papio doguera*, *Pongo pygmaeis*, *Cavia porcellus* i *Rattus norvergicus* ha donat lloc a un producte únic d'amplificació del tamany esperat predit a partir de la seqüència dels gens de protamina P1 existents (Fig 1). La seqüenciació de diversos clons ha revelat la seqüència nucleotídica de la regió promotora corresponent a les diverses espècies (Fig 2).

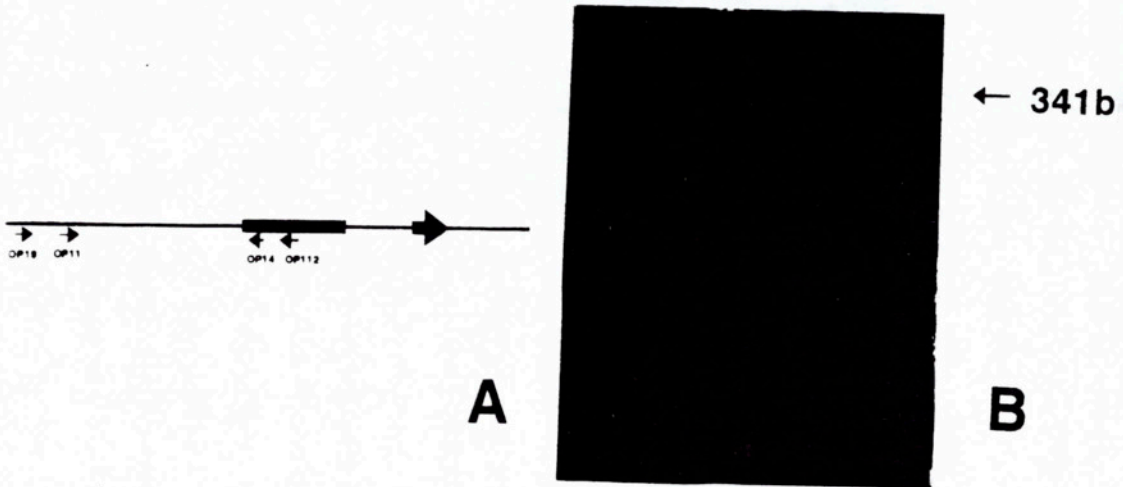
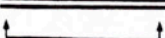


Figura 1: A: estructura del gen de protamina P1 amb un intró i el lloc d'hibridació dels oligonucleòtids. B: Producte amplificat de 341 bp 1: orangutà, 2: papió, 3-8: humans, 9: control negatiu.

PRHP1	CTTTGCCCTCACAATGACCAACGG CC--CCCTGGCATCTATAACAGGCCGCAGA
PRCER	CTTTGCCCTCACAATGACCAACGGCC--TCCTGGCATCTATAAGAACCCTGCAGA
PRGOR	CTTTGCCCTCACAATGACCAACGGCC--CCCTGGCATCTATAAGAGGCCGCAGA
PRORANGUTA	CTTAGCCCTCACAATGACCAACGGCC--CCCTGGCGTCTATAAGAGGCCGCAGA
PRPAPIO	CTTTGCCCTCACAATGACCAACAGCC--TCCTGGCATCTATAAGAGGCTGCACA
PRBP1	CTGTGACCTCACAATGACCAGGACCCTGCCCGGGTCTATATAAGAGGCCGGGAA
PRMP1	CTTGACTTCATAATTCCTAGGGG CCA---CTAGTATCTATAAGAGGAAGAGGG
PRRATA	CTTTGACTTCATAATTCAGGGG CCA---CTAGTATCTATAAGAGGAAGAGGG
PRCOBAYO	---TCACCTCACAATGTC-----CTGGGTTCTATAAGAGGCCAAGAG

  
**Tet-1**


  
**TATA**

Figura 2: Comparació de les seqüències amb el CRE i la TATA box. PRHP1: Humà, PRCER: *Cercopithecus*, PRGOR: Goril.la, PRORANGUTA: Orangutà, PRPAPIO: Papió, PRBP1: Brau, PRMP1: Ratolí, PRRATA: Rata, PRCOBAYO: Conill porquí.



## Discussió.

En aquest treball hem seqüenciat la regió promotora del gen de protamina P1 de diversos mamífers per tal de trobar regions conservades en l'evolució i per tant potencialment reguladores o determinants de l'expressió específica d'aquests gens. L'estratègia que hem seguit ha consistit en la comparació del limitat nombre de seqüències promotores de gens P1 conegudes per tal de dissenyar oligonucleòtids corresponents a les regions conservades i l'amplificació de l'ADN d'espècies addicionals fent servir aquests oligonucleòtids. Tal com es pot apreciar en la figura 1 l'amplificació ha sigut molt eficient generant-se únicament una banda del tamany esperat. L'enfoc que hem seguit no és pas trivial ja que la marcada variabilitat dels gens de protamina feia dubtar en un principi si l'estratègia seguida funcionaria o no. Així doncs disposem ara d'una nova estratègia per l'estudi dels gens de protamina en la filogènia. A més els resultats obtinguts són rellevants en l'estudi de la regulació de l'expressió d'aquests gens ja que demostren l'existència de regions conservades i per tant potencialment reguladores de l'expressió genètica (Fig. 2). Estudis previs comparant la seqüència d'un nombre limitat de gens de protamina P1 de vertebrats ja va demostrar l'existència d'algunes regions conservades a nivell del promotor (Oliva and Dixon, 1990). Una d'aquestes regions corresponia a la TATA-box mentre que l'altra demostrava similitud a l'element CRE (cAMP Response Element) però a més, immediatament després d'aquest element, es detectava el dinucleòtid CA en la majoria d'espècies. El fet que aquest dinucleòtid es trobés absent en la majoria d'elements CRE de gens constitutius va ésser un dels primers indicis de que si bé l'element CRE no era exclusiu i determinant d'expressió de gens en l'espermatogènesi, la variant que contenia l'extensió CA sí que ho podia ser. Per altra banda, ratolins transgènics obtinguts amb construccions formades pel promotor del gen de protamina de ratolí acoplats a un gen marcador demostren que les seqüències promotores contingudes en una regió d'aproximadament 500 nucleòtids 5' desde l'inici de transcripció són suficients per determinar l'especificitat d'expressió. Aquesta variant de l'element CRE estaria precisament inclosa dintre d'aquesta regió.

L'inclusió de 6 espècies addicionals de mamífers en la comparació dels promotors ha permès la determinació de les regions conservades d'una forma molt més precisa. Aquí demostrem que la regió conservada descrita prèviament amb homologia al CRE va seguida per una seqüència altament conservada (C/T A A T). La seqüència nucleotídica CA descrita prèviament correspondria a les primeres dos bases d'aquest nou consensus. Per altra banda demostrem que cap dels mamífers seqüenciats conté estrictament en el promotor de protamina P1 la seqüència canònica CRE (T G A C G T C A) i que en el seu lloc es trobaria el consensus T G C/A C T/C T C A. Tot en conjunt suggereix d'una forma molt més precisa que aquesta seqüència variant que conté el motiu addicional conservat ( T G C/A C T/C T C A C/T A A T; un 12mer) podria correspondre a un element específic que conferís expressió en testicle. Aquesta hipòtesi ve fortament recolzada per la recent descripció d'un factor específic de testicle (Tet-1) que s'uneix *in vitro* a la regió corresponent al gen de protamina P1 de ratolí (Tamura et al., 1992). La concordància entre la regió



protegida amb DNAsaI pel factor Tet-1 en el ratolí la seqüència conservada determinada mitjançant la comparació de la seqüència nucleotídica en les diverses espècies de mamífers es gairebé perfecta (Fig 2: Comparar la regió subrayada en doble línia en la regió delimitada per les dues sajetes. Es interessant notar que el nostre estudi comparatiu demostra la presència d'una T en el 3' d'aquest element altament conservada i present sense excepció en tots els mamífers estudiats. Aquesta T final no és evident a partir dels estudis de protecció amb DNAsaI possiblement perquè es tracta d'una regió que en absència de qualsevol factor és refractària a la digestió amb DNAsa I. Així doncs serà interessant determinar per altres tipus d'enzims o agents (per exemple dimetil-sulfat) el grau d'accessibilitat d'aquesta base en presència del factor Tet-1. Per altra banda, cal recordar que aquests estudis de protecció amb DNAsaI estan realitzats *in vitro* en absència de l'estructura nucleosòmica la qual cosa pot canviar d'una forma molt notable el grau d'unió d'aquests factors. Un avantatge dels estudis comparatius a nivell de seqüència d'ADN és que el seu resultat no ve limitat per les sondes moleculars existents al ésser determinada exclusivament per les característiques biològiques *in vivo*. La potencialitat de determinació d'expressió testicular d'aquest element que hem detectat és recolzada pel fet que el gen de la proteïna de transició TP1 de ratolí específicament expressada en testicle concorda també a la perfecció amb l'element Tet-1. Cal notar que la seqüència del conill porquí demostra una delecció de 10-12 nucleòtids després de la seqüència Tet-1. 10 nucleòtids correspon a una volta completa de l'hèlix en l'ADN i en el suposat que aquesta regió es trobés ensamblada formant un nucleosoma *in vivo*, el fet que la variació sigui d'una volta completa de l'hèlix de l'ADN, podria indicar la necessitat de situar aquesta seqüència accessible des de l'exterior del nucleosoma per tal de ser biològicament activa. Per altra banda, l'absència d'aquests 10-12 nucleòtids en el conill porquí correspon a una regió molt poc conservada en la resta de mamífers seqüenciats. Tot en conjunt suggereix una funció irrelevant per aquesta regió deleccionada en el conill porquí. Ara, serà interessant realitzar un estudi sistemàtic entre totes les seqüències del "Gen Bank" per buscar l'element altament conservat Tet-1, descrit en aquest article.

#### Agraïments.

Treball subvencionat per l'ajut d'inici a la Recerca de la Universitat de Barcelona, (1991) i, en part, pel projecte CICYT SAL90-0334 a R.O. S'agraeix la col.laboració de Justo Garasa, M<sup>a</sup> José Noto i resta de personal del parc zoològic de Barcelona per haver subministrat les mostres de les diferents espècies estudiades.

#### Referències.

- Adroer, R., Queralt, R., Ballabriga, J. and Oliva, R. (1992) *Nucleic Acids Res.*, 20, 609.  
Glaur et al., (1991) *Nature*, 351, 649-652.  
Oliva, R. and Dixon, G.H. (1990) *J. Mol. Evol.*, 30, 333-346.  
Oliva, R. and Dixon, G.H. (1991) *Progress in Nucleic Acid. Res. and Mol. Biol.*, 40, 25-94.  
Queralt, R. and Oliva, R. (1991) *Nucleic Acids Res.*, 19, 5786.  
Tamura et al., (1992) *J. Biol. Chem.*, 267, 4327-4332.