

# CARACTERITZACIO D'UNA FAMILIA DE SEQÜENCIES REPETIDES MOSAIC DISPERSES, ESPECIFIQUES DEL GENER ZEA.

Carlos M. Vicient, Amparo Monfort, Rosa Aledo i José Antonio Martínez-Izquierdo  
Departament de Genètica Molecular, CID (C.S.I.C.) Jordi Girona, 18. 08034 Barcelona

## RESUM

S'ha trobat una família d'elements repetits dispersos de com a mínim 14Kb presents tan sols als genomes de les espècies del gènere *Zea*, entre ells al blat de moro modern.

S'han obtingut seqüències de cinc elements corresponents a clons genòmics independents del teosinte *Zea diploperennis* (espècie silvestre del gènere *Zea*). Això ens ha permès comprovar com entre les regions equivalents existeix una semblança superior al 80%. També hem establert la seva estructura general que és de tipus mosaic, amb dos zones de repeticions tàndem entre seqüències úniques. En aquestes darreres s'han detectat diversos ORFs (marcs oberts de lectura).

Experiments d'hibridació *in situ* en cromosomes mostren que els elements repetits es distribueixen de manera dispersa per tot el genoma, essent presents a tots els cromosomes i localitzat parcialment a les zones telomèriques. Experiments de cinètiques de digestió amb l'enzim Bal31 indiquen que els elements repetits no tenen una localització telomèrica.

## INTRODUCCIO

Els genomes dels Eucariotes posseeixen un alt percentatge de seqüències repetides. Aquest percentatge és especialment alt a les plantes, sobre tot als cereals a on constitueixen més del 75% del genoma (5). Les seqüències repetides es poden classificar d'acord al seu grau de repetició en mitjanament repetides (de 100 a  $5 \times 10^4$  còpies) i altament repetides (de  $5 \times 10^4$  a  $\geq 10^6$  còpies); i en funció de la seva dispersió en agrupades en tàndem o disperses. Entre les distribuïdes en tàndem destaquen les altament repetides (satèl·lits) que solen tenir funcions estructurals (centròmers, telòmers, etc.) (13, 15). Les seqüències repetides disperses tenen funcions més variades que les anteriors: famílies multigèniques, RNAs ribosomals, de transferència o relacionats amb el *splicing* del RNA, transposons, retroelements, etc. (4, 6, 10, 13, 14). Les seqüències disperses curtes sense funcions conegudes es denominen SINES ("short interspersed repetitive elements") i les llargues LINES ("long interspersed repetitive elements") (14).

Dins d'un estudi evolutiu del gènere *Zea* i espècies relacionades, es van trobar 5 clons de *Zea diploperennis* que contenien còpies d'una nova família de seqüències repetides denominades ZEAR (8) que només es troben als genomes de les espècies del gènere *Zea*. Aquesta família fou caracteritzada i, en principi, catalogada com a SINE degut a que es creia que eren seqüències curtes (uns 250 pb) i disperses. Estudis posteriors han permès demostrar que el tamany dels elements és molt més gran ( $> 14$  Kb). Discutirem les característiques d'aquests elements i les seves possibles funcions.

## MATERIAL I METODES

Els clons genòmics lambda Charon 35 de *Zea diploperennis* s'han seqüenciat pel mètode de Sanger amb primers fluorescents mitjançant un seqüenciador automàtic A.L.F..

El nombre de còpies al genoma fou estimat per reconstrucció gènica mitjançant anàlisis per slot-blots de DNA.

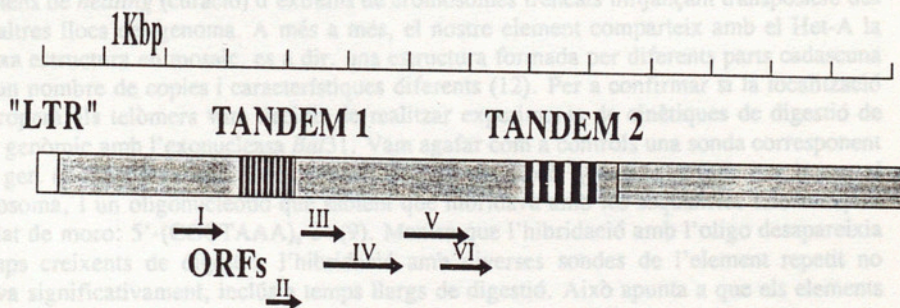
Per a les cinètiques de digestió amb *Bal31* es va fer servir DNA genòmic de *Zea diploperennis* extret pel mètode de Dellaporta (2). Les digestions es van fer amb 10 µg de DNA a concentració final de 0,2 µg/µl digerits amb 0,5 U/µg de *Bal31* a 30°C durant 0, 15, 30, 60 i 90 minuts en 12mM CaCl<sub>2</sub>, 24mM MgCl<sub>2</sub>, 0,2M NaCl, 20mM TrisHCl (pH 8,0) i 100 µg/ml de BSA. La reacció es va parar amb EGTA 20mM final.

Les hibridacions *in situ* de cromosomes metafàsics i nuclis interfàsics de *Zea mays* es van fer amb sondes marcades amb digoxigenina-11-dUTP pel mètode de nick-traslation i detectades amb anticossos anti-DIG conjugats a FITC, o a peroxidasa i amplificats amb argent.

### Distribució en el genoma

## RESULTATS

La seqüenciació completa del clon genòmic 7 (≈ 9kb), de gran part del clon 4 i de part del 6, dels quals ja es coneixien les seqüències ZEAR (també dels clons 3 i 5), ha permès establir l'estructura de l'element repetit com es mostra a la figura.



La comparació entre les zones corresponents dels diferents clons seqüenciats mostren que l'element està molt conservat. La mitjana del percentatge de semblança es superior al 80%.

### Identificació de l'extrem de l'element

S'ha establert l'extrem 5' de l'element. Això ha estat possible al seqüenciar la zona corresponent a l'extrem en dos clons, la qual cosa ens ha permès apreciar com a partir d'un nucleòtid determinat la semblança passa de ser del 34% al 92%. Les altes semblances comencen amb la seqüència 5'-TGTCGGT-3', molt semblant a la seqüència *consensus* d'iniciació dels LTR dels retrotransposons de la família *copia* (7). En la mateixa zona, just abans del començament de l'alt percentatge de semblança, existeix una seqüència amb un 63% de semblança amb el retrotransposó defectiu de blat de moro *Cin1* (11).

### Repeticions en tàndem

Dins dels elements repetits s'han detectat dos agrupacions de seqüències repetides en tàndem. El primer grup té 11 repeticions de 81pb, composades de dos subunitats de 40-41pb, molt semblants entre elles. El segon grup de repeticions en tàndem ha estat seqüenciat en dos clons diferents que presenten 5½ còpies i 7½ còpies, respectivament, essent molt alta la semblança entre els dos clons.

### Presència de ORFs

S'han identificat 6 ORFs a prop de les repeticions en tàndem, dels quals tan sols un presenta certa semblança amb proteïnes d'unió a àcids nucleics quan es comparen amb els bancs de seqüències. Hi han evidències de que aquest ORF I està present en transcrits, si bé no es coneix des de quin promotor s'inicia la transcripció.

### Distribució en el genoma

Per a confirmar la distribució dispersa de l'element s'han realitzat hibridacions *in situ* de cromosomes de blat de moro. Els resultats confirmen la dispersió però no descarten que estiguin organitzades en *clusters*.

Els resultats de les *in situ* ens indiquen una certa tendència a la localització telomèrica o subtelomèrica dels elements. Això ens va fer pensar en que el nostre element fos semblant als elements transposables subtelomèrics HeT-A de *Drosophila* (1), que intervien en fenòmens de *healing* (curació) d'extrems de cromosomes trencats mitjançant transposició des de d'altres llocs del genoma. A més a més, el nostre element comparteix amb el HeT-A la mateixa estructura en mosaic, es a dir, una estructura formada per diferents parts cadascuna amb un nombre de còpies i característiques diferents (12). Per a confirmar si la localització era propera als telòmers vam decidir de realitzar experiments de cinètiques de digestió de DNA genòmic amb l'exonucleasa *Bal31*. Vam agafar com a controls una sonda corresponent a un gen que codifica per a una HRGP de blat de moro que sabem que era intern al cromosoma, i un oligonucleòtid que sabem que hibridava amb les seqüències telomèriques del blat de moro: 5'-(CCCTAAA)<sub>n</sub>-3' (9). Mentre que l'hibridació amb l'oligo desapareixia a temps creixents de digestió, l'hibridació amb diverses sondes de l'element repetit no variava significativament, inclús a temps llargs de digestió. Això apunta a que els elements que donen hibridacions prop del final dels cromosomes a les *in situ*, no estan exactament als telòmers.

### DISCUSSIO

L'obtenció de noves seqüències dels clons genòmics que contenen les seqüències ZEAR, han confirmat els experiments previs de Southern que indicaven que els elements repetits eren llargs i que es tracta d'una família de LINES.

S'ha pogut establir un dels extrems (el 5') de l'element i presenta analogies amb els extrems LTR d'alguns retroelements. L'extrem 3' no s'ha pogut definir encara, i segons les semblances entre els clons genòmics estaria situat a més de 14 kpb de l'extrem 5'.

Algunes dades ens porten a pensar que es tractaria d'elements retrotransposables: tenen un extrem semblant a un LTR; la presència de l'element *Cin1* podria explicar-se per un fenomen de transposició de l'element gran sobre un *Cin1* ja present anteriorment; posseeixen

unes estructures en tàndem semblants a les que presenten d'altres elements retrotransponibles com el TOC1 de *Chlamydomonas reinhardtii* (3) que té unes repeticions en tàndem de 76 pb; existeixen evidències de que, al menys en part, l'element es transcriu. Tanmateix, no s'han trobat homologies significatives amb enzims típics dels retrolements transponibles com són la transcriptasa reversa i la integrasa.

Un altre possible funció d'aquests elements és estructural, com podria ser la dels elements Het-A, amb els quals tenen en comú l'estructura en mosaic. A favor d'aquesta possibilitat es troba el fet de que part de les seqüències estiguin molt metilades (8), la qual cosa ens suggereix que formen part de l'heterocromatina.

Una darrera possibilitat seria que es tractés d'una família multigènica. A favor d'això tenim la presència de les repeticions en tàndem presents en algunes famílies multigèniques i que es pensa que actuen com a llocs de recombinació i intervenen en fenòmens d'homogenització. No obstant això, no s'han trobat semblances importants entre els ORFs dels elements i gens ja coneguts, i, per altre part, el nombre de còpies (més de 2000) és molt gran per a tractarse d'una família multigènica.

Fins al moment no podem determinar a quin tipus pertanyen els elements repetits caracteritzats. Per les seves característiques estructurals sembla ser un tipus de retroelement transponible, tot i que no s'ha pogut demostrar clarament la seva transposició.

#### BIBLIOGRAFIA

- 1) Biessmann, H., Valgeirsdottir, K., Lofski, A., Chin, C., Ginther, B., Levis, R.W. i Pardue, M.L. (1991) *Mol. Cell. Biol.* **12**, 3910-1918.
- 2) Dellaporta, S.L., Wood, J. i Hicks, J.B. (1983) *Plant Mol. Biol. Rep.* **1**, 19-21.
- 3) Dai, A., Schirmer-Rahire, M., Kuchka, M.R., Maifield, S.P. i Rochaix, J.D. (1988) *EMBO J.* **7**, 1917-1927.
- 4) Döring, H.P. i Starlinger, P. (1986) *Annu. Rev. Genet.* **20**, 175-200.
- 5) Flavell, R.B. (1980) *Annu. Rev. Plant Physiol.* **31**, 569-596.
- 6) Heidecker, G. i Messing, J. (1986) *Annu. Rev. Plant Physiol.* **37**, 439-466.
- 7) Mount, S.M. i Rubin, G. (1985) *Mol. Cell. Biol.* **5**, 1630-1638.
- 8) Raz, R., Puigdomènech, P. i Martínez-Izquierdo, J.A. (1991) *Gene* **105**, 151-158.
- 9) Richards, E.J. i Ausubel, F.M. (1988) *Cell* **53**, 127-136.
- 10) Shapiro, J.A. (1983) *Mobile genetic elements*. Shapiro, J.A. (ed.), Academic Press, NY.
- 11) Shepherd, N.S., Schwarz-Sommer, Z., Blumberg vel Spalve, J., Gupta, M., Wienand, U. i Saedler, H. (1984) *Nature* **307**, 185-187.
- 12) Valgeirsdóttir, K., Traverse, K.L. i Pardue, M.L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7998-8002.
- 13) Vedel, F. i Delseny, M. (1987) *Plant Physiol. Biochem.* **74**, 247-251.
- 14) Weiner, A.M., Deiningen, P.L. i Efstratiadis, A. (1986) *Annu. Rev. Biochem.* **55**, 631-661.
- 15) Zakian, V.A. (1989) *Annu. Rev. Genet.* **23**, 579-604.