

GENE 05096

## A new family of repetitive nucleotide sequences is restricted to the genus *Zea*

(Interspersed repeats; transcribed repeats; teosinte; maize; corn; Gramineae; methylation; evolution; gene conversion)

Regina Raz, Pere Puigdomènech and José A. Martínez-Izquierdo

Departamento de Genética Molecular, CID-CSIC, Jordi Girona Salgado 18-26, 08034 Barcelona (Spain)

Received by G. Bernardi: 22 November 1990

Revised/Accepted: 29 March/12 May 1991

Received at publishers: 13 June 1991

---

### SUMMARY

We have isolated a new family of moderately repetitive nucleotide sequences (about 2500 copies per haploid genome) specific to the genus *Zea* and absent in other graminaceous species. These sequences are interspersed in the genome and they show the same genomic organization pattern and similar copy number in all the *Zea* species examined. These two facts, consistency in the copy number and the same organization pattern, would indicate on the one hand that these sequences were amplified before the divergence of *Zea* species, and on the other hand that maize and all the teosintes could be considered as the same evolutionary population. Independent clones corresponding to the repetitive sequences have been isolated and sequenced from a genomic library of the teosinte, *Zea diploperennis*. The repeats, flanked by *Hae*III sites, are more than 70% G + C-rich, on average 253 bp long and show 78% similarity to each other. These repetitive sequences are in a highly methylated-C context and they present some features resembling those of coding sequences, such as high CpG and low TpA content, and similar codon usage to maize genes in one of the reading frames. Moreover, the repetitive probe hybridizes with RNA extracted from different tissues of maize and from teosinte, indicating that these repeats or similar ones are present in transcribed sequences.

---

### INTRODUCTION

The genus *Zea* is a good system to study the variation rate among nt sequences, due to the close evolutionary relationships among its members (Walbot, 1983; Doebley et al., 1987). Variations in nt sequences during relatively

short evolutionary time scales occur at a higher rate in repetitive DNA (Hinegardner, 1976) than variations in coding regions that diverge more slowly. An intermediate situation is represented by the coding sequences of storage protein genes where selective pressures may not be so strong as for important structural or regulatory genes (Kreis and Shewry, 1989). Maize and other cereal genomes contain more than 70% of repeated sequences (Flavell et al., 1974; Bennett and Smith, 1976).

The nt sequences can be classified into three broad groups according to the frequency of reiteration: unique (one or a few copies per haploid genome), moderately repetitive (thousands to hundreds of thousands copies), and highly repetitive (millions of copies) (Britten and Kohne, 1968; Vedel and Delseny, 1987). Another classification refers to the organization of sequences in the genome: tandemly repeated sequences or sequences interspersed

---

Correspondence to: Dr. J.A. Martínez-Izquierdo, Departamento de Genética Molecular, CID-CSIC, Jordi Girona Salgado, 18-26, 08034 Barcelona (Spain)

Tel. (34-3)2046851; Fax (34-3)2045904.

Abbreviations: A, adenosine; bp, base pair(s); C, cytidine; cDNA, DNA complementary to RNA; G, guanosine; kb, kilobase(s) or 1000 bp; n, haploid number of chromosomes; nt, nucleotide(s); ORF, open reading frame; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na<sub>3</sub>·citrate pH 7.6; SSPE, 0.18 M NaCl/0.001 M Na<sub>2</sub>EDTA/0.01 M Na·phosphate pH 7.7; T, thymidine; *Z.*, *Zea*; *ZEAR*, *Zea* repeats.

throughout the genome. The latter seems to be the case for the majority of maize repetitive sequences which are mainly organized in a short interspersion pattern, where repeated sequences of 100–300 bp in length are interspersed with unique or other repetitive sequences (Hake and Walbot, 1980). The function and the origin of most repetitive sequences remain unknown, some of them might play a regulatory role in the control of gene expression or in recombination, while others might be involved in the structural organization of the genome. Some of the repetitive sequences, especially those with long-period and short-period interspersion patterns, are either transcriptional units for discretely sized RNAs themselves and/or they are extensively homologous to discretely sized RNAs that are transcribed elsewhere (Jelinek and Schmid, 1982).

Here, we report on a moderately repeated family of nt sequences interspersed in the genome of the teosinte *Z. diploperennis*, one of the putative ancestors of cultivated maize (Iltis et al., 1979; Mangelsdorf, 1986). These sequences are present in the DNA of all *Zea* spp. examined, with the same genomic organization and similar copy number. This fact will be discussed in relation to the putative mechanisms that have operated in the evolution of these sequences and the phyletic relationships among *Zea* species. These plant repetitive sequences are restricted to one genus and the first for which sequence information is available.

## RESULTS AND DISCUSSION

### (a) Isolation and sequence analysis of repeated *Hae*III sequences

A genomic library of *Z. diploperennis* DNA was made using the  $\lambda$ Charon35 vector. The library was screened with a cDNA probe corresponding to the maize storage protein glutelin-2 (Prat et al., 1985) and is 66% G + C-rich. A number of clones showed a 250-bp fragment hybridizing with the glutelin-2 cDNA probe in the *Hae*III digest. The sequences of *Hae*III fragments corresponding to five independent clones are shown in Fig. 1. These elements, referred to as *ZEAR* elements, have an average length of 253 bp, ranging from 231–270 bp. They show, on average, 78% similarity, the highest being between *ZEAR* 270 and *ZEAR* 266 (86%) and the lowest between *ZEAR* 260 and *ZEAR* 231 (70%). Differences have the character of deletions/insertions and nt substitutions, with 80% of the substitutions being transitions. They are not scattered randomly; the first 65 bp of the sequences are identical except for one substitution in *ZEAR* 266.

*ZEAR* sequences are very G + C-rich: 74% on average, compared to 49.5% for the maize genome as a whole (Hake and Walbot, 1980). Moreover, these sequences show a

pronounced strand asymmetry, with 47% G vs. 27% C and 20% A vs. 6% T, i.e., 67% purines. Strand asymmetry has also been reported for coding sequences of flax 5S DNA (Goldsbrough et al., 1982) and maize autonomously replicating sequences (Berlani et al., 1988), although no similarity can be found between these and *ZEAR* members. This unusual nt distribution is also reflected in the content of certain dinucleotides: the % of CpG and TpA are, on average, respectively, 22% higher and 39% lower than expected (corrected for nt frequencies). High CpG and low TpA content has been reported as typical of either DNA coding sequences or those that are transcribed but not translated (Beutler et al., 1989).

When the sequences are organized in the reading frame shown in Fig. 1 the codon distribution fits well with the codon usage described for maize genes (Murray et al., 1989). Interestingly, gaps introduced by the computer programme to increase the similarity of the sequences are multiples of three. However, the number of substitutions is roughly the same at the three positions of the putative codons depicted in Fig. 1, in contrast to coding sequences where the number of substitutions is higher at the third position of the codons (Li et al., 1985).

A data bank search revealed no significant similarities between *ZEAR* members and a number of repetitive sequences, as for example CAT repetitive sequences from yeast (Widelman et al., 1986). Although the observed sequence similarities most likely do not have any functional meaning, it is worthwhile to point out that the overwhelming majority of them was found with the last region of *ZEAR* molecules (data not shown).

The *ZEAR* elements are themselves very repetitive, having 19 perfect direct repeats ranging from 10 to 16 nt (for *ZEAR* 270), with only one present in the first 65 bp (results not shown). This fact, together with the conservation of the 5' region of *ZEAR* sequences and the absence of any similarities between that region and data bank sequences, suggests a different evolutionary pathway for the two regions of *ZEAR* sequences. This could be explained if these regions had evolved at different rates, and the conservation of the sequence from nt 1–65 might then reflect selection against mutation in it.

### (b) Genomic organization

Genomic organization of the *ZEAR* family was analyzed by Southern hybridization of different teosinte genomic digests probed with *ZEAR* 270 (Fig. 2A). Enzymes recognizing 6 bp give bands from approx. 20–1.8 kb with background hybridization, while *Hae*III, recognizing 4 bp, gives discrete bands (within a smear) ranging from 1600–250 bp. The latter band represents approx. 78% of the DNA hybridizing to the *ZEAR* probe and it has the same length as the *ZEAR* elements sequenced. This pattern



arrowheads in Fig. 2B) increase in intensity at the same time, suggesting that all these fragments are independently liberated. We conclude that the *ZEAR* repeat family is dispersed in the genome with some clustering. It is also possible that such a pattern of hybridization corresponds to larger units of repetition, with all or some of them containing *ZEAR* repeats (Evans et al., 1983; Gupta et al., 1984).

Estimation of copy number of *ZEAR* repeats by genomic reconstruction was done by slot blot analysis of total teosinte DNA, using *ZEAR* 270 element as a probe. The copy number was estimated to be 2500 per haploid genome of teosinte (results not shown) assuming 2.64 pg to be the haploid genome content of *Z. diploperennis* (Laurie and Bennett, 1985). The copy number may be underestimated by 2.2 to 10.5% (depending on the tissue) due to the contribution of chloroplast and mitochondrial DNA to the total amount of genomic DNA assayed, but corresponds to a moderately repetitive family of sequences (Britten and Kohne, 1968; Vedel and Delseny, 1987).

### (c) Methylation of *ZEAR* sequences

The high G + C content of *ZEAR* repeats, the fact that repetitive sequences seem to be prone to methylation (Selker, 1990), and the predominance of transition mutations among these repeats, prompted us to investigate the methylation status of this family of repetitive sequences. The nuclear DNA of higher plants is highly methylated at C residues, for example, 26.7–29.3% of maize DNA C's are methylated (Shapiro, 1976). The methylation of C residues in plants occurs at C-X-G trinucleotides as well as C-G dinucleotides (Gruenbaum et al., 1981). These nt arrangements are present in the recognition sequences of two pairs of restriction endonuclease isoschizomers: *HpaII/MspI* and *EcoRII/BstNI*, recognizing CCGG and CCAGG or CCTGG, respectively. *HpaII* and *EcoRII* do not cut DNA when the internal C is methylated, whereas *MspI* and *BstNI* do so in the same situation (McClelland and Nelson, 1985).

When genomic DNA of *Z. diploperennis* was digested with these four enzymes and probed with *ZEAR* 270, bands of low  $M_r$ , ranging from 2000–300 bp, were observed only in the case of *BstNI*, although slightly more digestion and background hybridization can be seen with *MspI* compared to *HpaII* (Fig. 3). The majority of DNA digested with *HpaII*, *MspI* and *EcoRII* remains as high  $M_r$ , but this is not due to incomplete digestion, because phage  $\lambda$  DNA used as an internal control was completely digested and bands of chloroplast DNA can be seen (results not shown). The inability of *MspI* vs. *BstNI* to digest teosinte DNA and to liberate *ZEAR* repeats can be explained by the high degree of methylation at external C's in the sequence CCGG. *MspI* does not cut DNA when these C's are methylated, whereas *BstNI* is insensitive to any methylation (McClelland and Nelson, 1985). These results indicate that teosinte DNA is

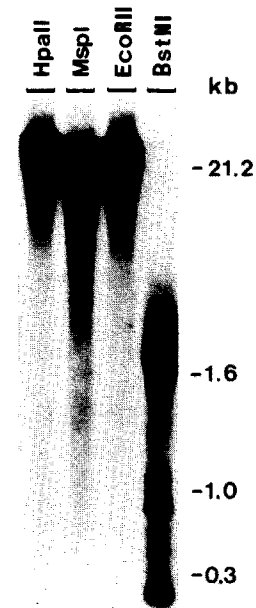


Fig. 3. Methylation context of *ZEAR* elements. *Z. diploperennis* genomic DNA was digested to completion with the isoschizomers *HpaII/MspI*, and *EcoRII/BstNI* (as was shown by  $\lambda$  DNA used as an internal control), electrophoresed in a 1.2% agarose gel and probed with *ZEAR* 270. Positions and sizes (in kb) of  $M_r$  markers are indicated. Conditions of blotting, hybridization, washing and exposure of the filters were as indicated in Fig. 2.

highly methylated at C residues recognized by methyl-sensitive endonucleases, and also that *ZEAR* members are in a highly methylated context.

We assume that many of the C's in *ZEAR* repeats have been methylated and considering that 17 out of 19 C  $\leftrightarrow$  T substitutions are followed by G's (and the other two by AG's), it is tempting to attribute these transitions to deamination of 5-methyl-C.

### (d) Species distribution

To determine the species distribution and the relative abundance of the *ZEAR* family, total DNA of maize, several other cereals and two dicotyledonous plants were compared to teosinte DNA for hybridization with *ZEAR* elements by Southern-blot analysis (Fig. 4A). Rice, wheat and barley DNAs gave rise to a smear and some bands, but in no case did the intensity and pattern of bands correspond with those of teosinte DNA. There was no detectable hybridization with sorghum, rapeseed and tobacco DNAs. Interestingly, maize and teosinte DNA gave exactly the same pattern of bands of approximately equal intensity, suggesting that *ZEAR* elements are specific to the genus *Zea* and that the structure is conserved. To confirm this hypothesis, DNAs from different species of the genus *Zea* and related taxa were digested with *HaeIII* and subjected to Southern hybridization analysis using *ZEAR* 270 as a probe. These species included several teosintes (*Z. mays*

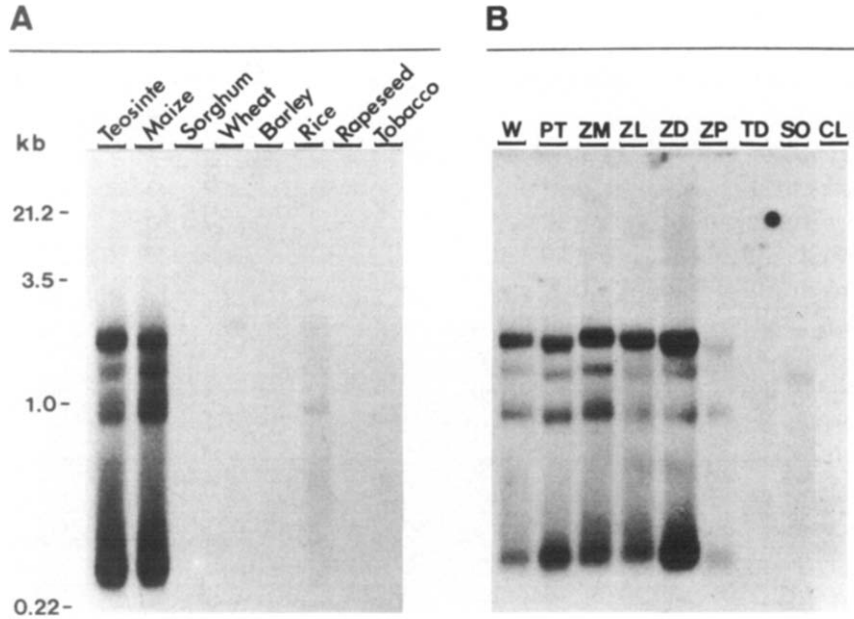


Fig. 4. Genomic Southern blots of different monocot and dicot DNAs hybridized against the *ZEAR* 270 element. Digestion (10 µg of DNA per lane) was with *Hae*III. Sizes (in kb) of *M<sub>r</sub>* markers and their position on the gel are indicated. (A) Southern corresponding to different gramineous and dicot species. (B) Southern corresponding to maize and species closely related with it: *Z. mays* ssp. *mays* inbred line W64 A (W) and Palomero Toluqueño race (PT), *Z. mays* ssp. *mexicana* (ZM), *Z. luxurians* (ZL), *Z. diploperennis* (ZD), *Z. perennis* (ZP), *Tripsacum dactyloides* (TD), *Saccharum officinale* (SO) and *Coix lacryma-jobi* (CL). Conditions of blotting, hybridization, washing and exposure of the filters were as indicated in Fig. 2.

spp. *mexicana*, *Z. luxurians*, *Z. diploperennis* and *Z. perennis*, *Z. mays* ssp. *mays* (race Palomero Toluqueño and inbred line W64A), *Tripsacum dactyloides*, *Saccharum officinale* and *Coix lacryma-jobi* (Fig. 5). The results are presented in Fig. 4B. All teosinte and maize spp. showed the same pattern of bands and all but *Z. perennis* similar band intensities, whereas *Tripsacum*, *Saccharum* and *Coix* DNAs

did not show significant hybridization with the *ZEAR* probe (Fig. 4B). This indicates that the *ZEAR* family is restricted to the genus *Zea*. Although *ZEAR* sequences from maize and other teosintes are not available, we presume that they are quite similar to those of *Z. diploperennis*, on the basis of Southern experiments performed with high-stringency hybridization and washing conditions. With one

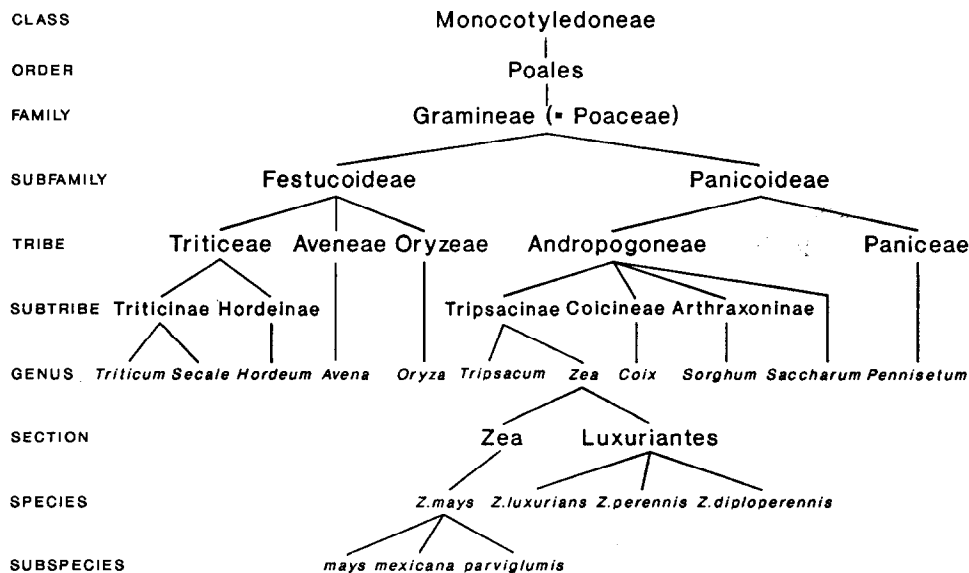


Fig. 5. Taxonomic tree of Gramineae and *Zea*. The tree has been made on the basis of taxonomic data reported by other authors (Doebley and Iltis, 1980; Bietz, 1982; Clayton, 1983, Mangelsdorf, 1986).

exception (*Z. perennis*), the variation in copy number is as much as 25% compared with *Z. diploperennis*, taking into account the difference in haploid genome DNA content (Laurie and Bennett, 1985) and the difference in intensity of bands stained with ethidium bromide. In *Z. perennis* there are 70% fewer copies than in *Z. diploperennis*.

The consistency of the organization and the constancy in copy number of *ZEAR* repeats within the *Zea* spp. indicates that these sequences were amplified before the divergence of *Zea* spp. and then selection against variation in the organization and copy number must have occurred. Alternatively, concerted evolution by homogenization mechanisms such as gene conversion may have occurred (Evans et al., 1983; Ganai et al., 1988). Gene conversion has been shown to occur between dispersed repeats within a chromosome, between homologous and nonhomologous chromosomes (Li et al., 1985 and references therein). As maize ( $2n = 20$ ) gives  $F_1$  fertile hybrids with all the teosintes having  $2n = 20$  (Wilkes, 1977; Mangelsdorf, 1986), the interchange of genetic information during meiosis between 'homologous' chromosomes of two different *Zea* spp. is probable. However, when maize is crossed with *Z. perennis* ( $2n = 40$ ), it gives  $3n F_1$  hybrids with a tendency to sterility (Mangelsdorf, 1986). In this case, the interchange of genetic material between 'homologous' chromosomes would be more difficult. As gene conversion causes no change in gene number (Li et al., 1985), it is reasonable to envisage that the conservation of the genomic organization and the consistency in the copy number of *ZEAR* elements among *Zea* spp. have taken place by gene conversion. On these bases *Zea* spp. should be considered as one evolutionary population and not only maize and Mexican teosintes (as *Z. mays* spp. *mexicana* and *Z. mays* ssp. *parviglumis* races) as has been suggested by Dennis and Peacock, 1984. The *ZEAR* family of repetitive sequences must be regarded as a differential trait of the genus *Zea* and may be useful for the analysis of somatic hybrids and the progeny of plants involving *Zea* species, or in taxonomic studies.

#### (e) Transcription of *ZEAR* elements

To assess whether the *ZEAR* repetitive sequences are present in transcribed RNAs, as has been shown for other dispersed repetitive sequences (review in Jelinek and Schmid, 1982), Northern experiments were carried out with total and poly(A)<sup>+</sup> RNA from several tissues of maize and total leaf RNA from *Z. perennis* and *Z. diploperennis* (Fig. 6). The *ZEAR* 270 element hybridized in all the cases with a discretely sized RNA of approx. 820 nt (Fig. 6). This RNA band hybridizing with the *ZEAR* probe is also present in the poly(A)<sup>+</sup> RNA fraction (Fig. 6, lane C+). These results can be explained by considering that *ZEAR* elements or related sequences are transcribed as a part of longer nt sequences (containing these elements), as it has been

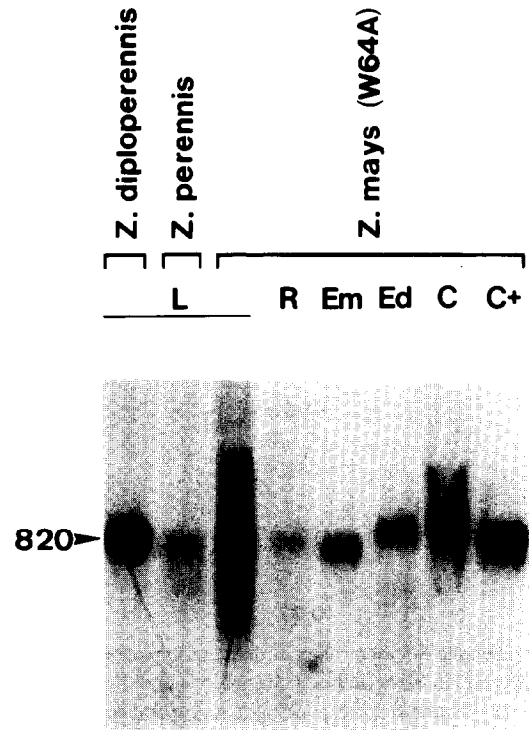


Fig. 6. Northern analysis of RNA extracted from several tissues of maize and from teosintes, probed with *ZEAR* 270. Total RNAs (10 µg/lane) were isolated from leaf (L), root (R), embryo (Em), endosperm (Ed) and coleoptile (C) (Logemann et al., 1987). Poly(A)<sup>+</sup>-enriched RNA from maize coleoptile (C+) (0.5 µg/lane) was purified by oligo(dT)-cellulose chromatography as described in Prat et al. (1985). RNAs were fractionated in 1.5% agarose/2.2 M formaldehyde gels (Lehrach et al., 1977) and blotted onto nylon (Hybond-N, Amersham, U.K.), according to the protocol recommended by the supplier. The probe was prepared as described in Fig. 2. The filter was prehybridized in 2.5 × SSPE/50% formamide/5 × Denhardt/0.5% SDS/0.25 mg of denatured salmon sperm DNA per ml at 42°C for 4 h. Hybridization was carried out under the same conditions for 18 h. Final washes were in 0.1 × SSC/0.1% SDS at 65°C. Exposure was with intensifying screen at -80°C for three days.

reported for *Alu* sequences (Jelinek and Schmid, 1982). Differences in band intensity, depending on tissues, might be explained by a distinct level of tissue expression (Jelinek and Schmid, 1982; Lewin, 1982).

The evidence that *ZEAR* or *ZEAR*-like sequences are present in RNA transcripts suggests that the generation of *ZEAR* members in the *Zea* genome could have taken place through an RNA intermediate, as it has been postulated for other interspersed transcribed families of repetitive sequences (Rogers, 1985; Weiner et al., 1986). The possibility that *ZEAR* elements are part of ORFs and that these elements are translated remains to be investigated.

#### (f) Conclusions

(1) We report here on the molecular characterization of a moderately repetitive family of dispersed nt sequences from *Z. diploperennis* which is present in all the *Zea* spp. but

not in related taxa. These *Zea* spp. share the same organization pattern and consistency in copy number (except for *Z. perennis*) of *ZEAR* repeats. Therefore, *Zea* spp. should be considered as one evolutionary population.

(2) *ZEAR* members can be regarded as a composite of two sequences. The first part (65 bp) is almost invariable, with few internal repetitions and no similarities with foreign sequences. The second part of the sequence is much more variable, with many internal direct repeats and with a certain degree of similarity with other data bank sequences. We suggest that the appearance of *ZEAR* sequences occurred simultaneously with that of the genus *Zea* and was followed by their amplification.

(3) Sequence comparisons reveal the predominance of transition vs. transversion mutations among single mutational events. The high methylation context of *ZEAR* and flanking sequences and the fact that all of the C↔T transition mutations occur in the C-methylation sequences observed in plants, suggest that these transitions may have resulted from the deamination of 5-methyl-C giving rise to T, during the evolution of *Zea* spp.

(4) The sequences of the five family members studied have interesting features themselves, namely: very high G + C content, pronounced strand asymmetry, high CpG and low TpA dinucleotide content and codon usage in one of the reading frames similar to that observed in maize genes. However, the frequencies of nt substitutions (roughly similar for the three codon positions) do not correspond to those usually observed in regions coding for proteins.

(5) *ZEAR* or *ZEAR*-like sequences are present in RNA transcripts from several tissues of maize and from *Z. diploperennis*.

#### ACKNOWLEDGEMENTS

The authors are indebted to Mr. Patricio Gómez for his technical assistance and to Dr. Shirley Burgess for carefully revising this manuscript. We thank CIMMYT (Mexico) and Angel Alvarez, Aula Dei (Zaragoza, Spain) for supplying the seeds. This work has been supported by a grant from PLANICYT (BIO 88/0242). R.R. is recipient of a fellowship from Programa de Formación del Personal Investigador.

#### REFERENCES

Bennett, M.D. and Smith, J.B.: Nuclear DNA amounts in angiosperms. *Phil. Trans. Roy. Soc. Ser. B* 274 (1976) 227–274.  
 Berlani, R.E., Walbot, V. and Davis, R.W.: Sequence analysis of three fragments of maize nuclear DNA which replicate autonomously in yeast. *Plant Mol. Biol.* 11 (1988) 173–182.

Beutler, E., Gelbart, T., Han, J., Koziol, J.A. and Beutler, B.: Evolution of the genome and the genetic code: selection at the dinucleotide level by methylation and polyribonucleotide cleavage. *Proc. Natl. Acad. Sci. USA* 86 (1989) 192–196.  
 Bietz, J.A.: Cereal prolamin evolution and homology revealed by sequence analysis. *Biochem. Genet.* 20 (1982) 1039–1053.  
 Biggin, M.D., Gibson, T.J. and Hong, G.F.: Buffer gradient gels and <sup>35</sup>S label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* 80 (1983) 3963–3965.  
 Britten, R.J. and Kohne, D.E.: Repeated sequences in DNA. *Science* 161 (1968) 529–540.  
 Burr, B. and Burr, F.A.: Controlling-element events at the shrunken locus in maize. *Genetics* 98 (1981) 143–156.  
 Clayton, W.D.: Notes on the tribe Andropogoneae (Gramineae). *Kew Bull.* 35 (1983) 813–818.  
 Dennis, E.S. and Peacock, W.J.: Knob heterochromatin homology in maize and its relatives. *J. Mol. Evol.* 20 (1984) 341–350.  
 Doebley, J. and Iltis, H.H.: Taxonomy of *Zea* I. Subgeneric classification with key to taxa. *Am. J. Bot.* 67 (1980) 982–993.  
 Doebley, J., Renfro, W. and Blanton, A.: Restriction site variation in the *Zea* chloroplast genome. *Genetics* 117 (1987) 139–147.  
 Evans, I.J., James, A.M. and Barnes, S.R.: Organization and evolution of repeated DNA sequences in closely related plant genomes. *J. Mol. Biol.* 170 (1983) 803–826.  
 Flavell, R.B., Bennett, M.D., Smith, J.B. and Smith, D.B.: Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* 12 (1974) 257–268.  
 Ganai, M.W., Lapitan, N.L.V. and Tanksley, S.D.: A molecular and cytogenetic survey of major repeated DNA sequences in tomato (*Lycopersicon esculentum*). *Mol. Gen. Genet.* 213 (1988) 262–268.  
 Goldsbrough, P.B., Ellis, T.H.N. and Lomonosoff, G.P.: Sequence variation and methylation of the flax 5S RNA genes. *Nucleic Acids Res.* 10 (1982) 4501–4514.  
 Gruenbaum, Y., Naveh-Many, T., Cedar, H. and Razin, A.: Sequence specificity of methylation in higher plant DNA. *Nature* 292 (1981) 860–862.  
 Gupta, M., Shepherd, N.S., Bertram, I. and Saedler, H.: Repetitive sequences and their organization on genomic clones of *Zea mays*. *EMBO J.* 3 (1984) 133–139.  
 Hake, S. and Walbot, V.: The genome of *Zea mays*, its organization and homology to related grasses. *Chromosoma* 79 (1980) 251–270.  
 Hinegardner, R.: Evolution of genome size. In: Ayala, F.J. (Ed.), *Molecular Evolution*. Sinauer, Sunderland, MA, 1976, pp. 179–199.  
 Iltis, H.H., Doebley, J.F., Guzmán, M.R. and Pazy, B.: *Zea diploperennis* (Gramineae): a new teosinte from Mexico. *Science* 203 (1979) 186–188.  
 Jelinek, W.R. and Schmid, C.W.: Repetitive sequences in eukaryotic DNA and their expression. *Annu. Rev. Biochem.* 51 (1982) 813–844.  
 Kreis, M. and Shewry, P.R.: Unusual features of cereal seed protein structure and evolution. *BioEssays* 10 (1989) 201–207.  
 Laurie, D.A. and Bennett, M.D.: Nuclear DNA content in the genera *Zea* and *Sorghum*. Intergeneric, interspecific and intraspecific variation. *Heredity* 55 (1985) 307–313.  
 Lehrach, W., Diamond, D., Wozney, J. and Boedtker, H.: RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* 21 (1977) 4743–4751.  
 Lewin, R.: Repeated DNA still in search of a function. *Science* 217 (1982) 621–623.  
 Li, W.-H., Luo, C.-C. and Wu, C.-I.: Evolution of DNA sequences. In: MacIntyre, R.J. (Ed.), *Molecular Evolutionary Genetics*. Plenum, New York, 1985, pp. 1–94.

- Loenen, W.A.M. and Blattner, F.R.: Lambda Charon vectors (Ch32, 33, 34 and 35) adapted for DNA cloning in recombination-deficient hosts. *Gene* 26 (1983) 171-179.
- Logemann, J., Schell, J. and Willmitzer, L.: Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* (1987) 16-20.
- Mangelsdorf, P.C.: The origin of corn. *Sci. Am.* 225 (1986) 80-86.
- Maniatis, T., Fritsch, E.F. and Sambrook, J.: *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.
- McClelland, M. and Nelson, M.: The effect of site specific methylation on restriction endonuclease digestion. *Nucleic Acids Res.* 13 Suppl. (1985) r201-r207.
- Murray, E.E., Lotzer, J. and Eberle, M.: Codon usage in plant genes. *Nucleic Acids Res.* 17 (1989) 477-498.
- Prat, S., Cortadas, J., Puigdomènech, P. and Palau, J.: Nucleic acid (cDNA) and amino acid sequences of the maize endosperm protein glutelin-2. *Nucleic Acids Res.* 13 (1985) 1493-1504.
- Queen, C. and Korn, C.J.: A comprehensive sequence analysis program for the IBM personal computer. *Nucleic Acids Res.* 12 (1984) 581-599.
- Rogers, J.H.: Long interspersed sequences in mammalian DNA. Properties of newly identified specimens. *Biochim. Biophys. Acta* 824 (1985) 113-120.
- Selker, E.V.: DNA methylation and chromatin structure: a view from below. *Trends Biochem. Sci.* 15 (1990) 103-107.
- Shapiro, H.S.: Distribution of purines and pyrimidines in deoxyribonucleic acids. In: Fasman, G.D. (Ed.), *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, Vol. II. CRC Press, Cleveland, OH, 1976, pp. 241-281.
- Vedel, F. and Delseny, M.: Repetitiveness and variability of higher plant genomes. *Plant Physiol. Biochem.* 25 (1987) 191-210.
- Walbot, V.: Morphological and genomic variation in plants: *Zea mays* and its relatives. In: Goodwin, B.C., Holder, N. and Wylie, C. (Eds.), *Development and Evolution.* Cambridge University Press, Cambridge, U.K., 1983, pp. 257-277.
- Weiner, A.M., Deininger, P.L. and Efstratiadis, A.: Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* 55 (1986) 631-661.
- Wilkes, H.G.: Hybridization of maize and teosinte, in Mexico and Guatemala and the improvement of maize. *Econ. Bot.* 31 (1977) 254-293.