Theor Appl Genet (1992) 84:555–559

Subfamilies of histone H3 and H4 genes are located on most, possibly all of the chromosomes in maize

N. Chaubet [1], G. Philipps [1], C. Gigot [1], C. Guitton [2], N. Bouvet [2], G. Freyssinet [2], M. Schneerman [3], and D. F. Weber [3]

[1] Institute of the Molecular Biology of Plants, CNRS, Université Louis Pasteur, 12, rue du Général Zimmer 67084 Strasbourg Cédex, France
[2] Rhône-Poulenc Agrochimie, BP 9163, 69263 Lyon Cédex 09, France
[3] Genetics group, Department of Biological Sciences, Illinois State University, Normal, IL 61761, USA

Received November 11, 1991; Accepted January 10, 1992
Communicated by J. Beckmann

Summary. It has been previously shown that in the genome of maize the multiple copies of the histone H3 and H4 multigenic families are organized into eight to ten subfamilies each containing a variable number of copies. Each subfamily is characterized by a specific proximal environment and thus can be revealed by blot-hybridization with its specific 5′ probe. Restriction fragment length polymorphism (RFLP) combined with monosomic analysis was used to localize several H3 and H4 subfamilies on maize chromosomes. H3 and H4 genes were found to be located on most, possibly all of the chromosomes, revealing a remarkably dispersed organization of these multigenic families.

Key words: Maize – H3-H4 multigenic families – Chromosomic localization – Monosomics

Introduction

The organization of the multigenic families encoding histones has been well documented in the genomes of lower eukaryotes and animals (Kedes 1979; Hentschel and Birnstiel 1981; Maxson et al. 1983). In a majority of these organisms these multigenic families are organized into gene clusters that show a remarkable variability in their topological arrangement from one species to another. Tandemly repeated sets of genes encoding the five major histones exist in the genome of sea urchins (Hentschel and Birnstiel 1981), fruit flies (Lifton et al. 1977), and newts (Stephenson et al. 1981). Histone genes have also been found to be clustered in birds and mammals, but from one cluster to another, the histone coding regions vary in identity, order, and spacing (Maxson et al. 1983). Nevertheless, in many organisms H2A and H2B genes on one hand and H3 and H4 genes on the other hand have been found to be paired. To add to such a topological variability, it has also been shown that both ordered and disordered gene arrangements may coexist in the same genome, as in the case of sea urchin (Childs et al. 1982; Maxson et al. 1983).

In the plant kingdom only the genes encoding the most conserved histones (i.e., H3 and H4) have been cloned from a small number of species (review by Gigot 1988), and data concerning their genomic organization are restricted to maize. The various cloned maize H3 and H4 genes encode the same proteins and show extensive sequence homology (95–98%) among their coding regions. In contrast, most of them have specific 5′ and 3′ flanking regions (Chaubet et al. 1989). When genomic blots of restricted maize DNA were hybridized to the H3 or H4 coding regions used as probes, 10–12 DNA fragments of different sizes hybridized at variable intensities, but whatever the restriction digest no DNA fragment hybridized to both H3 and H4 probes (Chaubet et al. 1986). This result suggests that these two types of genes are not closely linked, which is in contrast to the situation in animal genomes. When the 5′ flanking non-transcribed region of one particular gene was used as a probe, only one (at an exception, two) fragment hybridized. Using the 5′ non-coding regions of three H3 and two H4 genes as probes, we could demonstrate that in maize the multiple copies of both the H3 and H4 multigenic families are organized into eight to ten subfamilies (Chaubet et al. 1989). Each subfamily contains a different number of gene copies that all have the same proximal environment. These subfamilies were found to exist in more than 40 maize inbred lines as well as in the closely related ge-

Correspondence to: C. Gigot

nomes of sorghum, *Coïx*, teosinte, and sugar cane (Chaubet et al. 1989).

In this paper, we present an analysis of restriction fragment length polymorphisms on maize monosomics using 5′ non-transcribed regions of different H3 and H4 genes as well as H3 and H4 coding regions as probes. We localize three previously characterized H3 and H4 subfamilies on three different maize chromosomes and demonstrate that additional H3 and H4 genes are located on most, perhaps even all of the maize chromosomes.

Materials and methods

Monosomic plants

Monosomic maize plants analyzed in this study were produced using the *r-X1* system as described (Weber 1973; Helentjaris et al. 1986). Inbred W22 maize containing the *r-X1* deficiency (*R/r-X1*) were crossed as female parents by inbred Mangelsdorf's multiple chromosome tester (MT), and F_1 progeny which were monosomic for chromosomes 2, 3, 4, 6, 7, 8, 9, or 10 (eight of the ten chromosomes) or diploid were selected from the progeny of this cross. Monosomics generated with the *r-X1* system contain a complete haploid genome from the male parent (MT) and nine of the ten chromosomes from the female (W22) parent. Each of the monosomics analyzed in this study was examined cytologically or by RFLP analysis to establish that a complete chromosome was missing

DNA preparation and analysis

DNA was extracted from leaf tissues, restriction digested, separated on agarose gels, transferred onto nylon membranes as previously described and blot hybridized with the 5′ non-transcribed regions of different H3 (H3C2, C3 and C4) and H4 (H4C7 and C14) genes previously used to characterize the different multigenic subfamilies (Chaubet et al. 1989).

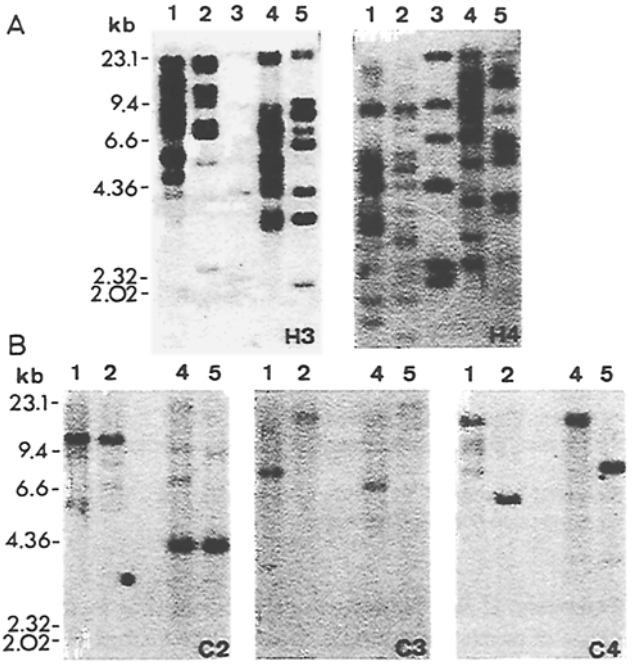

Fig. 1 A, B. Restriction fragment length polymorphism of the histone *H3* and *H4* genes (**A**) and 3 of their subfamilis (*H3C2, C3* and *C4*) (**B**) in the genomes of maize lines W22 and Mangelsdorf's multiple chromosome tester (MT). Genomic DNA from W22 (*lanes 1 and 4*) and MT (*lanes 2 and 5*) were digested with *Bgl*II (*lanes 1 and 2*) and *Xba*I (*lanes 4 and 5*), electrophoresed on 0.8% agarose gels and the fragments transferred to Hybond N nylon membrane. The blot was successively hybridized to the 5′-probes of the three H3 gene subfamilies C2, C3 and C4 (**B**), and finally to the H3 and H4 coding regions (**A**). *Lane 3* corresponds to size references whose lengths are indicated on the *left* of the figure

Results

Localization of histone H3 and H4 subfamilies using the 5′ non-transcribed regions as probes

When DNA purified from the two parental lines W22 and Mangelsdorf's multiple tester (MT) used to generate the eight maize monosomic types were restricted and blot hybridized to the maize H3 or H4 coding regions, the hybridization patterns showed 10–12 bands hybridizing to each probe (Fig. 1 A). The same blots were successively hybridized to the 5′ flanking regions specific to three H3 and two H4 subfamilies previously characterized (Chaubet et al. 1989). As an example we show here the genomic blots hybridized to the 5′ probes corresponding to three H3 subfamilies: H3C2, C3, and C4 (Fig. 1 B). Surprisingly, some of these 5′ probes like H3C3, H3C4, and H4C14 revealed polymorphism with almost all of the nucleases tested, while others like H3C2 and H4C7 did not.

DNA from the two parental lines, from a diploid F_1 plant, and from each of the eight F_1 monosomic types were digested with a nuclease-generating RFLP for a given 5′-probe and blot hybridized to this probe. The autoradiograms showing the localization of H3C3, H4C14, and H3C4 subfamilies are presented in Fig. 2. In the case of H3C3 and H4C14 subfamilies, a DNA fragment that was present in the female parent was absent in one of the monosomic types, clearly showing that the loci for H3C3 and H4C14 are located on chromosomes 9 and 2, respectively. However, when the H3C4 5′ region was used as a probe, each of the eight monosomic types analyzed showed a hybridization pattern that was identical to the pattern in diploid siblings; therefore, this locus is not present on any of these monosomic chromosomes and must probably be located on one of the two chromosomes (1or 5) for which no monosomics were available. These hybridizations thus enabled the localization of three of the five previously identified subfamilies. The two others could not be localized because polymorphisms were not found in the parental lines with the nucleases tested.

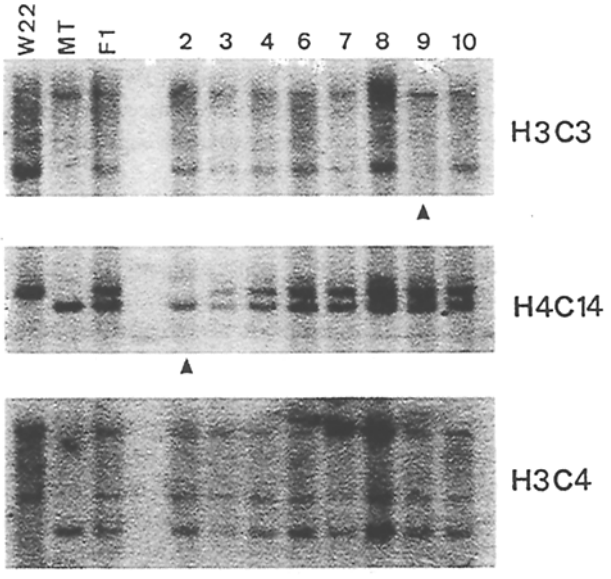

Fig. 2. Monosomic analysis of some H3 and H4 subfamilies. The 5′-probes of the subfamilies which were shown to produce RFLP in W22 and MT were used to probe Southern blots containing genomic DNAs prepared from the parental lines W22 and MT, a diploid F_1 and F_1 individuals monosomic for the chromosome listed on top of each lane. A *Bgl*II blot was successively probed with H3C3 and H3C4 5′-probes and a *Hin*dIII blot with H4C14 5′-probe. The *arrowheads* indicate the monosomic type missing the female parental DNA fragment

Localization using the H3 and H4 coding regions as probes

In order to gain further information on the location of other yet unidentified subfamilies, we hybridized the same genomic blots to H3 and H4 coding regions as probes. Because many DNA fragments of different sizes hybridize to H3 and H4 probes at different intensities, it is more difficult to understand the patterns, and we have restricted our conclusions to those data that could be unambiguously interpreted.

Three possibilities exist for the interpretation of these autoradiograms, considering the fact that one of the chromosomes from the female parent is missing in a monosomic genome: (1) if a DNA fragment hybridizing to a probe in the female genome has no counterpart in the male genome and is missing in one of the monosomic types, then the locus which codes for this fragment must be located on the missing chromosome (Fig. 3, black arrowheads); (2) if a DNA fragment is present in both parents but hybridizes less intensely to one of the monosomics than to the other fragments, it is suggestive of a location on the missing chromosome (Fig. 3, open arrowheads); (3) if a DNA fragment specific to the female genome is present in all monosomic types, then the locus which codes for this fragment is located on chromosome 1 or 5 (as for H3C4 in Fig. 2).

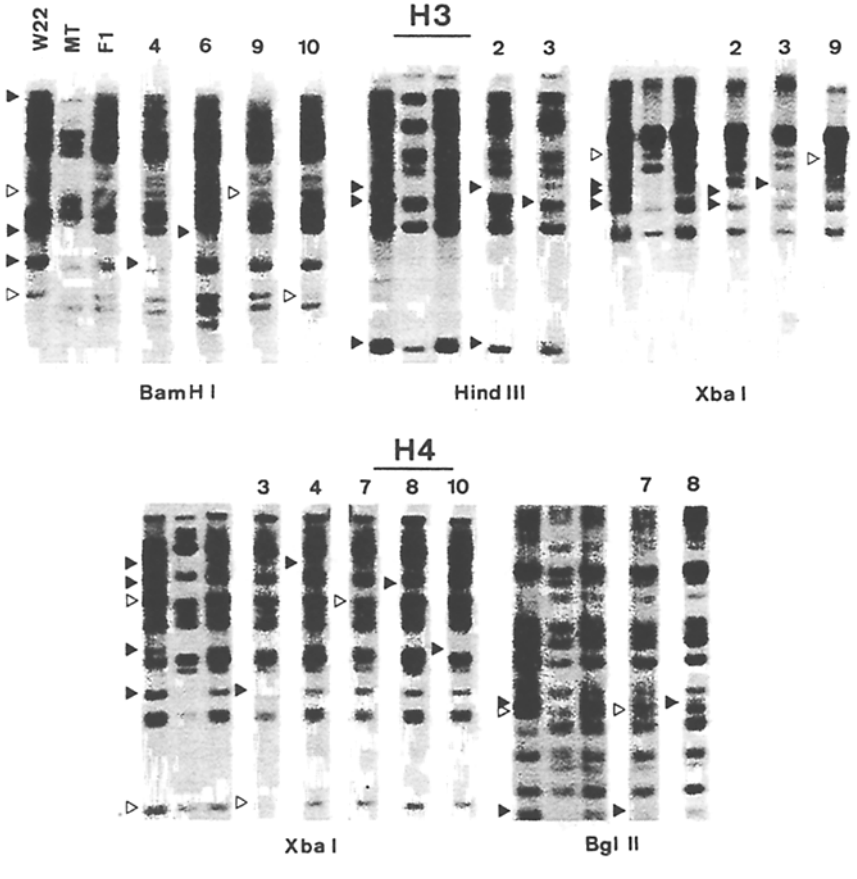

Fig. 3. Monosomic analysis of the H3 and H4 multigenic families. Monosomic genomic blots were probed with the maize histone H3 and H4 coding regions. For clarity only those lanes presenting a modified pattern are presented. For each particular blot the *three lanes* on the *left* correspond to the female (*W22*), male (*MT*) and F_1 lines. *Numbers* on *top* of the *other lanes* indicate the missing chromosome in the monosomic types. *Black arrowheads* point to female-specific missing fragments and *open arrowheads* to bands common to male and female parents but hybridizing with reduced intensity in some monosomic types

By combining the hybridization patterns obtained with the specific 5′-probes (Fig. 2) and H3 and H4 coding regions (Fig. 3) on monosomic genomic blots digested with four different endonucleases (*Bam*H1, *Bgl*II, *Hin*dI-II, *Xba*I), we have been able to map H3 subfamilies on chromosome 1 or 5 as well as on chromosomes 2, 3, 4, 6, 9, and 10 and H4 subfamilies on chromosome 1 or 5 as well as on chromosomes 2, 3, 4, 7, 8, and 10 at least. Because each blot contained many bands which could not be assigned to a particular chromosome due to the complexity of the hybridization pattern and to the absence of any visible RFLP, this result has to be considered to be the minimum possible. Thus, the H3 and H4 histone genes are likely to be distributed on (nearly) all of the chromosomes.

Discussion

We showed previously that the histone H3 and H4 multigene families are organized into eight to ten subfamilies, each of them being characterized by its proximal environment as well as by a different copy number (Chaubet et al. 1989). Using the 5′ non-transcribed regions of five genes (three H3 and two H4) as probes to hybridize to genomic blots of maize monosomic lines (Helentjaris et al. 1986), we were able to localize three subfamilies on three different chromosomes: 1 or 5, 2, and 9 for H3C4, H4C14, and H3C3, respectively. Moreover, using the H3 and H4 coding regions as probes, we could detect the presence of H3 and/or H4 genes on almost all the maize chromosomes.

As mentioned above, the histone genes have been generally found to be grouped in the majority of eukaryotes (Hentschel and Birnstiel 1981; Maxson et al. 1983). However, in several higher eukaryotes, groups of histone genes have been found at different loci. For example, in the genome of chicken they exist as two major groups of ten copies of each of the histone genes (D'Andrea et al. 1985; Grandy and Dodgson 1987); in mouse two loci have been identified on chromosomes 3 and 13 (Marzluff and Graves 1984), and in the human genome they are even more dispersed as several loci have been found on chromosomes 7 (Chandler et al. 1979), 1, 6, and 12 (Tripputi et al. 1986). However, in no case was such a highly dispersed organization found as the one we describe here in maize.

Long fragments of duplicated sequences shared by different chromosomes have recently been shown to exist in the genome of maize (Helentjaris et al. 1988). Among 217 probes used to establish a linkage map, 62 (29%) were shown to detect more than one locus. Also, several isoenzymes such as alcohol dehydrogenase (Dennis et al. 1984) and sucrose synthase (McCarty et al. 1986) have been shown to be encoded by two different loci, and genes encoding the zeins, the major family of storage proteins in maize, have been localized on three different chromosomes (Soave et al. 1981, 1982; Viotti et al. 1982). However, up to now there has been no evidence for a multigenic family to be distributed on least, seven to eight and perhaps all of the chromosomes of maize.

Multiple homologous sequences existing in the same genome may originate from (1) duplication and subsequent dispersal of an ancestral gene, and/or (2) a combination of related genomes. It is difficult to imagine that the insertion of reverse-transcribed histone mRNA in the genome could account for the actual organization of these genes, as it is hardly conceivable that such reinserted histone "genes", whose expression is rigorously controlled, could be set under their proper promoter elements when integrated at random into the genome.

On the contrary, tandem duplication and subsequent dispersal could certainly be responsible for such a dispersed organization. Originally, an ancestral H3 or H4 gene could have undergone duplication, leading to an ancestral gene cluster. From such a cluster some copies could have been transferred to other regions of the genome and undergone independent evoluton giving rise to the actual H3 and H4 subfamilies by successive duplications and translocations. Such events likely arose at different stages in the evolution of the genome, as suggested by the existence of more or less extensive sequence homologies between the 5′ flanking regions of some of the genes: for instance, among the three cloned H3 genes, two of them, namely H3C2 and H3C4, show limited but significant sequence homology between their 5′-flanking regions, whereas the third gene, H3C3, does not (Chaubet et al. 1989). Thus, it is likely that these two genes diverged from each other more recently than did their common ancestor from the C3 gene.

As suggested earlier (Helentjaris et al. 1988), the fusion of related genomes could also have contributed to such a complex organization. However, as all five H3 and H4 subfamilies identified in maize also exist, with a very similar if not identical environment, in several related genomes such as sorghum, *Coïx*, teosinte, and sugar cane (Chaubet et al. 1989), the genome fusion, as well as the gene duplication/dispersal, should have occurred among ancestral plants before the divergence of the andropogonoideae into different tribes.

It is well known that histone genes and particularly those encoding H3 and H4 are under a high and constant selective pressure. On the other hand, as it is known that clustering of the genes facilitates sequence conservation by conversion mechanisms, such a dispersed organization of maize histone genes is surprising. It will be very interesting to determine whether this organization relates to some regulatory mechanisms particular to each subfamily or if a special mechanism exists to ensure the concerted expression of these so highly dispersed genes.

Acknowledgements. We are very grateful to Mrs. B. Clement and M. Ehling for skillful technical assistance. We express our appreciation to Ciba Geigy Seed Division for providing the summer nursery facilities in which the plants analyzed in this study were grown. This work was supported by research funds from USDA 6-CRC-1-2213 and 89-37140-4863.

References

Chandler ME, Kedes LH, Cohn RH, Yunis JJ (1979) Genes coding for histone proteins in man are located on the distal end of the long arm of chromosome 7. Science 205:908–910

Chaubet N, Philipps G, Chaboute ME, Ehling M, Gigot C (1986) Nucleotide sequences of two corn histone H3 genes. Genomic organization of the corn histone H3 and H4 genes. Plant Mol Biol 6:253–263

Chaubet N, Philipps G, Gigot C (1989) Organization of the histone H3 and H4 multigenic families in maize and in related genomes. Mol Gen Genet 219:404–412

Childs G, Nocente-McGrath C, Lieber T, Holt C, Knowles JA (1982) Sea urchin (*Lytechinus pictus*) late-stage histon H3 and H4 genes: characterization and mapping of a clustered but non-tandemly linked multigene family. Cell 31:383–393

D'Andrea RJ, Coles LS, Lesnikowski C, Tabe L, Wells JRE (1985) Chromosomal organization of chicken histone genes. Preferred associatons and inverted duplications. Mol Cell Biol 54:3108–3115

Dennis ES, Sachs MM, Gerlach WL, Finnegan EJ, Peacock WJ (1985) Molecular analysis of the alcohol dehydrogenase 2 (*Adh2*) gene of maize. Nucleic Acids Res 13:727–743

Gigot C (1988) Histone genes in higher plants. In: Kahl G (ed) Architecture of eukaryotic genes. VCH Verlagsgesellschaft, Weinheim, pp 1367–1373

Grandy DK, Dodgson JB (1987) Structure and organization of the chicken H2B histone gene family. Nucleic Acids Res 15:1063–1080

Helentjaris T, Weber D, Wright S (1986) Use of monosomics to map cloned DNA fragments in maize. Proc Natl Acad Sci USA 83:6035–6039

Helentjaris T, Weber D, Wright S (1988) Identification of the genomic locations of duplicate nucleotide sequences in maize by analysis of restriction fragment length polymorphism. Genetics 118:353–363

Hentschel CC, Birnstiel ML (1981) The organization and expression of histone gene families. Cell 25:301–313

Kedes LH (1979) Histone genes and histone messengers. Annu Rev Biochem 48:837–870

Lifton RP, Goldberg ML, Karp RW, Hogness DP (1977) The organization of the histone genes in *Drosophila melanogaster*: Functional and evolutionary implications. Cold Spring Harbor Symp Quant Biol 42:1047–1051

McCarty DR, Wright S, Helentjaris T, Shaw J, Hannah LC (1986) The sucrose synthetase-2 gene is genetically linked to *sh1*. Maize Genet Coop News 60:60–61

Marzluff WF, Graves RA (1984) Organization and expression of mouse histone genes. In: Stein G, Stein J, Marzluff WF (eds) Histone genes: structure, organization and function. John Wiley and Sons, New York, pp 281–315

Maxson R, Cohn R, Kedes L (1983) Expression and organization of histone genes. Annu Rev Genet 17:239–277

Soave C, Rhigetti R, Di Fonzo N, Salamini F (1981) Clustering of genes for 20 K subunits in the short arm of maize chromosome 7. Genetics 97:363–377

Soave C, Reggiani R, Di Fonzo N, Salamini F (1982) Genes for zein subunits on maize chromosome 4. Biochem Genet 11:1027–1037

Stephenson EC, Erba HP, Gall JG (1981) Characterization of a cloned histone cluster of the newt *Notophthalmus viridescens*. Nucleic Acids Res 9:2281–2295

Tripputi P, Emanuel BS, Croce CM, Green LG, Stein GS, Stein JL (1986) Human histone genes map to multiple chromosomes. Proc Natl Acad Sci USA 83:3185–3188

Viotti A, Abildstein D, Pogna N, Sala E, Pirrotta V (1982) Multiplicity and diversity of cloned zein cDNA sequences and their chromosomal localization. EMBO J 1:53–58

Weber DF (1973) A test of distributive pairing in *Zea mays* utilizing doubly monosomic plants. Theor Appl Genet 43:167–173