Nucleic Acids Research, 1995, Vol. 23, No. 11 1901–1911

Chromosomal mapping of the MADS-box multigene family in *Zea mays* reveals dispersed distribution of allelic genes as well as transposed copies

Achim Fischer, Nikolai Baum, Heinz Saedler and Günter Theißen*

Max-Planck-Institut für Züchtungsforschung, Abteilung Molekulare Pflanzengenetik, Carl-von-Linné-Weg 10, D-50829 Köln, Germany

Received February 28, 1995; Revised and Accepted April 20, 1995

ABSTRACT

A linker PCR procedure has been developed for preparing repetitive DNA-free probes from genomic clones, which is especially efficient for members of gene families. Using this procedure as well as standard methods to prepare hybridization probes, chromosomal map positions of MADS-box genes were determined in recombinant inbred lines of maize (*Zea mays* ssp. *mays*). It appears that MADS-box genes are scattered throughout the maize genome. While there is evidence that this genomic distribution is representative for plant MADS-box genes in general, the following two other observations probably reflect *Zea* genome organization. First, at least one family of MADS-box-carrying elements contains line-specific versions, which are present in some maize lines at certain chromosomal positions, but are absent from these loci in other lines. The members of this family resemble transposable elements in some respects. Secondly, the finding of pairs of highly related MADS-box genes which are accompanied by other duplicated markers is a further indication of the ancestral polyploid genome constitution revealed with other markers. The importance of these findings for an understanding of the genomic organization of MADS-box genes and the evolution of the MADS-box gene family is discussed.

INTRODUCTION

One of the major efforts in current biological research is the elucidation of developmental mechanisms in multicellular organisms. During the past few years it has turned out that many genes of central importance for development are members of multigene families, with the most prominent examples being the homeobox genes (1). The proteins encoded by these genes function as transcription factors and, among other roles, largely determine the body plan of animals (2). Accordingly, mutations in homeobox genes can lead to severely impaired phenotypes, e.g. homeotic transformations, in which body parts develop in the wrong positions. A well-known example is the *Antennapedia* mutant of *Drosophila melanogaster*, which has legs growing in place of antennae (3).

Cloning of the first homeotic genes from plants yielded the surprising result that most of them also belong to just one gene family (4–9; for reviews, see 10,11), termed the MADS-box genes (12). All MADS-box genes share a highly conserved, ~180 bp long DNA sequence, the MADS-box, which encodes the DNA-binding domain of the respective MADS-domain transcription factors. The occurrence of the MADS-box gene family is not restricted to plants, but has also been established in other eukaryotes from yeast to man (10). In plants, MADS-box genes have become known due to their important role in flower development (12). Those genes which have been characterized in some detail function either as flower meristem or organ identity genes. The floral meristem identity genes, such as *SQUA*, from *Antirrhinum majus*, function in establishing the identity of the floral primordium (7). The flower organ identity genes, like *DEFA*, *GLO* and *PLE*, from *A.majus*, determine the identity of different flower organs, the petals, stamens and carpels (4,6,13).

Although to date a considerable number of studies on plant MADS-box genes have been published, most of them focus on dicotyledonous model plants like *A.majus* and *Arabidopsis thaliana* (for a review, see 11). Little is known about MADS-box genes in monocots, though this taxon comprises the most important crop plants, whose inflorescence and flower development is of central agronomical importance. Another obviously neglected aspect of today's studies on MADS-box genes is their chromosomal localization. This seems unfortunate for two reasons. First, in case of homeobox genes, the organization into gene complexes of defined order is obviously of functional and evolutionary relevance (14), and one may ask whether the same is also true for other developmental control genes. Secondly, precisely determined chromosomal map positions of isolated genes could be compared with genomic loci defined by morphological mutants in order to determine whether the mapped genes are good candidates for the mutant loci.

Therefore, we started to investigate the chromosomal organization of the MADS-box gene family in maize (*Zea mays* ssp. *mays*), a typical monocotyledonous plant which has been well characterized genetically and molecularly. Maize seemed especially promising for these studies because a large set of interesting inflorescence and flower mutants has been described (15), some of which could well be due to mutations in MADS-box genes.

* To whom correspondence should be addressed

To localize the map position of a given clone the use of recombinant inbred (RI) lines has found wide application in many organisms, among them maize (16). The standard mapping experiment requires the preparation of a labeled probe, which is then hybridized to Southern blots of genomic DNA isolated from the members of the RI population. Many studies have successfully used cDNAs as probes. However, in certain cases it may be desirable to use genomic clones for preparing gene specific probes. This in turn may become laborious for two reasons. First, one has to get rid of the conserved regions shared by other gene family members, which often makes it necessary to determine the sequences and exon–intron-structures of the genomic clones. The second requirement is to avoid repetitive sequences, because their presence in a hybridization probe for genomic DNA obscures other signals. Unfortunately, the genomes of many higher organisms are rich in repetitive DNA, with maize being no exception (17–21).

In order to facilitate the RFLP mapping of large gene families, a technique based on linker (or ligation mediated) PCR (22) has been developed, which enables an efficient preparation of repetitive DNA-free probes from genomic clones. Applying this technique as well as the use of conventional probes for the mapping of MADS-box genes provided some unexpected and intriguing insights into the chromosomal organization of these genes in maize.

MATERIALS AND METHODS

Gene isolation and conventional preparation of hybridization probes

Genomic MADS-box clones were obtained from a maize λEMBL4 library, and cDNAs were prepared by RACE cloning as described recently (23). The MADS-box regions of all clones were sequenced using the '*fmol* DNA sequencing system' (Promega).

Conventional isolation of hybridization probes from genomic clones was carried out as follows. After restriction digest of phage DNA, fragments were separated by gel electrophoresis and blotted to Hybond N^+ membranes (Amersham). Blots were hybridized with a radiolabeled MADS-box probe (23). Insert bands not containing MADS-box sequences were recovered from agarose gels (24) and were radiolabeled using random oligonucleotide primers (24) either directly or after subcloning. Unincorporated radionucleotides were removed from probes on Nick columns (Pharmacia).

Probes representing the last exon of genes *ZAG2* and *ZMM1* were obtained by PCR employing λEMBL4-I17b or λEMBL4-I16b (23), respectively, as templates, and P029 (5′-TCAGATTGAGGAGGGAGAGCAA-3′) and P030 (5′-CTCGTACATGCATCCAGTTTAG-3′) in both cases as primers, followed by radiolabeling as cited above. Oligonucleotides were synthesized on a '392 DNA/RNA Synthesizer' (Applied Biosystems).

cDNAs present in BRACE (23) clones were cut out of their plasmid vectors, purified by agarose gel electrophoresis, and recovered and labeled as cited above, yielding 'complete RACE probes'. Alternatively, preferential labeling of the 3′ end of a cDNA was achieved by 'linear PCR' as follows, yielding '3′-end probes'. Labeling was done in a total volume of 10 µl containing 1 × *fmol* sequencing buffer (Promega), 20 pmol RACE adapter primer (25), 100 ng BRACE9-* (template; with '*' being 1, 12, 22, 37, 60), 2 µM each of dATP, dGTP and dTTP, 2.5 U *Taq* polymerase (Promega) and 50 µCi (1.7 µM) [α-^{32}P]dCTP (Amersham). A PCR program was run for 25 cycles as suggested in the protocol of the '*fmol* DNA sequencing system'.

Preparation of hybridization probes by linker PCR

For the preparation of linker L27/LR28 ('blunt end linker'), equal molar amounts of oligonucleotides L27 (5′-ATCTATACTTTCGGGTTGAGAATCGAG-3′) and LR28 (5′-ACTCGATTCTCAACCCGAAAGTATAGAT-3′) (30 pmol/µl each) were hybridized as described elsewhere (53). Similarly, a '*Mse*I-linker' with a two base 5′-overhang 'TA' was synthesized for ligation to DNA ends generated by restriction digest with *Mse*I.

Ligation of linkers to phage DNA was done as follows: 1 µg phage DNA was digested with 5 U of restriction enzymes *Hae*III (Boehringer) (yielding 'blunt ends') or *Mse*I (NEB) (yielding 'staggered ends'), respectively, for 3 h at 37°C in a volume of 30 µl 1 × incubation buffer. Digests were phenolized and ethanol precipitated. Approximately 150 ng digested phage DNA, resuspended in TE buffer, was ligated to 500 ng (some 50-fold molar excess) suitable linker DNA by 0.2 U T4 DNA ligase (Boehringer) in 4 µl 1 × ligation buffer (Boehringer) overnight at 16°C. Unincorporated linkers were removed on Qiagen tips 5 (Qiagen Inc.). Eluted ligation products were ethanol precipitated and resuspended in 4 µl of TE buffer.

Ligation products were amplified by PCR in Trioblock thermocyclers (Biometra). In a total volume of 50 µl PCR buffer [67 mM Tris–HCl, pH 8.8 (RT), 17 mM $(NH_4)_2SO_4$, 0.1% Tween 20, 1.5 mM $MgCl_2$, 200 µM dNTPs each], 0.5 µl linker ligation product were combined with 25 pmol primers LR28 and P008 (5′-CGAGAGCTCGTACGCCTTCTTGAG-3′) each and overlaid with paraffin oil. After initial denaturation (3 min at 94°C) and 'hot start' with 2.5 U *Taq* polymerase (Boehringer), a cycling program (25 or 30 cycles) was run as follows: denaturation 30 s, 93°C; primer annealing 30 s, 60°C; extension 90 s, 72°C. Amplification products were purified on Qiagen tips 5. Subsequent radiolabeling of PCR products was done using random oligonucleotide primers (24).

Mapping of genomic loci

Chromosomal map positions of MADS-box genes were determined by applying RFLP technology to maize recombinant inbred lines T × CM and CO × Tx (16,26). The maize lines were from B. Burr, Brookhaven National Laboratory, NY. Plants were grown in the greenhouse, harvested after 8–9 weeks and used for preparing genomic DNA following standard protocols (27).

Ten micrograms DNA of each maize plant was digested with an appropriate restriction enzyme, run on a 0.8% agarose gel and blotted to Hybond N^+ membranes. RFLP searchfilters with parental DNAs cut with six different enzymes (*Bam*HI, *Bgl*II, *Dra*I, *Eco*RI, *Eco*RV and *Hin*dIII) were prepared to detect RFLPs; if no RFLP was found, another six enzymes (*Kpn*I, *Pst*I, *Pvu*II, *Sca*I, *Sac*I and *Xba*I) were tried. RFLP filters with parental and RI DNAs cut with a particular enzyme showing RFLP with a probe of interest were prepared to analyze the strain distribution patterns of restriction fragment lengths. Filters were hybridized with radiolabeled probes obtained as described above at 65°C overnight in 3 × SSPE (24)/0.02% ficoll/0.02% polyvinyl

pyrolidone/0.1% SDS/100 μg/ml salmon sperm DNA, and washed 3 × 20 min in 0.2× SSPE/0.1% SDS at 65°C.

The band patterns revealed by hybridization of RFLP filters were analyzed by identifying each of two segregating RI bands as being either 'parental 1' (for parents CM37 or Tx303, respectively) or 'parental 2' (for parents T232 or CO159, respectively) according to (26). In the case of line specific gene versions, however, scoring strategy had to be modified. Because respective bands appeared only in one parent and in ~50% of the progeny, classifying a band as having the size of the corresponding band of either parent 1 or parent 2 had to be replaced by dealing with 'allelic states'. For example if the respective allele was present in parent 1 and absent in parent 2, presence of the respective band in a RI plant was evaluated as 'state 1', its absence as 'state 2'. Note that the normal way of mapping allelic bands is a special case of this more general procedure, as the two strain distribution patterns obtained by two allelic segregating bands are identical. Mapping positions were computed from the obtained strain distribution patterns by B. Burr.

RESULTS

A linker PCR-technique for the preparation of gene-specific probes

In the course of a project dealing with the structure and function of MADS-box genes in maize, genomic and cDNA clones of maize MADS-box genes were isolated from a *wx-844* maize line and from line C, respectively, as described elsewhere (23). Here, we are concerned with the genomic organization of these genes. Since the MADS-box gene family is quite large in maize (28), we became interested in a procedure to prepare hybridization probes suitable for RFLP mapping even from genomic clones without the need of a detailed structural characterization of the isolated clones.

In the beginning of the mapping efforts, restriction fragments of genomic phage inserts were tested as hybridization probes. From a whole series of experiments, however, we had to conclude that within the vicinity of a few kb of MADS-box sequences in maize, repetitive DNA is regularly found, and thus, arbitrarily chosen DNA restriction fragments of genomic clones are generally not suitable as gene-specific probes (GT, unpublished).

The only type of standard experiment leading to satisfactory hybridizations (data not shown) and subsequently to mapping data with a genomic clone (Table 1) was the use of a probe obtained from an exon region of *ZMM1*, a gene which had been structurally characterized in detail (23). However, this latter approach of course has the disadvantage of requiring detailed information about sequence and exon–intron structure of the respective clone. Therefore, a new, convenient technique of probe preparation is highly desirable to effectively study the arrangement of genomic loci of a gene family.

It was assumed that if the highly conserved MADS-box was flanked at least on one side by a unique sequence, it should be possible to amplify this gene section from a genomic clone using the MADS-box as an anchor. For several reasons (see Discussion) the method of choice to accomplish such an amplification is linker PCR as previously described (22). The amplification product should then yield a suitable template for preparing hybridization probes.

Therefore, the DNA of the MADS-box genes containing phages was digested with four-base cutters *Hae*III or *Mse*I into short subfragments, one of which contains part of the MADS-box appropriate as primer binding region. This set of fragments was then treated with a ligase mix containing linkers in excess. One linker end was tailored to match the fragments' ends ('blunt end linker' for ligation to *Hae*III-ends, '*Mse*I-linker' with a overhang for *Mse*I-ends), the other end was characterized by a single base overhang to avoid linker self ligation to long aggregates. Considering all available MADS-box protein sequences, primer P008 was designed as MADS-box specific primer corresponding to the nucleotides encoding the highly conserved 'LKKAYELS' amino acid motif. P008 was in an orientation to allow amplification of the region directly upstream of the MADS-box. A PCR was then performed using LR28 as linker primer and P008 as MADS-box specific primer (Fig. 1). As the linkers are 3′- and 5′-unphosphorylated, during the ligation step one strand of each linker molecule (oligonucleotide L27) was not covalently linked to the DNA backbone of the phage fragments. This strand dissociates under PCR denaturation conditions and therefore cannot serve as linker primer binding site. For this reason amplification can only start from the internal primer binding site, in our case from the MADS-box motif.

Thus, MADS-box gene specific PCR-products are obtained which represent the region directly flanking the primer binding site, here the upstream region adjacent to the MADS-box. Using the DNA of 14 different phage clones, at this time identified as MADS-box genes only by hybridization experiments with a MADS-box probe (23), a specific amplification product was obtained with P008 in every case (for examples, see inset of Fig. 1). Later, the presence of a MADS-box was verified for every clone also by sequencing (see below). However, this was done after the mapping experiments had been finished, demonstrating that a detailed sequence information was not a prerequisite for the application of the linker PCR procedure.

PCR products were purified, radiolabeled and hybridized against genomic Southern blots of maize DNA. This procedure proved to be equivalent to the use of probes derived from identified exon regions in that it reproducibly provided satisfactory hybridization results. An example is shown in Figure 1. The obtained clear band patterns allowed scoring of the allele distribution patterns of the genomic clones in the mapping populations and the assignment of chromosomal locations (Fig. 2 and Table 1). The number of the loci determined with probes from genomic clones is smaller than the total number of genomic clones, because some of these clones represent the same genes, as became evident by identical map positions and subsequent sequence analyses (redundant data not shown).

Because of the success of probes generated with the upstream orientated primer P008, probes obtained with downstream orientated primers were tried only in two cases (clones λEMBL4-I7a and -I15a). In these cases, however, hybridizations of genomic maize Southern blots revealed the presence of repetitive DNA on the probes (data not shown).

MADS-box genes are dispersed throughout the maize genome

Using linker PCR products prepared as outlined above, seven different chromosomal map positions were determined (Table 1). Moreover, 13 different loci were mapped using either defined exon regions or cDNA sequences as hybridization probes (Table 1). Six map positions were taken from the literature. In Figure 2, all loci are placed together on a chromosomal map. It can be seen

Table 1. Map positions of MADS–box carrying elements determined in this work

Clone[a]	Mapping probe[b]	Marker[c]	Chromosomal localization[d]	Gene[e]	Mapping population and presence of alleles[f] T×CM (CO×Tx)
BRACE9-1	cDNA	***mpik24***	8L @ *ucsd78B* (8L055.6)	***ZMM2***	+ +
BRACE9-12	cDNA	***mpik23A***	1L between *uwo2* (1L130.8) and *umc58* (1L134.0)	***ZMM6***	+ +
BRACE9-22	cDNA	***mpik28***	9L between *bnl14.28* (9L111.2) and *npi93A* (9L114.7)	***ZMM8***	(+ +)
BRACE9-37	cDNA	***mpik25***	9S between *bz1* (9S035.2) and *npi266* (9S043.9)	***ZMM3***	+ +
BRACE9-60	cDNA	***mpik27***	7L between *npi389* (7L088.8) and *ast(amy3)* (7L103.5)	***ZMM7***	+ +
EMBL4-I16b	*Hae*III/L-PCR	***mpik31***	10L @ *ucsd81B, glu1, uaz97, uaz98* (10L054)	***ZMM1***	+ +
"	last exon	***mpik21***	"	"	+ +
EMBL4-I17b	*Hae*III/L-PCR	***mpik32A***	3L between *abp1* (3S083.9) and *bnl6.06* (3L085.7)	***ZAG2***	+ +
EMBL4-I120	*Mse*I/L-PCR	***mpik22A***	1L near *bnl8.29A* (1L243.2) and *ucsd104B*	***ZMM4***	+ +
EMBL4-I5a,-I15a -I20a[g]	*Hae*III/L-PCR[h]	*mpik17E*	10S @*ucsd64D* (10S007.8)	***TMZ1-9***	+ -
"	*Hae*III/L-PCR[h]	***mpik16C***	8L @ *ucsd78B* (8L055.6)	***TMZ1-7***	+ +
"	segment B[i]	*mpik11B*	6S between *umc85* (6S004.5) and *bnl6.29* (6L006.9)	***TMZ1-5***	+ -
"	*Hae*III/L-PCR[h]	*mpik12D*	@ *mpik11C* (see below)	***TMZ1-8***	- +
"	segment B[i]	*mpik11C*	9S between *bz1* (9S035.2) and *npi266* (9S043.9)	"	- +
"	*Hae*III/L-PCR[h]	*mpik12E*	@ *mpik11D* (see below)	***TMZ1-3***	+ -
"	segment B[i]	*mpik11D*	4L between *dnap4* (4L075.4) and *uaz71A* (4L078.5)	"	+ -
EMBL4-I20a	segment C[j]	***mpik331***	2L between *umc2B* (2L099.3) and *npi356*(2L102.5)	***TMZ1-2***	(+ +)
"	"	*mpik33H*	1S @ *umc11* (1S070.5)	***TMZ1-1***	(+ -)
"	"	*mpik33D*	6L between *enp1* (6L025.2) and *npi373* (6L027)	***TMZ1-6***	+ -
"	"	*mpik33E*	5S between *mdh5* (5S085.6) and *bnl6.10* (5S088.4)	***TMZ1-4***	- +

[a]λEMBL4 ('EMBL4') clones contain genomic DNA fragments from a maize line carrying the *wx-844* allele (23), BRACE clones ('BLUESCRIPT vector with RACE inserts') contain partial cDNAs of maize line C, obtained by RACE cloning (23). If not stated otherwise, all bands visible with the respective probes on stringently washed mapping filters have been mapped.

[b]Probes from genomic clones are either linker PCR products representing DNA segments directly upstream of the MADS-box ('*Hae*III/L-PCR' or '*Mse*I-/LPCR', respectively, depending on the restriction enzyme the DNA had been digested with), or conventional PCR products. These latter were obtained from the last exon of the respective MADS-box gene ('last exon' probe) employing primers designed according to the λEMBL4-I17b sequence (see Materials and Methods), or from different segments of *TMZ1* family members, present on genomic clones λEMBL4-I5a, -I15a and -I20a and some others (see text) ['segment B', 'segment C'; the region upstream of the MADS-box ('segment A', see text) is represented by the respective *Hae*III/L-PCR products]. The probes of the *TMZ1* segment B and C regions were provided by Katharina Montag. cDNA probes ('cDNA') were 'complete RACE probes' in case of BRACE9-1, and '3′-end probes' in all other cases (see Materials and Methods).

[c]Under the mpik marker designation the loci can be found on the BNL maize map, of which an updated version is published every year in the 'Maize Genetics Cooperation Newsletter' (Ben Burr, personal communication). Markers representing loci which are present in both parents of a mapping population ('allelic markers') are shown in bold.

[d]The chromosomal localization on the BNL map (16; Ben Burr, personal communication) is given in relation to flanking markers or framework loci. @ ('at') means at the same locus within the resolution power of the data available and methods used. The strain distribution patterns on which the localizations are based are available upon request.

[e]*ZAG2*, *ZMM1* and *ZMM2* have been described elsewhere (23,28). The *TMZ1* (Transposed MADS-box elements of *Zea* No. 1) family is described in the text. Members of that family may have compositional differences, as have been found with elements present on the phage clones mentioned in the table (see also below, and text). An identification number (*TMZ1-1* to *-9*) was given for each locus in order to indicate also possible structural differences between the elements. The other clones were defined as putative MADS-box genes (*ZMM* = *Zea mays* MADS) on the basis of a typical MADS-box sequence, a 3′ sequence similar to those of dicotyledonous MADS-box genes (data not shown) and an allelic mapping behavior.

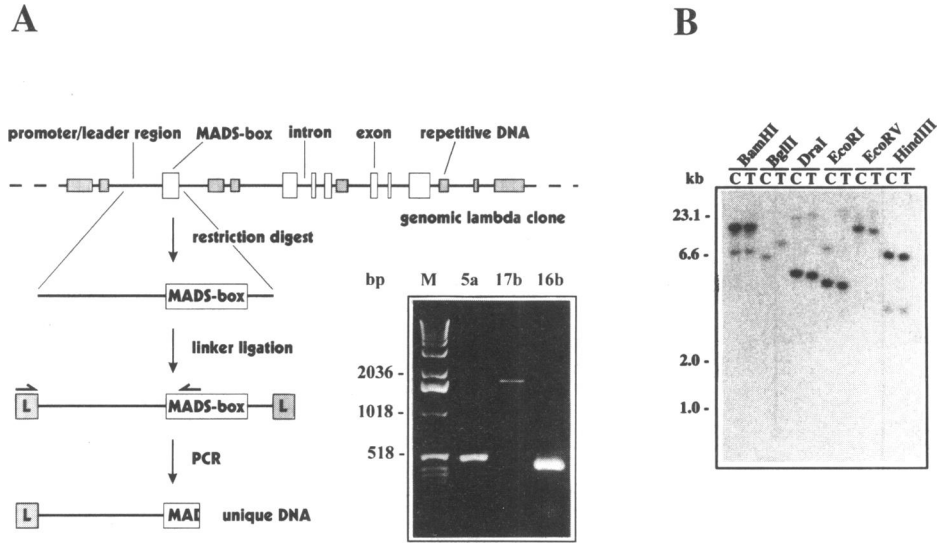

Figure 1. The linker PCR procedure for obtaining hybridization probes. (**A**) Schematic description. Phage DNA containing MADS-box genes with interspersed repetitive DNA (shaded areas) is cut with a frequent cutting restriction enzyme like *Hae*III or *Mse*I and ligated to linkers. PCR is then performed using a linker primer and a primer directed against the MADS-box, thus amplifying the region upstream of the MADS-box. Since in most genes the MADS-box is situated at the 5′-most end of the coding unit (10), a leader or promoter sequence will often be amplified. However, an analogous procedure could be used for amplification of the region downstream of the MADS-box, where regularly an intron is found. The inset shows PCR products obtained from genomic clones λEMBL4-I5a, -I17b and -I16b, respectively, using *Hae*III digested phage DNA, linker L27/LR28, and P008 and LR28 as primers for PCR. The PCR products were separated on a 1.5% agarose gel. (**B**) Suitability of a linker PCR product as hybridization probe. A Southern blot containing genomic DNA from maize inbred lines T232 (T) and CM37 (C) which had been digested with the restriction enzymes indicated above the lanes, was hybridized with a radioactive probe obtained from the region upstream of the MADS-box of clone λEMBL4-I16b, containing gene *ZMM1* (23), amplified according to the linker PCR protocol outlined here, which yielded the fragment shown in the inset of (A), lane '16b'. An autoradiography is shown on which in most lanes a pair of bands is visible because besides recognizing *ZMM1* (strong bands), the probe cross-hybridizes with a second gene, *ZAG2* (weak bands).

that the loci are dispersed throughout the genome. Between one and four loci were located on each chromosome, and there is no chromosome without any. No extensive clustering of MADS-box loci was evident. Nevertheless, a slight preference of centromeric regions appears to be ascertainable, as six out of 23 MADS-box loci are situated within a few centimorgans around the centromeres (see Fig. 2).

However, the genomic organization of MADS-box genes in the maize genome has some subtleties which are not apparent from such a superficial consideration.

Some map positions of MADS-box genes reveal the duplicate character of the maize genome

During the course of the mapping experiments, there were few cases of just one band per lane, indicating one single locus in every maize line (Fig. 3). With the majority of probes, different levels of complexity could be observed. A slightly more complicated situation than that shown in Figure 3 arose with some clones making up pairs of rather similar, but not identical, sequences. One example is the pair of genomic clones

[f] '+ +' means that an allele is present in both parents of the respective recombinant inbred (RI) population; '+ –' in case of the T × CM (CO × Tx) RI family means, that an allele is present in T232 (CO159), but absent ('null allele') in CM37 (Tx303); '– +' vice versa. These symbols are in brackets if the loci were mapped in the CO × Tx family of recombinant inbred lines; otherwise loci were mapped employing T × CM.

[g] These clones carry a family of MADS-box containing sequences (*TMZ1* family), which all share some domains, but differ in others. According to phage Southern blot hybridization studies with the respective clones, the λEMBL4-I20a insert contains a segment A region, a MADS-box, a segment B and a segment C; the λEMBL4-I5a insert deviates from -I20a in that it has no segment C; this is also true for λEMBL4-I15a, but in addition, this element has a segment downstream of the MADS-box which is different from that of the other clones. Similarity between the respective segments of the different clones is >90%, so that each segment-specific probe detects the respective segments of all three clones. For more details, see text.

[h] The linker PCR products obtained with the respective clones represent the segment A region. The probes detect one more locus which could not be mapped due to lack of RFLP; in some maize lines, this locus possibly contains all three segments together (segments A, B, C plus a MADS-box downstream of segment A). Line specific loci on chromosomes 4L (*Bgl*II, *Eco*RV, *Hin*dIII), 9S (*Eco*RV, *Hin*dIII) and 10S (*Eco*RI, *Eco*RV, *Hin*dIII) and the number of loci were found in parallel mapping experiments employing two to three different restriction enzymes as given in parentheses. The allelic locus on chromosome 8L, which is represented by bands on Southern blots which are significantly weaker than those of the line-specific elements, was also mapped with three different restriction enzymes (*Bgl*II, *Eco*RI, *Hin*dIII).

[i] The segment B region is present on λEMBL4-I5a, -I15a and -I20a. Probes from this segment detect two more bands which could not be mapped because no RFLP was found. Positions on chromosomes 4L and 6S were mapped with *Eco*RI and *Hin*dIII as line specific versions, the 9S position was only mapped with *Hin*dIII (no RFLP with *Eco*RI).

[j] With the segment C region, three different bands are visible on each, T232 and CM37 Southern blots (*Bgl*II), thereof two with identical length in both lines (thus not mappable, but probably allelic), and one non-allelic version in each line (*mpik33D* and *E*). The mapping in CO × Tx is complete, i.e.one allelic locus and one extra version explain all the bands segregating on mapping filters. Results were obtained with two different restriction enzymes (*Bam*HI and *Eco*RI).

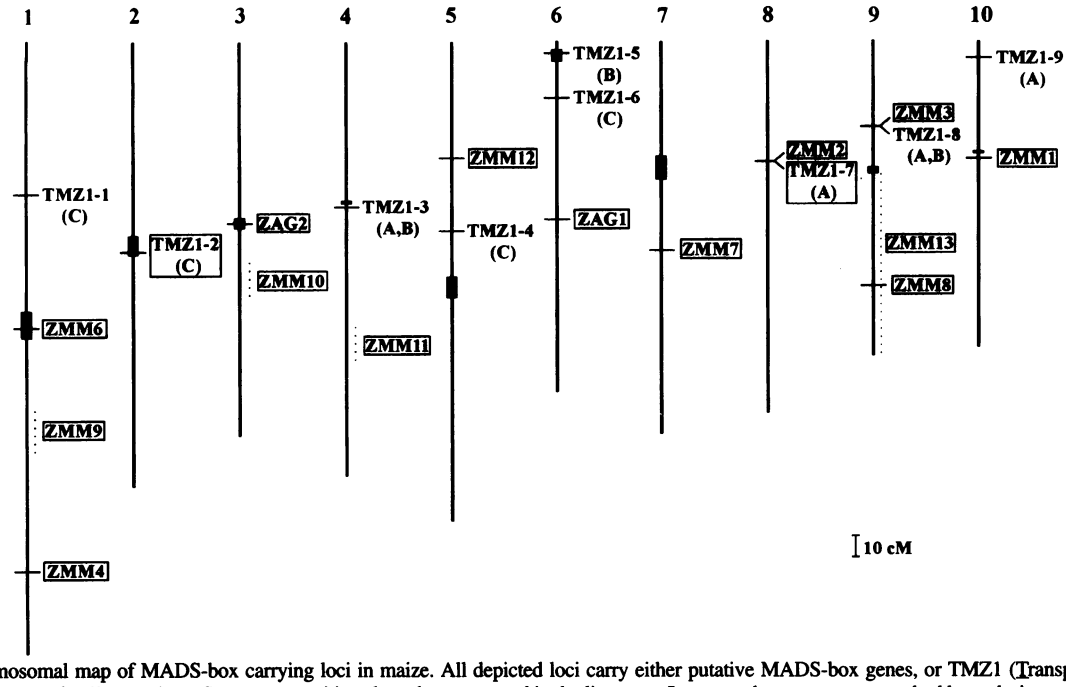

Figure 2. Chromosomal map of MADS-box carrying loci in maize. All depicted loci carry either putative MADS-box genes, or TMZ1 (Transposed MADS-box elements of *Zea* No. 1) family members. Some map positions have been reported in the literature. In cases where no gene name had been designated before, we have continued the ZMM nomenclature (ZMM9–13). Map positions of ZAG1 (28), ZMM12 (=csuh137) (54) and ZMM13 (=uaz231) (55) have been determined by others. This is also true for ZMM9–11, whose approximate locations have been reported by Veit and co-workers (15). The other loci have been determined during the course of this work (Table 1). Loci found in both parents of the respective mapping populations are boxed, line specific gene versions are not. For line specific gene versions of the TMZ1 family, the segments which have been mapped are given in parentheses. Note that in some cases more work is necessary to definitively prove that the different TMZ1 versions lack the respective segments (A, B, C) at those genomic positions where they have not been mapped yet (see text). To unambiguously mark the elements mapped in this work, they have been numbered from TMZ1-1 to -9, thus reflecting the different chromosomal locations, but also indicating potential structural differences. Black bars indicate the approximate positions of the centromeres. The dots beneath ZMM9–11 and ZMM13 indicate that the respective map position was not reported precisely. Note that sequence comparison demonstrates that ZMM8 and ZMM13 are two different genes (our unpublished observations).

λEMBL4-I16b and -I17b, containing genes *ZMM1* and *ZAG2*, respectively. Sequence analysis showed that the derived gene products are 94% identical (23). Using hybridization probes derived from the last, most divergent exon led to strong crosshybridization due to the high level of sequence identity, but allowed the two loci to be distinguished by considering differences in band intensities (Fig. 3). However, with the exon probe, no suitable RFLP was detected for clone λEMBL4-I17b (see, for example, Fig. 3). The regions upstream of the MADS-box of these two clones obtained by linker PCR were also sufficiently different from each other [inset of Fig. 1A and (23)] to unambiguously assign each clone to the corresponding bands on Southern blots via distinct differences in band intensity (Fig. 1). Since in this case a RFLP for λEMBL4-I17b could easily be found with our standard repertoire of restriction enzymes, also this second genomic clone could be mapped (Table 1).

ZMM1 and *ZAG2* on chromosomes 10 and 3, respectively, are accompanied by some duplicated markers; the same is true for *ZAG1* and *ZMM2* on chromosomes 6 and 8, respectively, which are also quite similar in their sequence (23). Therefore, it is very likely that these gene pairs represent recent paralogues which were created by duplication of chromosomal segments and thus represent the well known duplicate (or 'tetraploid') character of the maize genome (23).

From these data we conclude that probes obtained by linker PCR from a polymorphic promoter, leader or intron region will be well suited for distinguishing and mapping genes. Due to the duplicate character of the maize genome comprising pairs of highly related genes, our method may facilitate gene mapping studies in maize.

Line-specific MADS-box containing elements

A new dimension of complexity was provided by some 'genes' where 'solitude' bands were observed, i.e. bands present in one of two RI parents as well as in part of the RI offspring without a counterpart band in the other individuals (see Fig. 3 as an example). Our explanation for this phenomenon is the occurence of line-specific sequence elements. This phenomenon was investigated in some detail for a family of elements which are present in genomic clones λEMBL4-I3a, -I5a, -I7a, -I9a, -I15a, -I18a and -I20a. Sequencing of the MADS-boxes (Fig. 4) and adjacent regions (GT, unpublished results) revealed that (i) these clones very likely represent three different elements, with λEMBL4-I3a carrying the same element as λEMBL4-I5a and -I9a, λEMBL4-I7a containing the same element as -I15a and λEMBL4-I18a the same as -I20a; (ii) even the non-identical elements share >90% sequence identity in the region upstream of the MADS-box and the MADS-box itself (see Fig. 4); (iii) the MADS-boxes present in the elements are very similar to those of *AGAMOUS*-like genes (Fig. 4). From (ii) we conclude that hybridization probes representing the respective regions of the elements probably will not distinguish between the different elements due to the high sequence similarity.

A probe representing the region directly upstream of the MADS-box prepared by our linker PCR procedure, from now on

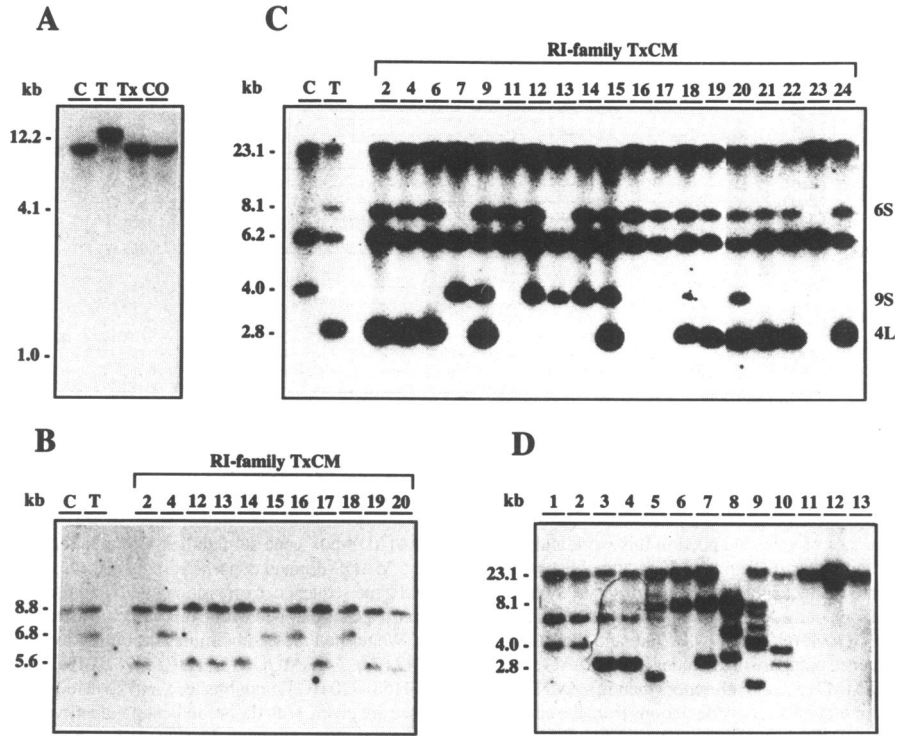

Figure 3. Hybridization of genomic Southern blots revealed different degrees of complexity in genomic representation among the members of the MADS-box gene family in maize. (**A**) A single copy-gene: The hybridization of a '3'-end probe' derived from MADS-box cDNA-clone BRACE9-12 with genomic maize DNA clearly demonstrated the clone to be derived from a single-copy gene (*ZMM6*). (**B**) A gene pair. The figure shows a mapping experiment with a 'last exon' probe from genomic clone λEMBL4-I17b of *ZAG2*. The *ZAG2* locus is represented by the non-segregating bands. The segregating bands indicate a second locus, *ZMM1*. The respective gene is present on λEMBL4-I16b, and encodes a protein with 94% sequence identity to *ZAG2* (23) (compare with Fig. 1B). In other hybridization experiments crosshybridization between members of other gene pairs were observed (not shown). (**C**) Line-specific elements. An even more complicated situation than in (B) can be seen here, where a probe prepared from genomic clone λEMBL4-I5a, representing the segment B region, not only recognized two loci not showing RFLPs (non-segregating bands), but also a set of three non-allelic, line-specific element versions, whose map location (chromosome arm) is indicated at the right margin. (**D**) Hybridization of the same probe as in C with several maize relatives shows that the segment B sequence is present throughout the tribe *Andropogoneae*. Genomic DNA was prepared from maize lines CM37 (lanes 1 and 2) and T232 (lanes 3 and 4), several teosintes (*Zea mays* ssp. *parviglumis*, lane 5; *Zea mays* ssp. *mexicana* 'Nobogame', lane 6; *Zea mays* ssp. *mexicana* 'Northern', lane 7; *Zea mays* ssp. *huehuetenangensis*, lane 8, *Zea luxurians* (lane 9) and *Zea diploperennis* (lane 10), and the non-*Zea Andropogoneae Coix aquatica* (lane 11), *Coix lacryma-jobi* (lane 12) and *Trilobachne cookei* (lane 13). In (A), (C) and (D), *Hin*dIII had been used for digestion of genomic DNA; in (B), the mapping filter had been prepared with *Bgl*II digested RI DNA.

called segment A, detects two to five bands per maize line on genomic Southern blots, depending on maize lines and restriction enzymes used. Two non-allelic positions on chromosomes 10S and 4L, respectively, were mapped in parallel experiments employing three different restriction enzymes, one other position (on 9S) was mapped with two different enzymes (see Table 1 and Fig. 2), which makes it very unlikely that an allelic band with the respective distribution pattern has been overlooked. Thus, these positions are clearly line-specific.

Similar results were obtained with two other DNA fragments, segment B and segment C (gift of K. Montag). Hybridization of Southern blots, carrying DNA of the genomic MADS clones (see above) with probes representing segment B and segment C demonstrated that segment B is present on all the clones discussed here (type λEMBL4-I*a, with '*' being 3, 5, 7, ... as outlined above), whereas segment C was only found in clones λEMBL4-I18a and -I20a (data not shown). Thus, according to the phage Southern experiments, at least one of the cloned elements contains all four regions considered here, namely a MADS-box, a segment A region upstream of it, a segment B and a segment C region, whereas others seem to lack the segment C. The fact that both types of segment combinations were cloned several times independently makes it unlikely that they are simply cloning artifacts.

In two different cases (on chromosomes 4L and 9S), segment B was mapped as a line-specific copy at the same locations as the segment A region, in one case (on chromosome 6S) only a segment B locus was mapped (Table 1; Fig. 2). The positions on 4L and 6S were identified with two different restriction enzymes (described in Table 1). Note that thus one line-specific version and one allelic position on chromosomes 10 and 8, respectively, were found only with the segment A probe, but not with the segment B probe (Table 1; Fig. 2). Due to lack of RFLP two bands obtained with the segment B probe could not be mapped. The invariable presence of these bands in all RI individuals, as exemplified in Figure 3, however, suggests them to represent allelic loci rather than line-specific versions. Thus it is unlikely that a segment B is present at the above mentioned line-specific segment A locus on chromosome 10. Similarly, one can conclude that on chromosome 6S there is a line-specific segment B locus without a segment A (Fig. 2), since the only segment A locus which has not been mapped due to a lack of RFLP (Table 1) seems to be allelic.

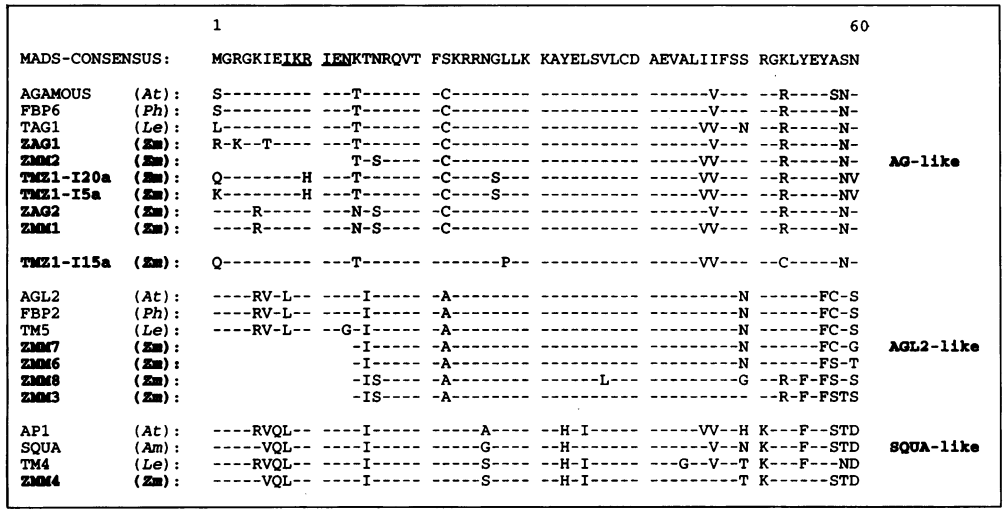

Figure 4. MADS-domain sequences of genes mapped in this work and assignment to MADS-box gene subfamilies. Compilation of all published MADS-domain sequences from yeast, animals and plants (GT and Jan Kim, unpublished data) yielded a MADS-domain consensus sequence, shown in the upper line. Moreover, a clustering of MADS-domains into distinct subfamilies was observed, which reflects total gene sequence, expression pattern and function (GT, unpublished data). As is already obvious from a comparison of the MADS-domain sequences shown here, the maize genes mapped within this work are members of three of these subfamilies, the *AGAMOUS*- (*AG*-) like, the *SQUAMOSA*- (*SQUA*-) like and the *AGL2*-like genes. We termed these subfamilies according to the first members that have been molecularly described. The sequence comparison includes ZAG1, ZAG2, ZMM1, ZMM2 (23), AGAMOUS (5), FBP2 (37), TM4, TM5 (38), AGL2 (35), SQUA (7), AP1 (9), FBP6, TAG1, ZMM3, ZMM7 (53, and references therein), ZMM4, TMZ1-I5a, -I15a, -I20a (GT, unpublished), and ZMM6, ZMM8 (GT and AF, unpublished). For the individual MADS-domain sequences, only deviations from the consensus sequence are given, with dashes indicating identity with consensus. TMZ1-I5a, -I15a and -I20a are encoded by the MADS-boxes present in genomic clones λEMBL4-I5a, -I15a and -I20a, respectively. Due to cross-hybridization of the respective probes (see text), the exact correspondence of the cloned and sequenced elements to the mapped family members TMZ1-1 to -9 is unknown. Note that the MADS-domain of TMZ1-I15a does not fit perfectly into any of these subfamilies but, however, is somehow quite AGAMOUS-like. Some of the ZMM sequences are incomplete due to the RACE procedure applied for cloning (23,53). Underlined amino acids within the consensus sequence indicate the primer binding site employed during RACE cloning. The respective genes are from *Arabidopsis thaliana* (*At*), *Petunia hybrida* (*Ph*), *Lycopersicon esculentum* (*Le*), *Antirrhinum majus* (*Am*) and *Zea mays* (*Zm*) as indicated in the figure.

Hybridizing 'botanical garden blots' (i.e. genomic Southern blots with the DNA of several plant species) with a segment B probe demonstrated that this DNA region is not only present in the genus *Zea*, but also in other *Andropogoneae* like *Coix*, *Trilobachne* (Fig. 3D) and *Chionachne koengii* (data not shown).

The chromosomal positions of the segment C region could be completely mapped in the recombinant inbred population CO × Tx, where one allelic position at 2L and one copy specific for CO159 at 1S were found independently with two different restriction enzymes (Table 1; Fig. 2). In T × CM, the loci of the segment C domain could not be completely mapped, since two bands lacked RFLP, suggesting the presence of one to two allelic positions. Moreover, two loci of line-specific copies were found on chromosomes 5S and 6L which do not match those in CO × Tx.

Botanical garden blots hybridized with a segment C probe demonstrated its presence in all 13 maize lines tested (comprising inbred lines as well as primitive land races), as well as in all six different teosinte species and subspecies (data not shown). With *Tripsacum dactyloides*, only a very faint signal was obtained, and no signal at all was found with more distantly related *Andropogoneae* and with dicotyledonous plants. Note that with probes obtained from 'regular' maize MADS-box genes, under the same experimental conditions a signal is generally obtained throughout all members of the tribe *Andropogoneae* (data not shown).

Our data clearly show that line-specific versions of domains of MADS-box containing sequence elements exist in the maize genome. However, it is not yet clear whether an allelic master copy carrying all mentioned domains, like that cloned as λEMBL4-I20a, is present in all maize lines, because except segment C in CO × Tx, mapping data for all other domains are incomplete due to a lack of RFLPs. However, the data obtained by hybridization of mapping filters with the different domains obtained so far, make it quite unlikely that such a complete 'master copy' is present in maize lines CO159 and Tx303, since no element version seems to exist there which comprises at least those segments (A, MADS-box, B, C) found in clones λEMBL4-I18a and -I20a of the maize line used for genomic cloning, which is different from the mapping lines (see above).

Due to the unusual 'transposed' mapping behaviour we have termed the element family present in the λEMBL4-I*a clones *TMZ1* (Transposed MADS-box elements of *Zea* No. *1*).

DISCUSSION

A linker PCR procedure for the preparation of gene-specific probes

During initial stages of the work outlined here, preparation of gene-specific hybridization probes from poorly characterized genomic MADS-box clones proved difficult due to the ubiquitous presence of repetitive DNA in close proximity, or even within, these genes. Although the maize genome has been estimated to contain 60–80% repetitive DNA (17), it was unexpected to find repetitive DNA regularly near potentially active genes. Meanwhile, however, similar results have also been reported by others (20,21).

Therefore, the idea was to look for gene-specific DNA stretches, such as flanking introns, leader or promoter regions, directly adjacent to the highly conserved MADS-box; with other

gene families, also exons or trailer regions may be considered. To accomplish amplification of regions outside a known sequence, inverse PCR has been employed by others (29). However, this technique has several disadvantages: first, the circularization step is quite difficult. Secondly, two separate primer binding sites within the specific domain have to be derived. In case of scarce sequence data this reduces the chance of successful amplification by the probability of the second primer's non-proper binding. Moreover, in most cases flanking sequences from both sides of the anchor domain will be amplified, which in turn doubles the risk to catch repetitive DNA again.

As a simple alternative to inverse PCR, a linker PCR procedure was established that amplifies either the region upstream or downstream of a primer binding site common to all gene family members, using the MADS-box as an anchor. This procedure could be routinely used for the preparation of gene-specific hybridization probes. In contrast to inverse PCR, only one primer binding site is critical, and only one flanking region is amplified, thus providing two independent chances for the preparation of specific probes. One of the most obvious advantages of the linker PCR procedure is that only a minimal amount of sequence information is required, namely a primer binding site where even some mismatches may be allowed, which makes the method quite convenient and independent of detailed sequence information.

The procedure should prove useful also for other organisms and gene families, mainly for two reasons. First, the presence of large amounts of repetitive DNA is typical for many eukaryotic genomes (30), so that similar problems during the preparation of gene-specific probes may occur. Secondly, many other gene families are defined by a highly conserved sequence element, which could be used as a family-specific primer binding site (31).

The distribution of MADS-box genes in the genome of maize: functional and evolutionary implications

In dicotyledonous plants there is evidence that MADS-box genes interact with each other. Given the high structural similarity between some MADS-box genes of monocots and dicots (23,28), it seems likely that this is also the case in monocotyledonous plants. Well studied examples for a direct interaction of gene products are the factors determining 2. and 3. whorl flower organ identity (6,32). Furthermore, according to genetic data, some organ identity genes have additional functions as cadastral genes, i.e. they demarcate the expression of other organ identity genes (11,33). Thus it becomes increasingly clear that many of the genes involved in flower development mutually regulate each other at the transcriptional level, though not necessarily directly. One may assume, therefore, that they constitute a gene 'network' as defined by Wagner (34). If so, their genomic distribution becomes of evolutionary importance.

Gene duplication events were probably very important in the evolution of the MADS-box gene family (23,35). According to mathematical model calculations (34), evolution of gene networks should preferentially occur either by duplication of single genes or by duplicating all genes involved in a network. Therefore, tight linkage ('clustering') or strong dispersal are the two evolutionary most favorable forms of genomic organization of genes forming such networks. We suggest that both predicted types of organization of developmental control genes are realized in nature, with the first type represented by *Hox* genes and, as suggested by Figure 2, the second type by MADS-box genes. The striking and extremely conserved clustering of *Hox* genes and its correlation with genes' expression patterns and functions has provoked a lot of functional arguments (discussed in 14,36), but none of them is compelling, and experimental evidence is largely missing. Thus, the driving force of clustering network genes suggested recently (34) should be discussed favorably.

We interpret it as a reinforcement of the suitability of the mathematical model that also the opposite scenario which is predicted, i.e. strong dispersal of 'network' genes, can be found in nature, namely for the MADS-box genes, as outlined here. Though the genes considered in this work are only a fraction of the total number of MADS-box genes in the maize genome (our unpublished results), they belong already to at least three different subfamilies of MADS-box genes, as is evident from their MADS-domain sequence (Fig. 4). Given this diversity of genes, the strong dispersal of MADS-box genes, as evident from Figure 2, seems to be representative for maize MADS-box genes in general. This does not exclude that minor clusters will become apparent if more genes are placed on the map, as is possibly the case in regions on chromosomes 8 and 9 (see Fig. 2). However, more sequence analyses have to be carried out to exclude that one of the genes is artificially mapped there by cross-hybridization.

Strong dispersal is probably not only characteristic for maize MADS-box genes, but typical for the MADS-box genes of higher plants in general. For example, Pnueli *et al.* (38) have localized five different MADS-box genes on the five different chromosomes of the tomato genome. Similarly, eight MADS-box genes mapped in the *Arabidopsis* genome (*AG*, *AGL1*, *AGL2*, *AGL3*, *AP1*, *AP3*, *CAL* and *PI*) appeared to be located on five out of five different chromosomes (35,39,40).

It is safe to assume that flowers are essential structures for flowering plants, and that therefore the gene network regulating flower development will be under strict functional constraints. Thus strong genomic dispersal could have been established to avoid that in case of chromosomal changes a significant fraction of these genes is doubled, whereas duplication of either very few or all the genes would have only a small or no deleterious effect, respectively (34). In line with this, polyploidy is a frequent phenomenon in the plant kingdom (41), whereas aneuploids are generally not very vigorous.

Moreover, strong chromosomal dispersal of MADS-box genes could maximize physical separation of related genes in the interphase nucleus, thus reducing undesirable homologous recombination among gene family members (42).

Speculations on the nature and importance of the line-specific elements

The most unexpected observation was the existence of line-specific elements. To our knowledge, such a phenomenon has not been reported so far for MADS-box genes in other organisms. The nature and functionality of these 'gene' versions are unclear, and the possibility remains that they may be transposons or pseudogenes.

Indeed, the *TMZ1* family members studied so far have some properties that fit to transposable elements rather than to 'ordinary' MADS-box genes. First, they display an unusual primary structure. Sequencing of the MADS-box and adjacent regions of all isolated genomic clones revealed that *TMZ1* family members have a MADS-box very similar to that of *AGAMOUS* (Fig. 4), but the flanking sequences show no similarity to

AGAMOUS-like genes at all (unpublished results). Note that 'regular' MADS-box genes share considerable subfamily-specific sequence similarity also outside the MADS-box [see, e.g. (23)]. The *TMZ1* structure, therefore, suggests that a MADS-box of an *AGAMOUS*-like gene was captured by a sequence element of whatsoever identity. Precedent cases of such unusual combinations of domains in maize are provided by the recent finding that the retroelement *Bs1* has acquired a portion of a membrane proton ATPase (43), and that *Irma* is a transposon *En*-related receptor element which carries a 1.7 kb sequence that also may have originated from a gene locus (44).

Secondly, the mapping data of this work show that some *TMZ1* segments are at different places in different lines. This is the distribution expected for a transposable element, as opposed to a fixed, allelic gene locus of an 'ordinary' gene. For example, a similar line-specific genomic distribution has been reported for *Bs1* (45). Moreover, as is suggested by mapping data as well as the structure of genomic clones, the different putative segments of *TMZ1* elements occur probably in different combinations in the genome (for example with or without segment C, and segment C with or without other segments). Compositional differences are well known among maize transposable elements like *En* and *Bs1* which can incorporate genomic regions which are not present in the standard elements (see above, and references 43,44). The *Mutator* element system is especially well known for the existence of multiple subfamilies with apparently unrelated internal sequences (56). Furthermore, the maize genome is full of truncated elements and defective elements carrying internal deletions of various degrees (for reviews, see 46,47). Therefore, the observed modular line-specific distribution of *TMZ1* segments is quite diagnostic for a transposable element.

The third aspect is dispensability of at least some of the loci. The independent segregation of the line specific versions mapped with the segment B region results in individuals of the recombinant inbred population without any of them, for example, T × CM No. 23 (Fig. 3). Moreover, preliminary results (data not shown) demonstrated that one locus, detected by a probe specific for the segment C region, is present only in parental line Tx303, but absent in all individuals of the CO × Tx offspring. This suggests that the respective copy was lost during the production of the RI lines, most likely in the F_1 generation. The loss of a locus is by no means an usual process for an 'ordinary' gene, but has been observed in excessive amount in case of the maize transposable element *En*, where it is caused by transposition during chromosome replication (48).

Taken together, these data suggest that a so far unknown transposable element has captured a MADS-box of an *AGAMOUS*-like gene, and was then distributed in maize genomes and lines. Our data indicate that the segment B region is present throughout the *Andropogoneae*, whereas presence of the segment C region seems to be restricted to the genera *Zea* and *Tripsacum*. These findings are compatible with the assumption that a transposon with a segment C sequence was introduced into the genome of a common ancestor of the members of the genus *Zea*, was fused there to a pre-existing segment B sequence and then was distributed in the *Zea* genomes. However, also other scenarios are conceivable, e.g. different rates of sequence change during molecular evolution of the segment B and segment C regions.

If the *TMZ1* family members are really transposons, the intriguing question arises whether their MADS-box is functional at the protein level. Since MADS-boxes encode DNA-binding domains which bind with a certain sequence specificity to DNA regions known as CArG-boxes (10), it is conceivable that the *TMZ1* encoded MADS-domain provides DNA-binding specificity to a *TMZ1* gene product.

Alternatively, line specific versions could also have originated by reverse transcription and genomic insertion of transcripts of an 'ordinary' gene, thus they could be retropseudogenes (49). Since MADS-box genes are preferentially expressed in flowers, with the expression of some of them even being concentrated in ovules (35), cDNAs of their transcripts would have an exceptionally high chance of entering the 'germ line'. A lot of putative retroelements have been characterized in maize, among them *Cin4* (50) and *Bs1* (43,51). For *Bs1* direct evidence of transposition has been provided (51,52), thus the needed reverse transcriptase activity is clearly present in maize.

The combination of MADS-boxes with non-MADS DNA segments in at least some *TMZ1* elements might be an evolutionarily interesting scenario, and perhaps we are looking at new transcription factors in their cradle. In any case, the line-specific elements may provide excellent tools for evolutionary studies concerning the origin of certain maize races. Since their 'strange' distribution in the genomes of different maize races suggests a recent origin, they should facilitate tracing the fate of certain chromosomal segments on an evolutionary short time scale. In contrast, the line-specific elements are no suitable markers for conventional RFLP programs, since the presence of every individual is limited to specific maize lines. Therefore, they should be marked as 'line-specific' in the public domain maize maps to avoid that they are requested for this purpose.

ACKNOWLEDGEMENTS

We thank Ben Burr and Eileen Matz for assistance with RFLP mapping, Alexander Yephremov for the linkers and oligonucleotides used in this work, Alfons Gierl for the genomic maize library, Udo Wienand for advice during initial stages of the project and Uwe Schreiber for cloning λEMBL4-I120. Many thanks also to Tim Strater for providing Figure 3, and to Jorge Cacharrón for help during determination of *ZMM2* and *ZMM6* map position. We are indebted to Katharina Montag and Richard Thompson for providing probes and for critical reading of the manuscript. The technical assistance of Tüzün Akmandor and Irmgard Weißkirchen is highly acknowledged. Two of us (GT and AF) were supported by fellowships from the Deutsche Forschungsgemeinschaft (DFG).

REFERENCES

1 Gehring, W.J. (1992) *Trends Biochem. Sci.* **17**, 277–280.
2 Slack, J.M.W., Holland, P.W.H. and Graham, C.F. (1993) *Nature* **361**, 490–492.
3 McGinnis, W. and Kuziora, M. (1994) *Scientific American* **February 1994**, 36–42.
4 Sommer, H., Beltrán, J.-P., Huijser, P., Pape, H., Lönnig, W.-E., Saedler, H. and Schwarz-Sommer, Z. (1990) *EMBO J.* **9**, 605–613.
5 Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A. and Meyerowitz, E.M. (1990) *Nature* **346**, 35–39.
6 Tröbner, W., Ramirez, L., Motte, P., Hue, I., Huijser, P., Lönnig, W.-E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1992) *EMBO J.* **11**, 4693–4704.
7 Huijser, P., Klein, J., Lönnig, W.-E., Meijer, H., Saedler, H. and Sommer, H. (1992) *EMBO J.* **11**, 1239–1249.
8 Jack, T., Brockman, L.L. and Meyerowitz, E.M. (1992) *Cell* **68**, 683–697.

9 Mandel, M.A., Gustafson-Brown, C., Savidge, B. and Yanofsky, M.F. (1992) *Nature* **360**, 273–277.
10 Davies, B. and Schwarz-Sommer, Z. (1994) In Nover, L. (ed.) *Results and Problems in Cell Differentiation*. Springer-Verlag, Berlin, **Vol 20**, *Plant promoters and transcription factors*. pp. 235–258.
11 Ma, H. (1994) *Genes Dev.* **8**, 745–756.
12 Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H. and Sommer, H. (1990) *Science* **250**, 931–936.
13 Bradley, D., Carpenter, R., Sommer, H., Hartley, N. and Coen, E. (1993) *Cell* **72**, 85–95.
14 Duboule, D. and Morata, G. (1994) *Trends Genet.* **10**, 358–364.
15 Veit, B., Schmidt, R.J., Hake, S. and Yanofsky, M.F. (1993) *Plant Cell* **5**, 1205–1215.
16 Burr, B. and Burr, F.A. (1991) *Trends Genet.* **7**, 55–60.
17 Flavell, R.B., Bennett, M.D., Smith, J.B. and Smith, D.B. (1974) *Biochem. Genet.* **12**, 257–269.
18 Hake, S. and Walbot, V. (1980) *Chromosoma* **79**, 251–270.
19 Gupta, M., Shepherd, N.S., Bertram, I. and Saedler, H. (1984) *EMBO J.* **3**, 133–139.
20 Bennetzen, J.L., Schrick, K., Springer, P.S., Brown, W.E. and SanMiguel, P. (1994) *Genome* **37**, 565–576.
21 Springer, P.S., Edwards, K.J. and Bennetzen, J.L. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 863–867.
22 Mueller, P.R. and Wold, B. (1989) *Science* **246**, 780–786.
23 Theißen, G., Strater, T., Fischer, A. and Saedler, H. (1995) *Gene* **156**, 155–166.
24 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
25 Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998–9002.
26 Burr, B., Burr, F.A., Thompson, K.H., Albertson, M.C. and Stuber, C.W. (1988) *Genetics* **118**, 519–526.
27 Dellaporta, S.L., Wodd, J. and Hicks, J.B. (1983) *Plant Mol. Biol. Reporter* **1**, 19–21.
28 Schmidt, R.J., Veit, B., Mandel, M.A., Mena, M., Hake, S. and Yanofsky, M.F. (1993) *Plant Cell* **5**, 729–737.
29 Ochmann, H., Gerber, A.S. and Hartl, D.L. (1988) *Genetics* **120**, 621–623.
30 Charlesworth, B., Sniegowski, P. and Stephan, W. (1994) *Nature* **371**, 215–220.
31 Campbell, I.D. and Downing, A.K. (1994) *Trends Biotech.* **12**, 168–172.
32 Goto, K. and Meyerowitz, E.M. (1994) *Genes Dev.* **8**, 1548–1560.
33 Meyerowitz, E.M. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 5735–5737.
34 Wagner, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4387–4391.
35 Ma, H., Yanofsky, M.F. and Meyerowitz, E.M. (1991) *Genes Dev.* **5**, 484–495.
36 Kenyon, C. (1994) *Cell* **78**, 175–180.
37 Angenent, G.C., Franken, J., Busscher, M., Weiss, D. and van Tunen, A.J. (1994) *Plant J.* **5**, 33–44.
38 Pnueli, L., Abu-Abeid, M., Zamir, D., Nacken, W., Schwarz-Sommer, Z. and Lifschitz, E. (1991) *Plant J.* **1**, 255–266.
39 Hauge, B.M., Hanley, S.M., Cartinhour, S., Cherry, J.M., Goodman, H.M., Koornneef, M., Stam, P., Chang, C., Kempin, S., Medrano, L. and Meyerowitz, E.M. (1993) *Plant J.* **3**, 745–754.
40 Kempin, S.A., Savidge, B. and Yanofsky, M.F. (1995) *Science* **267**, 522–525.
41 Hilu, K.W. (1993) *Am. J. Botany* **80**, 1494–1499.
42 Matsuo, K., Clay, O., Künzler, P., Georgiev, O., Urbanek, P. and Schaffner, W. (1994) *Biol. Chem. Hoppe-Seyler* **375**, 675–683.
43 Jin, Y.-K. and Bennetzen, J.L. (1994) *Plant Cell* **6**, 1177–1186.
44 Muszynski, M.G., Gierl, A. and Peterson, P.A. (1993) *Mol. Gen. Genet.* **237**, 105–112.
45 Fuerstenberg, S.I. and Johns, M.A. (1990) *Theor. Appl. Genet.* **80**, 680–686.
46 Fosket, D.E. (1994) *Plant Growth and Development: A Molecular Approach*. Academic Press, San Diego.
47 Gierl, A., Saedler, H. and Peterson, P.A. (1989) *Annu. Rev. Genet.* **23**, 71–85.
48 Dash, S. and Peterson, P.A. (1994) *Genetics* **136**, 653–671.
49 Weiner, A.M., Deininger, P.L. and Efstratiadis (1986) *Annu. Rev. Biochem.* **55**, 631–661.
50 Schwarz-Sommer, Z., Leclercq, L., Göbel, E. and Saedler, H. (1987) *EMBO J.* **6**, 3873–3880.
51 Johns, M.A., Babcock, M.S., Fuerstenberg, S.M., Fuerstenberg, S.I., Freeling, M. and Simpson, R.B. (1989) *Plant Mol. Biol.* **12**, 633–642.
52 Mottinger, J.P., Johns, M.A. and Freeling, M. (1984) *Mol. Gen. Genet.* **195**, 367–369.
53 Fischer, A., Saedler, H. and Theißen, G. (1995) *Proc. Natl. Acad. Sci. USA* **92**, in press.
54 Chao, S., Baysdorfer, C., Heredia-Diaz, O., Musket, T., Xu, G. and Coe, E.H. (1994) *Theor. Appl. Genet.* **88**, 717–721.
55 Shen, B., Carneiro, N., Torres-Jerez, I., Stevenson, B., McCreery, T., Helentjaris, T., Baysdorfer, C., Almira, E., Ferl., R.J., Habben, J.E. and Larkins, B. (1994) *Plant Mol. Biol.* **26**, 1085–1101.
56 Bennetzen, J.L., Springer, P.S., Cresse, A.D. and Hendrickx, M. (1993) *Critical Rev. Plant Sci.* **12**, 57–95.