Molecular Characterization of the *Abp1* 5′-Flanking Region in Maize and the Teosintes[1]

Nabil Elrouby and Thomas E. Bureau*

Department of Biology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada H3A 1B1

Auxin-binding protein 1 subsp. *mays* (ABP1) has been suggested as a receptor mediating auxin-induced cell expansion and differentiation. In maize (*Zea mays*), ABP1 is encoded by a single gene, *Abp1*. The TATA and CAAT promoter elements as well as the transcriptional start site were previously identified and all were found to be located within a transposable element (TE), *Tourist-Zm11*. In this study we report the cloning and characterization of the *Abp1* 5′-flanking region in maize and its wild relatives, the teosintes. We provide evidence for insertion polymorphism corresponding to *Tourist-Zm11* and two novel TEs, *Batuta* and *Pilgrim*. Despite this polymorphic structure, the *Abp1* core promoter in maize and the teosintes is conserved, is located downstream of the TE insertions in the 5′-flanking region, and is TATA-less. We discuss the potential evolutionary impact of these TEs on the regulation of *Abp1* gene expression.

Auxins are phytohormones that regulate various aspects of plant growth and development including elongation growth, photo- and gravitropism, apical dominance, lateral root initiation, the differentiation of vascular tissues, embryogenesis, and fruit ripening (Macdonald, 1997). Although the mechanism by which auxins are perceived by the cell is still unclear, several auxin-binding proteins (ABPs) have been identified and are thought to play a role in auxin perception (Jones and Prasard, 1992). Of these, ABP1 has been implicated as an auxin receptor because it binds the most active auxins in vitro (Ray et al., 1977; Lobler and Klambt, 1985). However, studies that showed that ABP1 is localized predominantly to the endoplasmic reticulum and to a much lesser extent to the cell membrane were puzzling since it is expected that ABP1 binds auxins at the cell membrane (Lazarus et al., 1991; Napier, 1997). Two main lines of evidence subsequently established ABP1 as an auxin receptor. First, ABP1 was found to bind auxins at the cell membrane and not the endoplasmic reticulum despite the latter being its predominant location (Barbier-Brygoo et al., 1989, 1991; Tian et al., 1995). Second, both transgenic tobacco plants and maize (*Zea mays* subsp. *mays*) cell lines overexpressing *Abp1* displayed an increased capacity for auxin-induced cell expansion (Jones et al., 1998). A model was suggested in which ABP1 is secreted to the outer surface of the cell membrane through its association with a membrane-spanning docking protein, possibly a G-protein-coupled receptor (Macdonald, 1997). Auxin binding at the cell membrane would induce a conformational change in ABP1 that activates the auxin signal transduction pathway.

Comparison of the maize (*Z. mays* subsp. *mays*) *Abp1* genomic and cDNA clones failed to reveal any TATA box motifs in the genomic sequences immediately upstream of the cDNAs considered to be full-length (Lazarus et al., 1991). Initial attempts to determine the transcriptional start site (+1) yielded inconsistent results (Lazarus et al., 1991). However, the *Abp1* +1 was mapped to the CC A CT at 320 bp upstream of the start of translation (ATG) by consensus sequence analysis and primer extension (Schwob et al., 1993). Although this +1 is located 45 bp downstream from a consensus TATA motif, the predicted transcript is much longer than the mRNA detected by northern analysis (Inohara et al., 1989) and the longest cDNA sequenced (Hesse et al., 1989). The TATA and CAAT box motifs as well as the +1 were reported to be located within a transposable element (TE), *Tourist-Zm11*, inserted 299 bp upstream of the ATG (Bureau and Wessler, 1992; Bureau et al., 1996). *Tourist-Zm11* was, thus, suggested to contribute the *Abp1* core promoter sequences.

Tourist-Zm11 belongs to a novel superfamily of TEs called miniature inverted-repeat TEs (MITEs). MITEs are characterized by their small size, presence of conserved terminal inverted repeats (TIRs), and a target site preference (Bureau and Wessler, 1992; Bureau et al., 1996). MITEs and MITE-like sequences are frequently associated with the non-coding regions of normal (wild-type) plant genes (Bureau and Wessler, 1992, 1994a, 1994b; Bureau et al., 1996; Casacuberta et al., 1998; Charrier et al., 1999; Surzycki and Belknap, 1999) and are also present in non-plant systems including the mosquito (*Aedes aegypti*; Tu, 1997), human (Smit and Riggs, 1996), and teleost fish (Izsvak et al., 1999). In addition to MITEs, remnants of *copia*-like retrotransposons have been shown to flank many normal plant genes (White et al., 1994). In contrast, short and long interspersed nuclear elements are the predominant TEs in association with mammalian

[1] This work is funded by a research grant from the National Science and Engineering Research Council (grant no. 183960–96).

* Corresponding author; e-mail Thomas_Bureau@maclan.mcgill.ca; fax 514–398–5069.

genes (Eickbush, 1992). Although some TE sequences have been demonstrated to contribute to the control of expression of nearby genes (McDonald, 1993), their role in the evolution of normal gene expression is not well understood.

To determine the impact of *Tourist-Zm11*, we characterized the *Abp1* 5′-flanking region in maize and its wild relatives, the teosintes. We show that this region is highly polymorphic due to the insertion of several TEs, and we discuss their significance in the regulation of *Abp1* gene expression.

RESULTS

The *Abp1* 5′-Flanking Region Contains Multiple TE Insertions

Tourist-Zm11 was first identified in the 5′-flanking region of maize *Abp1* by sequence similarity searches (Bureau and Wessler, 1992). Database searches using the published maize *Abp1* sequence (GenBank accession no. L08425) as a query also revealed that the 5′ upstream-most region (870–1240 bp upstream of the ATG) shares similarity with the *Ds1* element insertion of the maize *Bz1* gene (Schiefelbein et al., 1988; EMBL accession no. X14155) and with a *Ds1* insertion in *Zea perennis* (MacRae and Clegg, 1992; EMBL accession no. X54711). Although the 3′ TIR could be recognized (5′-ATCCATCCCTA-3′), the L08425 sequence did not extend far enough upstream to include the 5′ TIR (Fig. 1).

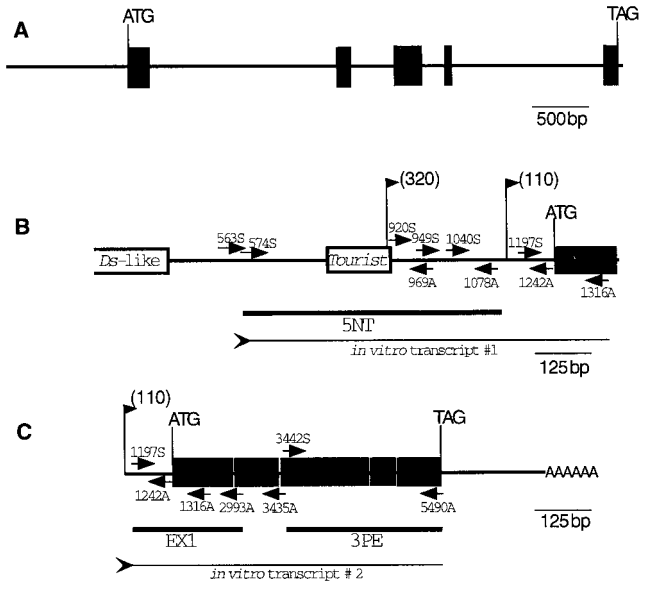

Figure 1. Genomic (A), 5′-flanking region (B), and transcript (C) organization of maize (*Z. mays* subsp. *mays*) *Abp1*. Exons and TEs are represented by shaded and empty rectangles, respectively. Arrows represent primers. Primer names are derived from their positions on the genomic sequence (accession no. L08425). The thick black lines represent the probes used in this study and the thin lines starting with arrowheads represent the two in vitro transcripts used in Figure 5. The flags labeled 320 and 110 represent the transcriptional start sites determined by Schwob et al. (1993) and in this report and located 320 and 110 nt upstream of the ATG, respectively.

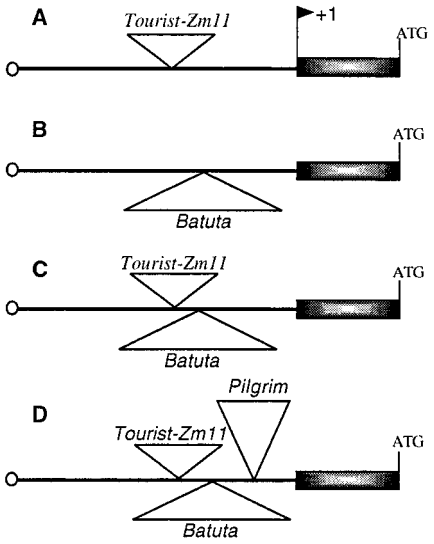

Figure 2. Schematic depiction of the four types of the *Abp1* 5′-flanking sequences. The first, second, third, and fourth types (as described in the text) are represented by A, B, C, and D, respectively. The shaded rectangles represent the 5′-untranslated region (UTR), and the TE insertions are represented by triangles. The transcriptional start site is represented by a flag. All sequences were amplified using primer 574S or 563S (represented by the circle) and 1316A.

To search for insertion polymorphism and to further characterize the *Abp1* upstream sequences, the *Abp1* 5′-flanking regions in maize (cv W22) and the teosintes (*Z. mays* subsp. *parviglumis*, *Z. mays* subsp. *heuheutenangenesis*, *Z. mays* subsp. *mexicana*, *Zea diploperennis*, and *Zea luxurians*) were amplified using primers ZMAUX563S or ZMAUX574S and ZMAUX1316A shown in Figure 1. Nucleotide sequence (GenBank accession nos. AF292696–AF292709) comparison revealed the presence of several TE insertions. Based on their TE content, the *Abp1* 5′-flanking sequences can be grouped into four main types as represented in Figure 2 and Table I. The first group (type A) is represented by the only maize sequence amplified (ZmW22) and one sequence amplified from *Z. mays* subsp. *huehuetenangenesis* (ZmH3). Both sequences contain only one TE insertion, *Tourist-Zm11*. The second group (type B) of *Abp1* sequences lacks *Tourist-Zm11* and instead contains a different insertion located 21 bp downstream of the position where *Tourist-Zm11* has inserted in ZmW22 and ZmH3. This new element was called *Batuta* (after Ibn Batuta, a traveler and author), is flanked with a 5-bp target site duplication (TSD, consensus; 5′-TTCTT-3′), and is 221 to 222 bp in length. Members of this group of sequences include ZmP1, ZmH1, ZmM7, ZL1, ZL5, ZD1, and ZD2. ZL1 and ZL5, and ZD1 and ZD2 are allelic sequences that share 91.8% and 96.7% nucleotide identity, respectively. The third group (type C) of *Abp1* sequences contains both *Tourist-Zm11* and *Batuta*. This latter *Batuta* insertion is also located 21

Table I. *TE insertions in the maize and teosinte Abp1 5′-flanking regions*

Sequence Name	TE Content			Type[a]
	Tourist-Zm11	*Batuta*	*Pilgrim*	
W22[b]				
ZmW22	√[c]			A
PARV				
ZmP1		√		B
ZmP3	√	√		C
HUE				
ZmH1		√		B
ZmH3	√			A
ZmH8	√	√		C
MEX				
ZmM1	√	√	√	D
ZmM6	√	√		C
ZmM7		√		B
LUX				
ZL1		√		B
ZL4	√	√		C
ZL5		√		B
DIP				
ZD1		√		B
ZD2		√		B

[a] *Abp1* 5′-flanking types as depicted in Fig. 2. [b] The *Zea taxa* are *Z. mays* subsp. *mays* (W22), subsp. *parviglumis* (PARV), subsp. *huehuetenangenesis* (HUE), subsp. *mexicana* (MEX), *Z. luxurians* (LUX), and *Z. diploperennis* (DIP). [c] √ indicates the presence of the TE.

bp downstream of the *Tourist-Zm11* target site, has a 5-bp TSD (5′-TTCTT-3′) but is 12 to 13 bp shorter than its counterpart of the second group. *Abp1* sequences of the third group include ZmP3, ZmH8, ZmM6, and ZL4. The fourth group (type D) of *Abp1* sequences is represented by a sequence obtained from *Z. mays* subsp. *mexicana*, ZmM1. This sequence contains three TE insertions; *Tourist-Zm11*, *Batuta*, and a third insertion designated *Pilgrim*. *Pilgrim* is located 109 bp downstream of *Batuta*, has an 11-bp TSD (5′-CTGCGTGGTGC-3′), and is 165 bp in length.

The Maize and Teosinte *Abp1* Transcripts Are Approximately the Same Size

Northern-blot hybridization was performed using a probe spanning the translated region of exon I and a part of the 5′-UTR (EX1; Fig. 1) and RNA extracted from maize (cv W22 and cv Seneca Horizon) and the teosintes. Figure 3 shows that only one *Abp1* mRNA species of about 1 kb in size was observed in all samples tested. This suggested that the *Abp1* transcripts of maize and the teosintes may have the same 5′-UTRs regardless of the TE insertions at their 5′-flanking regions. They may alternatively have different 5′-UTRs and the differences in sizes are compensated for by differences at their 3′-UTRs. The maize *Abp1* transcript was reported to start at 320 bp upstream of the ATG (Schwob et al., 1993) and to contain an open reading frame of 603 nucleotides (nt; Hesse et al., 1989; Inohara et al., 1989; Tillmann et al., 1989). In addition, our 3′ RACE results (data not shown), and primer-directed RNaseH digestion of maize and teosinte *Abp1* mRNA (Fig. 4) revealed that the size of the 3′-UTR including the poly(A) tail is approximately 300 nt. Taken together, the predicted transcript size should be approximately 1,225 nt. Although the transcript size reported here (approximately 1.0 kb) is in agreement with (approximately 1.0 kb in Inohara et al., 1989) or larger than (approximately 0.8 kb in Hesse et al., 1989; approximately 850 nt in Tillmann et al., 1989) that previously re-

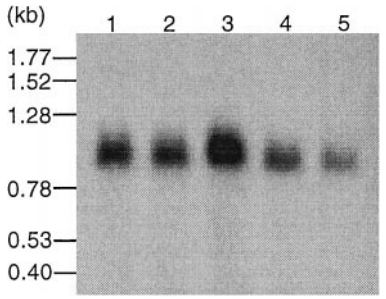

Figure 3. Northern-blot analysis of *Abp1*. Total RNA from *Z. mays* subsp. *mays* cv Seneca Horizon (lane 1), *Z. mays* subsp. *mays* cv W22 (lane 2), *Z. mays* subsp. *parviglumis* (lane 3), *Z. mays* subsp. *mexicana* (lane 4), and *Z. diploperennis* (lane 5) was separated on formaldehyde agarose gels, blotted to nylon membranes, and probed with EX1. RNA size standards (in kb) are given on the left-hand side.

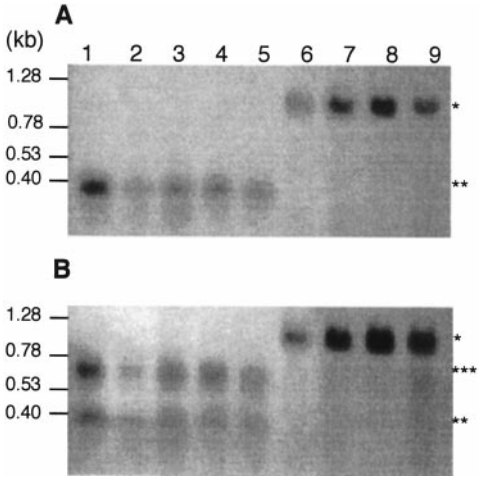

Figure 4. Primer-directed RNaseH digestion of *Abp1* mRNA. Total RNA was incubated with RNaseH and ZMAUX3435A (lanes 1–5), ZMAUX3435A only (lane 8), or RNaseH only (lane 9), or was untreated (lanes 6 and 7), followed by northern analysis. The membrane was probed with EX1 and subjected to autoradiography (A), and then probed with 3PE and re-exposed (B). Lane 1, *Z. mays* subsp. *mays* cv Seneca Horizon; lanes 2 and 6, *Z. mays* subsp. *parviglumis*; lanes 3 and 7, *Z. mays* subsp. *mexicana*; lanes 4 and 8, *Z. diploperennis*; lanes 5 and 9, *Z. luxurians*. The asterisks indicate the positions of the intact *Abp1* transcript (*), the 5′ fragment (**), and the 3′ fragment (***).

ported, it does not correspond to the predicted size of approximately 1,225 nt.

The *Abp1* Transcript 5′ End Is Conserved in Maize and the Teosintes

To resolve the structure of the 5′-UTR and the discrepancy in mRNA size, we adopted a number of approaches to map the 5′ end of *Abp1* mRNAs. 5′ RACE was performed with maize and teosinte mRNA as templates and using primers that spanned the junction of exons I and II (ZMAUX2993A) and the junction of exons II and III (ZMAUX3435A) for first-strand cDNA synthesis (Fig. 1). The longest 5′ RACE products terminated at 110 nt upstream of the ATG. Since Schwob et al. (1993) mapped the maize *Abp1* mRNA 5′ end to 320 nt upstream of the ATG, we attempted to amplify maize sequences upstream of nt number 110 using reverse transcriptase (RT)-PCR and different combinations of primers ZMAUX-920S, ZMAUX949S, and ZMAUX1040S and primers ZMAUX1078A, ZMAUX1242A, ZMAUX1316A, ZMAUX2993A, and ZMAUX3435A (see Fig. 1 and "Materials and Methods"). Despite using several different RTs (AMV, M-MLV, and *rTth*), no amplification products were obtained (data not shown).

The same primers were used successfully to amplify the corresponding maize genomic sequences. This suggested that the transcript 5′ end for maize and probably the teosintes may actually be located further downstream from that previously reported for maize *Abp1* (Schwob et al., 1993). RT-PCR using primer ZMAUX1197S, which is located 44 nt upstream of the ATG, and any of primers ZMAUX-1316A, ZMAUX2993A, or ZMAUX3435A yielded the expected amplification products (data not shown). In addition, RT-PCR was performed using mRNA purified from seven different tissues and developmental stages (young seedling leaf, 10-week-old leaf, immature tassel, post-pollen tassel, immature ear, mature ear, and fertilized ear). Whereas the expected amplification products were obtained when primers ZMAUX1197S and ZMAUX2993A were used, amplification using primers ZMAUX949S and ZMAUX-2993A yielded no products.

To get a precise measure of the size of the *Abp1* mRNA in maize and the teosintes, we performed primer-directed RNaseH digestion of total RNA followed by northern-hybridization analysis. Figure 4 shows that the size of the 5′ fragment of the mRNA in both maize and the teosintes is approximately 330 nt, which puts the transcript 5′ end at approximately 110 nt upstream of the ATG. The sum of the sizes of the 5′ (approximately 330 nt) and 3′ (approximately 650 nt) fragments, and primer ZMAUX3435A (25 nt) yields an *Abp1* mRNA of approximately 1 kb, which is consistent with our (Fig. 3) and previous (Inohara et al., 1989) northern results.

Northern hybridization was performed on maize total RNA using three different probes; EX1, 3′ end (3PE, which spans exons III–V), and 5′-non-transcribed region (5NT, which covers sequences upstream of nucleotide number 110 from the ATG) (Fig. 1 and "Materials and Methods"). According to Schwob et al. (1993) all three probes should hybridize to *Abp1* mRNA. As seen in Figure 5A, although both the EX1 and 3PE probes gave the expected 1-kb signal, the 5NT probe failed to recognize any *Abp1* RNA. The 5NT probe hybridizes specifically, however, to an in vitro transcript covering the ZmW22 sequence but not to that covering full-length cDNA (Figs. 1 and 5B).

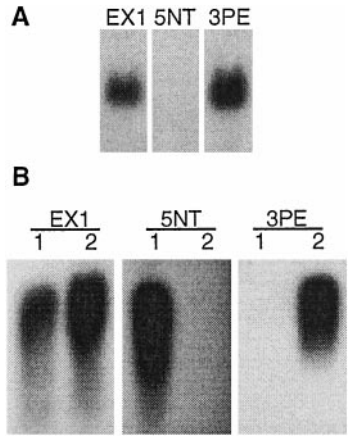

Figure 5. Northern-blot analysis of maize total RNA (A), and in vitro-transcribed maize *Abp1* sequences (B, also see Fig. 1). A, Ten micrograms of *Z. mays* subsp. *mays* total RNA was separated on three replica lanes, blotted onto a nylon membrane, and the membrane was cut into three strips. Each strip was then hybridized with the probe indicated. B, Ten nanograms of in vitro transcripts of ZmW22 (lanes 1) and full-length *Abp1* cDNA (lanes 2) was separated and probed with the same probes used in A.

Primer extension was performed to confirm our 5′ RACE results. As seen in Figure 6, primer ZMAUX-1242A gave rise to two extension products that put the 5′ end of *Abp1* mRNA at the A and G residues located at positions 110 and 109 upstream of the ATG. Primers ZMAUX1078A and ZMAUX969A, on the other hand, did not yield extension products (data not shown). Primer extension with cv Seneca Horizon RNA also yielded two extension products terminating at the T and C residues located at positions 108 and 107 upstream of the ATG, respectively (data not shown).

DISCUSSION

TE Insertions within the *Abp1* 5′-Flanking Region of Maize and the Teosintes

Sequence comparison between the 5′-flanking region of maize and teosinte *Abp1* genes reveals the presence of two novel TE insertions, *Batuta* and *Pilgrim*. Computer-assisted database searches (in September 1999) using either the *Batuta* or *Pilgrim* sequences as queries did not reveal similarity to known TEs. However, several lines of evidence sug-

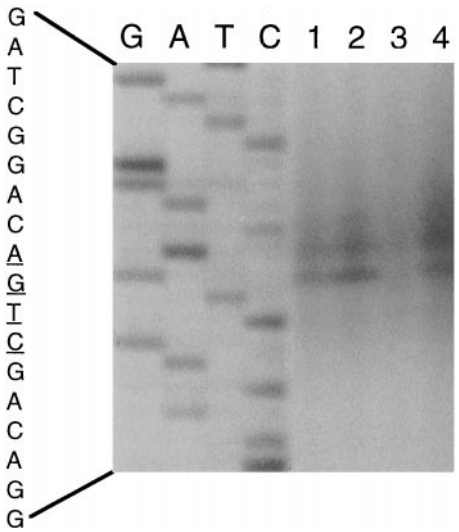

Figure 6. Primer extension mapping of the *Abp1* transcript 5′ end. ^{32}P-Labeled primer ZMAUX1242A was used in extension reactions with 20 μg of total RNA extracted from *Z. diploperennis* (lane 1), *Z. mays* subsp. *mexicana* (lane 2), and *Z. mays* subsp. *mays* cv W22 (lane 4). Lane 3 is blank. Extension products were separated on 8% (w/v) polyacrylamide/urea gels alongside with ^{35}S-labeled sequencing reactions primed with ZMAUX1242A. The sequence of the region flanking the +1 is shown and residues 107, 108, 109, and 110 upstream of the ATG are underlined.

gest that both insertions are actually TEs. Insertion polymorphisms were identified for both *Batuta* and *Pilgrim* in the different teosintes. Furthermore, *Pilgrim* contains TIRs with sequence similarity to those of *Ac*-like TEs (Kunze, 1996) and has an 11-bp TSD. *Batuta,* on the other hand, appears to be a solo long terminal repeat of a retroelement (LTR). Solo LTRs are presumably generated by homologous recombination of the two LTRs of an ancestral retroelement (Mager and Goodchild, 1989). Like retroelement LTRs (Bingham and Zachar, 1989), *Batuta* has a 5-bp TSD and contains potential promoter (e.g. multiple TATA-, CAAT-, and GC-like sequences), polyadenylation, and enhancer sequences. The existence of LTR fragments within or near genes has been documented in several eukaryotes (White et al., 1994; Noma et al., 1997; Goodchild and Poulter, 1998; Liao et al., 1998).

Abp1 Transcriptional Initiation Is Conserved among Maize and the Teosintes

The *Abp1* core promoter was previously reported to be contributed by sequences present in *Tourist-Zm11*. This is difficult to reconcile with our results for two reasons. First, some teosinte *Abp1* 5′-flanking types (Fig. 2, type B) lack *Tourist-Zm11*. In this case, the core promoter would have to be provided by other sequences. Second, the *Abp1* 5′-flanking regions of some teosintes contain not only *Tourist-Zm11*, but also *Batuta* (Fig. 2, type C) or *Batuta* and *Pilgrim* (Fig. 2, type D) further downstream. If the core promoter is provided by *Tourist-Zm11*, the resulting transcripts will contain unusually long 5′-UTRs.

Although the maize *Abp1* +1 was previously mapped to 320 nt upstream of the ATG, we provide strong evidence that the *Abp1* +1 for both maize and the teosintes is in fact located 110 nt upstream of the ATG. First, we were unable to amplify, by RT-PCR, sequences upstream of the ATG (the 5′-UTR) using primers positioned more than 110 nt upstream of the ATG. However, amplification products were obtained using primers located within 110 nt upstream of the ATG. Second, we investigated the possibility that *Abp1* may be differentially transcribed from two promoters; one that generates the 1-kb transcript that starts 110 nt upstream of the ATG, whereas the other generates a transcript that starts 320 nt upstream of the ATG. Northern-blot analysis using a full-length *Abp1* cDNA probe and total RNA extracted from 18 different tissues and developmental stages did not reveal the presence of a transcript originating 320 nt upstream of the ATG (data not shown). Likewise, RT-PCR products from seven of the RNA samples only revealed the transcript predicted to originate 110 nt upstream of the ATG. Third, primer-directed RNaseH digestion of *Abp1* mRNA followed by northern hybridization indicated that the 5′ end of the transcript is located approximately 110 nt upstream of the ATG. Fourth, using a probe upstream of 110 nt from the ATG (5NT), we were unable to detect any hybridization signal on northern blots. However, a band of the expected size corresponding to a transcript initiating 110 nt upstream of the ATG was detected using the EX1 probe. Fifth, using two methods, 5′ RACE and primer extension, to determine the precise position of the *Abp1* transcript 5′ end in maize and the teosintes, we found that the *Abp1* transcript starts at two positions located at 109 and 110 nt upstream of the ATG. When a primer (ZMAUX969A) located 28 nt downstream from what was previously reported to be the *Abp1* +1 (i.e. 320 nt upstream of the ATG; Schwob et al., 1993), no extension products were obtained. Our mapping of +1 to 110 bp upstream of the ATG fits well with previous (Inohara et al., 1989) and our determination of transcript size.

We also show that the *Abp1* +1 is conserved between maize and the teosintes despite the structural diversity (conferred mostly by different TE insertions) of their 5′-flanking regions. The sequence upstream of the ATG up to the position of the first insertion is highly conserved between maize and the teosintes, including the region immediately upstream of +1 (approximately 170 bp; >95% identity). In addition, although *Tourist-Zm11* and *Batuta* both contain putative TATA box motifs, these motifs are located much further upstream (> −180) than the canonical position of approximately −25 to −40 for a typical TATA box. Therefore, we propose that the sequences immediately upstream of +1 provide the *Abp1* core promoter. This proposal has two implica-

tions. First, the TE insertions in the *Abp1* upstream region do not contribute to the *Abp1* core promoter sequence. Second, since the region from +1 up to the first TE insertion lacks any putative TATA box motifs, this suggests that *Abp1* transcription may be derived by a TATA-less promoter.

Transcription from TATA-less promoters depends on the presence of an initiator element (Inr). The Inr is a strong core promoter element that is functionally analogous to the TATA box (Smale, 1997). The Inr extends from nucleotide position −6 to +11 with the consensus Py Py A N T/A Py Py (underlined residue = +1). Although the Py residues are not needed in all four positions, they must be present in at least a few positions. The most critical residues are central CA, A at +1, T or A at +3, and Py at −1 (Smale, 1997). As seen in Figure 7, the sequence flanking the maize and teosinte *Abp1* +1 (110 nt from the ATG) is GGA C*A* GTC, which conforms to the Inr consensus. Sequence conservation at this region between maize and the teosintes provides additional support that it functions as an Inr (Fig. 7). Many TATA-less promoters are also characterized by a GC-rich 5′-UTR and 5′-flanking region that contains binding sites for transcription factors such as Sp1 and AP (Dufau et al., 1995; Liu and Hla, 1997; Smale, 1997). The *Abp1* 5′-UTR and 5′-flanking region (−1 to −100) are similarly GC-rich (GC content: 60% and 66%, respectively). In summary, the *Abp1* core promoter is conserved between maize and the teosintes, is TATA-less, and is located downstream of the TE insertions.

Potential Significance of the *Abp1* 5′-Flanking TE Insertions in *Abp1* Gene Expression

The finding that the *Abp1* core promoter sequence is not provided by the TE insertions in the *Abp1* 5′-flanking regions raises questions about whether they play any role in *Abp1* expression. Schwob et al. (1993) investigated the ability of different *Abp1* promoter deletions to promote transient chloramphenicol acetyl transferase (CAT) reporter gene expression in maize leaf protoplasts. They describe a region from 684 to 449 bp upstream of the ATG that contains a negative regulatory element. When the sequence from 449 to 256 bp upstream of the ATG was deleted, expression declined almost 4-fold. It is interesting that this region contains the entire *Tourist-Zm11* element plus an additional 66 bp. In addition, the sequence from 256 to 17 bp upstream of the ATG (which we suggest to contribute the core promoter) promotes expression levels higher than those obtained from full-length *Abp4* and *Abp5* promoters (other members of the ABP gene family; Lazarus et al., 1991; Schwob et al., 1993).

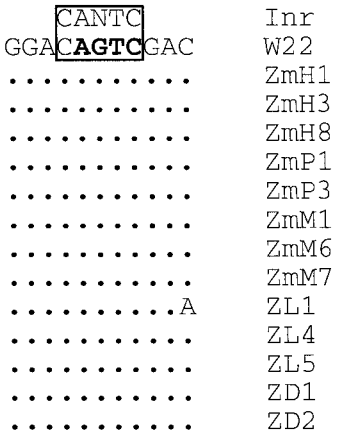

Figure 7. Alignment of the consensus Inr and *Abp1* sequences flanking +1. The critical Inr residues are enclosed in the box and the four *Abp1* transcriptional start sites are in bold.

In light of our mapping of the *Abp1* transcript 5′ end and the presence of a TATA-less promoter, we suggest that *Tourist-Zm11* may act as a positive regulatory element and that the sequence 684 to 449 bp upstream of the ATG may contain a negative regulatory element. *Abp1* minimal expression would be conferred by the core promoter, and different expression levels and possibly inducible, tissue-specific, and/or temporal expression would be the net action of both the positive and negative regulatory elements. Furthermore, eukaryotic promoter database (Cavin Périer et al., 1999) searches using the *Abp1* 5′-flanking regions as queries revealed that *Tourist-Zm11*, *Batuta*, and *Pilgrim* all contain several putative cis-acting regulatory sequences (data not shown). This opens the possibility that differential expression may be determined by cis-acting regulatory elements contributed by these TEs. In view of the fact that *Abp1* is differentially expressed in coleoptiles and female reproductive organs (Hesse et al., 1993; N. Elrouby and T.E. Bureau, unpublished data), the presence of a TATA-less core promoter may further indicate a possible function of the upstream cis-acting regulatory sequences. Many TATA-less promoters have been shown to confer a greater potential for differential expression (e.g. Dufau et al., 1995; Liu and Hla, 1997) than the tightly controlled TATA-containing promoters.

In general, TEs may influence the evolution of wild-type genes and gene expression in a variety of mechanisms (for review, see McDonald, 1995; Britten, 1996, 1997; Kidwell and Lisch, 1997). TE insertions may induce changes in chromatin structure that may, for example, insulate a gene's promoter from enhancer sequences located distal to the site of insertion (Gerasimova et al., 1995). TEs inserted in the 5′-flanking regions of genes may alternatively influence gene expression by altering transcriptional initiation in a temporal or tissue-specific manner. This change may result from the generation of a novel transcript initiated from the TE promoter. For example, a retrotransposon insertion in the upstream region of the maize 19-kD zein genes results in the generation of an additional zein transcript initiated

from the TE promoter (Kriz et al., 1987; White et al., 1994). TEs may also influence normal gene expression by contributing positive or negative regulatory sequences (McDonald, 1995). For instance, an enhancer sequence contained within a cryptic retrotransposon located in the 5′-flanking region of the murine *slp* gene confers androgen-specific regulation (Stavenhagen and Robins, 1988). Likewise, a retroelement insertion in the 5′-flanking region of the human *Amy1* gene confers salivary gland-specific expression (Meisler and Ting, 1993). In addition, an endogenous retroviral LTR inserted in intron 19 of the human leptin receptor gene *OBR* induces steroid-mediated alternative splicing of its mRNA (Kapitonov and Jurka, 1999). Selection on short sequences within TEs associated with the regulatory regions of normal genes may provide a mechanism for the evolution of eukaryotic enhancers (McDonald et al., 1997). Whether TEs, namely *Tourist*, *Batuta*, and *Pilgrim*, are involved in the expression or regulation of the maize and/or teosinte *Abp1* genes will require fine-scale promoter characterization. This type of study will ultimately help determine if TEs are for the most part "junk DNA" or, alternatively, are fundamental generators of phenotypic variation.

MATERIALS AND METHODS

Plant Material and Nucleic Acids Extraction

Maize (*Zea mays* subsp. *mays* cv W22 and cv Seneca Horizon) germplasms were obtained from Susan Wessler (University of Georgia, Athens) and Stokes Seeds (Ontario, Canada), respectively. The teosinte germplasm was obtained from John Doebley (University of Minnesota, St. Paul). For genomic DNA isolation, seedlings were grown under greenhouse conditions for 3 weeks and genomic DNA was extracted as described (Dellaporta et al., 1985). Total RNA was isolated from 7- to 10-d-etiolated seedlings following a LiCl precipitation method (Ausubel et al., 1996). Poly(A^+) RNA was purified using the PolyAtract mRNA Isolation System III kit (Promega, Madison, WI).

PCR Amplification, RT-PCR, and 5′ RACE

PCR amplifications were performed in 10 mM Tris [tris(hydroxymethyl)aminomethane]-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% (w/v) gelatin, dNTPs (200 μM each), 35 ng of each primer, 75 ng of genomic DNA, and 1 unit of AmpliTaq Gold DNA polymerase (Perkin Elmer, Foster City, CA). PCR conditions were as described (Bureau and Wessler, 1994b) except that the annealing step was performed for 2 min at 65°C. First-strand cDNA synthesis for RT-PCR and 5′ RACE was performed using AMV RT (Pharmacia Biotech, Piscataway, NJ) and 20 to 30 ng of poly(A^+) RNA. We used the *Abp1* genomic sequence (Schwob et al., 1993, accession no. L08425) to design the primers used in this study. The primer name consists of two parts; ZMAUX for *Z. mays* ABP 1, and a number that corresponds to the position of the first nucleotide of the primer sequence in the L08425 sequence followed by the suffix "S" for the sense sequence or "A" for antisense sequence. The primers used to amplify the maize and teosinte 5′-flanking region were ZMAUX574S (5′-AGTAAGACCGCCCCAACAGG-3′) or ZMAUX563S (5′-CATATTTTCTTAGTAAGACCGCCCCAACAGGTTACCC-3′) and ZMAUX1316A (5′-GAAGGAGGCAGCGAGGAGGAC-3′). The primers used for RT-PCR and 5′ RACE were ZMAUX920S (5′-ACTGGAACCGAACAAGCCC-3′), ZMAUX949S (5′-GCCGGTTGGATTCTT-3′), ZMAUX1040S (5′-CTGGCAAGGCGAGGCTTGTCCTGCGTG-3′), ZMAUX1078A (5′-GCCCAGAGACAAGCCGCACC-3′), ZMAUX1197S (5′-CTCATTCCACTCCGACATTC-3′), ZMAUX1242A (5′-ATTGCCTGCTCCCGACAGCTGC-3′), ZMAUX1316A, ZMAUX2993A (5′-CTCTCACCAATGAGTTATCTCGCAC-3′), ZMAUX3435A (5′-GCCACACTTCCACCTCCTTCATC-3′), ZMAUX3442S (5′-CAATAAGTCCAGGTCAAAGGACGCCAATC-3′), and ZMAUX5490A (5′-GGAAACACTTGTGACCTAGAG-3′).

5′ RACE was performed as described (Troutt et al., 1992). For this purpose, first-strand cDNA was ligated to an anchor oligonucleotide (5′-TTTAGTGAGGGTTAATAAGCGGCCGCGTCGTGACTGGGAGCGC-3′) using T4 RNA ligase (Epicentre Technologies, Madison, WI), and then amplified using a gene-specific primer (ZMAUX3435A or ZMAUX2993A) and an oligonucleotide complementary to the anchor oligonucleotide (5′-GCGGCCGCTTATTAACCCTCACTAAA-3′).

Cloning, Sequencing, and Sequence Analysis

All PCR fragments were cloned into pCR2.1 as a part of the Original or Topo TA cloning kit (Invitrogen, Carlsbad, CA). Clones were given names that describe the species or subspecies and a number that corresponds to the clone number (Table I). Sequencing was performed following the dideoxy termination method using the SequiTherm Excel II DNA sequencing kit (Epicentre Technologies). Database searches were performed using the algorithm BLAST (version 2.0; National Center for Biotechnology Information, Bethesda, MD). Sequence analysis was performed using the Gap and PileUp programs as part of the UWGCG program suite (version 9.0, University of Wisconsin Genetics Computer Group, Madison).

Probes

Abp1 exon 1 (EX1) and 3PE cDNAs were isolated by RT-PCR using the primers ZMAUX1197S and ZMAUX2993A, and ZMAUX3442S and ZMAUX5490A, respectively (Fig. 1), and cloned into pCR2.1 as described above. To generate the DNA probes used for northern hybridizations, the EX1 and 3PE cDNAs were released from pCR2.1 by *Eco*RI digestion. The 5NT probe, which covers genomic sequences from 110 to 676 bp upstream of the ATG (Fig. 1), was released from ZmW22 by *Eco*RI-*Sal*I digestion. After gel purification, all DNA fragments were radiolabeled by random priming using the Oligolabelling kit (Pharmacia Biotech).

Northern Analysis

Total RNA(10–12 μg) from each species or subspecies was separated on 1.5% (w/v) formaldehyde agarose gels, transferred to nylon membranes (GeneScreen Plus, NEN Life Science Products, Boston), and hybridized to the EX1, 3PE, or 5NT probes in a solution of 0.25 M Na_2HPO4 (pH 7.4), 1 mM EDTA, and 7% (w/v) SDS at 60°C overnight. Membranes were then washed in 0.1× SSC, 1% (w/v) SDS for 30 min at 60°C and subjected to autoradiography. To generate the run-off transcripts used as controls in Figure 5, pCR2.1 clones containing ZmW22 (in vitro transcript no. 1) or a full-length *Abp1* cDNA (in vitro transcript no. 2) were linearized with *Bam*HI and used in transcription reactions following the RiboMax RNA production system (Promega).

Primer-Directed RNaseH Digestion of *Abp1* mRNA

Site-specific enzymatic cleavage of RNA was as previously described (Donis-Keller, 1979). For this purpose, 12 μg of total RNA from each species or subspecies was incubated with 50 ng of the antisense primer ZMAUX3435A at 90°C for 5 min, chilled on ice for 5 min, and then allowed to anneal for 15 min at 42°C in 50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, and 100 mM NaCl. Five units of RNaseH (Epicentre Technologies) was added and the reaction was incubated at 37°C for 30 min. Nucleic acids were ethanol-precipitated, separated on a 1.5% (w/v) formaldehyde gel, transferred onto nylon membranes, hybridized to the EX1 probe, and subjected to autoradiography. The same blot was then hybridized to the 3PE probe and re-exposed.

Primer Extension

Primer extension was performed as described (Ausubel et al., 1996). γ-^{32}P end-labeled primers (ZMAUX1242A, ZMAUX1078A,orZMAUX969A[5′-GGCAACAAGAATCC-AACCGGC-3′]) were mixed individually with 20 μg of total RNA and extended using AMV RT (Pharmacia Biotech). The extension products were separated on 8% (w/v) polyacrylamide/urea gels alongside with ^{35}S-labeled sequencing reactions primed with the same primer used for the extension reaction.

ACKNOWLEDGMENTS

We thank Candace Waddell, Stephen Wright, and Quang Hien Le for critically reading our manuscript. We also thank John Doebley for providing the teosinte germplasm.

Received February 18, 2000; accepted May 8, 2000.

LITERATURE CITED

Ausubel F, Brent R, Kingston RE, Moore D, Seidman JG, Smith JA, Struhl K (1996) Current Protocols of Molecular Biology. Greene and Wiley, New York

Barbier-Brygoo H, Ephritikhine G, Klambt D, Ghislan M, Guern J (1989) Functional evidence for an auxin receptor at the plasmalemma of tobacco mesophyll protoplasts. Proc Natl Acad Sci USA **86:** 891–895

Barbier-Brygoo H, Ephritikhine G, Klambt D, Maurel C, Palme K, Schell J, Guern J (1991) Perception of the auxin signal at the plasma membrane of tobacco mesophyll protoplasts. Plant J **1:** 83–93

Bingham PM, Zachar Z (1989) Retrotransposons and the FB transposon from *Drosophila melanogaster*. *In* DE Berg, MM Howe, eds, Mobile DNA. American Society for Microbiology, Washington, DC, pp 485–502

Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. Proc Natl Acad Sci USA **93:** 9374–9377

Britten RJ (1997) Mobile elements inserted in the distant past have taken on important functions. Gene **205:** 177–182

Bureau TE, Ronald PC, Wessler SR (1996) A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci USA **93:** 8524–8529

Bureau TE, Wessler SR (1992) *Tourist*: a large family of small inverted-repeat elements frequently associated with maize genes. Plant Cell **4:** 1283–1294

Bureau TE, Wessler SR (1994a) *Stowaway*: a new family of inverted repeat elements associated with the genes of both monocotyledonous and dicotyledonous plants. Plant Cell **6:** 907–916

Bureau TE, Wessler SR (1994b) Mobile inverted-repeat elements of the *Tourist* family are associated with the genes of many cereal grasses. Proc Natl Acad Sci USA **91:** 1411–1415

Casacuberta E, Casacuberta JM, Puigdomenech P, Monfort A (1998) Presence of miniature inverted-repeat transposable elements (MITEs) in the genome of *Arabidopsis thaliana*: characterization of the Emigrant family of elements. Plant J **16:** 79–85

Cavin Périer R, Junier T, Bonnard C, Bucher P (1999) The eukaryotic promoter database (EPD): recent developments. Nucleic Acids Res **27:** 307–309

Charrier B, Foucher F, Kondorosi E, d'Aubenton-Carafa Y, Thermes C, Kondorosi A, Ratet P (1999) *Bigfoot*: a new family of MITE elements characterized from the *Medicago* genus. Plant J **18:** 431–441

Dellaporta SL, Wood J, Hicks JB (1985) Maize DNA mini prep. *In* R Malmberg, J Messing, I Sussex, eds, Molecular Biology of Plants: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 36–37

Donis-Keller H (1979) Site-specific enzymatic cleavage of RNA. Nucleic Acids Res **7:** 179–192

Dufau ML, Tsai-Morris CH, Hu ZZ, Buczko E (1995) Structure and regulation of the luteinizing hormone receptor gene. J Steroid Biochem Mol Biol **53:** 283–291

Eickbush TH (1992) The non-LTR class of retrotransposable elements. New Biol **4:** 430–440

Gerasimova TI, Gdula DA, Gerasimov DV, Simonova O, Corces VG (1995) A *Drosophila* protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. Cell **82:** 587–597

Goodchild TJD, Poulter RTM (1998) The *Care-2* and *Rel-2* repetitive elements of *Candida albicans* contain LTR fragments of a new retrotransposon. Gene **218:** 85–93

Hesse T, Feldwisch J, Balshusemann D, Bauw G, Puype M, Vandekerckhove J, Lobler M, Klambt D, Schell J, Palme K (1989) Molecular cloning and structural analysis

of a gene from *Zea mays* (L.) coding for a putative receptor for the plant hormone auxin. EMBO J **8:** 2453–2461

Hesse T, Garbers C, Brzobohaty B, Kreimer G, Soll D, Melkonian M, Schell J, Palme K (1993) Two members of the *ERabp* gene family are expressed differentially in reproductive organs but to similar levels in the coleoptile of maize. Plant Mol Biol **23:** 57–66

Inohara N, Shimomura S, Fukui T, Futai M (1989) Auxin-binding protein located in the endoplasmic reticulum of maize shoots: molecular cloning and complete primary structure. Proc Natl Acad Sci USA **86:** 3564–3568

Izsvak Z, Ivics Z, Shimoda N, Mohn D, Okamoto H, Hackett PB (1999) Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. J Mol Evol **48:** 13–21

Jones AM, Im K-H, Savka MA, Wu M-J, DeWitt NG, Shillito R, Binns AN (1998) Auxin-dependent cell expansion mediated by overexpressed auxin-binding protein 1. Science **282:** 1114–1117

Jones AM, Prasard PV (1992) Auxin-binding proteins and their possible roles in auxin-mediated plant cell growth. Bioassays **14:** 43–48

Kapitonov VV, Jurka J (1999) The long terminal repeat of an endogenous retrovirus induces alternative splicing and encodes an additional carboxy-terminal sequence in the human leptin receptor. J Mol Evol **48:** 248–251

Kidwell MG, Lisch D (1997) Transposable elements as sources of variation in animals and plants. Proc Natl Acad Sci USA **94:** 7704–7711

Kriz AL, Boston RS, Larkins BA (1987) Structural and transcriptional analysis of DNA sequences flanking genes that encode 19 kilodalton zeins. Mol Gen Genet **207:** 90–98

Kunze R (1996) The maize transposable element *Activator* (*Ac*) in transposable elements. Curr Top Microbiol Immunol **204:** 161–194

Lazarus CM, Napier RM, Yu L-X, Lynas C, Venis MA (1991) Auxin-binding proteins: antibodies and genes. Symp Soc Exp Bot **45:** 129–148

Liao D, Pavelitz T, Weiner AM (1998) Characterization of a novel class of interspersed LTR elements in primate genomes: structure, genomic distribution, and evolution. J Mol Evol **46:** 649–660

Liu CH, Hla T (1997) The mouse gene for the inducible G-protein-coupled receptor *edg-1*. Genomics **43:** 15–24

Lobler M, Klambt D (1985) Auxin-binding protein from coleoptile membranes of corn (*Zea mays* L.): I. Purification by immunological methods and characterization. J Biol Chem **260:** 9848–9853

Macdonald H (1997) Auxin perception and signal transduction. Physiol Plant **100:** 423–430

MacRae AF, Clegg MT (1992) Evolution of *Ac* and *Ds1* elements in select grasses (Poaceae). Genetica **86:** 55–66

Mager DL, Goodchild NL (1989) Homologous recombination between the LTRs of a human retrovirus-like element causes a 5-kb deletion in two siblings. Am J Hum Genet **45:** 848–854

McDonald JF (1993) Evolution and consequences of transposable elements. Curr Opin Genet Dev **3:** 855–864

McDonald JF (1995) Transposable elements: possible catalysts of organismic evolution. Trends Ecol Evol **10:** 123–126

McDonald JF, Matyunina LV, Wilson S, Jordan IK, Bowen NJ, Miller WJ (1997) LTR retrotransposons and the evolution of eukaryotic enhancers. Genetica **100:** 3–13

Meisler MH, Ting C-N (1993) The remarkable evolutionary history of the human amylase genes. Crit Rev Oral Biol Med **4:** 503–509

Napier RM (1997) Trafficking of the auxin-binding protein. Trends Plant Sci **2:** 251–255

Noma K, Nakajima R, Ohtsubo H, Ohtsubo E (1997) *RIRE1*, a retrotransposon from wild rice *Oryza australiensis*. Genes Genet Syst **72:** 131–140

Ray PM, Dohrmann U, Hertel R (1977) Specificity of auxin-binding sites on maize coleoptile membranes as possible receptors for auxin. Plant Physiol **60:** 585–591

Schiefelbein JW, Furtek DB, Dooner HK, Nelson OE Jr (1988) Two mutations in a maize bronze-1 allele caused by transposable elements of the *Ac-Ds* family alter the quantity and quality of the gene product. Genetics **120:** 767–777

Schwob E, Choi S-Y, Simmons C, Migliaccio F, Ilag L, Hesse T, Palme K, Soll D (1993) Molecular analysis of three maize 22 KDa auxin-binding protein genes: transient promoter expression and regulatory regions. Plant J **4:** 423–432

Smale S (1997) Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. Biochim Biophys Acta **1351:** 73–88

Smit A, Riggs A (1996) *Tiggers* and other DNA transposon fossils in the human genome. Proc Natl Acad Sci USA **93:** 1443–1448

Stavenhagen JB, Robins DM (1988) An ancient provirus has imposed androgen regulation on the adjacent mouse sex-limited protein gene. Cell **55:** 247–254

Surzycki SA, Belknap WR (1999) Characterization of repetitive DNA elements in Arabidopsis. J Mol Evol **48:** 684–691

Tian H, Klambt D, Jones AM (1995) Auxin-binding protein 1 does not bind auxin within the endoplasmic reticulum despite this being the predominant subcellular location of this hormone receptor. J Biol Chem **270:** 26962–26969

Tillmann U, Viola G, Kayser B, Siemeiter G, Hesse T, Palme K, Lobler M, Klambt D (1989) cDNA clones of the auxin-binding protein from corn coleoptile (*Zea mays* L.): isolation and characterization by immunological methods. EMBO J **8:** 2463–2467

Troutt AB, McHeyzer-Williams MG, Pulendran B, Nossal GJV (1992) Ligation-anchored PCR: a simple amplification technique with single-sided specificity. Proc Natl Acad Sci USA **89:** 9823–9825

Tu Z (1997) Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes aegypti*. Proc Natl Acad Sci USA **94:** 7475–7480

White SE, Habera LF, Wessler SR (1994) Retrotransposons in the flanking regions of normal plant genes: a role for *copia*-like elements in the evolution of gene structure and expression. Proc Natl Acad Sci USA **91:** 11792- 11796