Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 3894–3898, May 1992
Developmental Biology

HD-Zip proteins: Members of an *Arabidopsis* homeodomain protein superfamily

(plant homeobox genes/leucine zipper motif)

MARK SCHENA AND RONALD W. DAVIS

Department of Biochemistry, Beckman Center, Stanford University Medical Center, Stanford, CA 94305-5307

Contributed by Ronald W. Davis, January 10, 1992

ABSTRACT **Homeobox genes encode a large family of homeodomain proteins in animal systems. To test whether such genes are also abundant in higher plants, degenerate oligonucleotides complementary to sequences encoding the recognition helix (helix three) of the homeodomain were used to screen genomic and cDNA libraries from the plant *Arabidopsis thaliana*. Analysis of 8 of the 41 cDNAs isolated revealed that each encodes a presumptive homeodomain; interestingly, most of these clones also contain a leucine zipper motif tightly linked to the homeodomain. It is concluded that *Arabidopsis* encodes a large family of homeodomain proteins, including members that contain a homeodomain/leucine-zipper (HD-Zip) motif.**

The homeobox, a DNA sequence motif, was first identified in a set of *Drosophila* developmental genes (1, 2), though it is now clear that large numbers of homeobox genes are present in many animal systems, including insects, amphibians, fish, rodents, and mammals (3). These genes control a vast array of developmental decisions in animals, apparently acting as molecular switches to control the fate of cells during development (4). The homeobox encodes a 61-amino acid sequence, the homeodomain, which confers the site-specific DNA binding function of homeodomain proteins (5, 6). Genetic and structural analyses of the homeodomain suggest a general model for homeodomain binding to DNA, in which the most highly conserved of three α-helices (helix 3) fits directly into the major groove (7–10). Homeobox genes thus constitute a large family of genes whose products determine the fate of cells in developing animals by directly modulating gene expression (1–10).

In contrast to genetic regulators of animal development, factors that control the development of higher plants remain largely unidentified. Understanding plant development at the molecular level is interesting, particularly given the fact that certain aspects of plant development differ considerably from animal development (11). These differences include the following: plant cells exhibit no cell movement during development; plant tissues retain totipotency following differentiation; organogenesis is continuous during development; and plant development is intimately linked to environmental signals (11). Despite these differences, the conservation of basic transcriptional regulatory mechanisms in plants and animals (12–15) suggests that higher plants might utilize genes analogous to those used by animals to control development. To investigate this notion, it seemed reasonable as a first step to determine whether homeobox genes could be recovered from the plant *Arabidopsis thaliana*.

MATERIALS AND METHODS

Oligonucleotide Synthesis. Oligonucleotides were synthesized by using a Millipore 8700 DNA synthesizer by the PAN Facility (Beckman Center, Stanford University).

Plasmids and DNA Sequencing. Sequencing of *HAT4*, *-5*, *-22*, and *-24* was performed by inserting the cDNAs (Fig. 1*C*, clones 4, 5, 22, and 24) into the *Eco*RI site of Bluescript (Stratagene), followed by double-stranded plasmid sequencing (PAN Facility, Beckman Center, Stanford University) using *Taq* DNA polymerase, dye primers, and a 373A automated DNA sequencer as recommended by the manufacturer (ABI).* The 5′ portion of the *HAT4* cDNA (nucleotides 1–310) was obtained by sequencing an overlapping cDNA (Fig. 1*C*, clone 13). Data were analyzed by using the Macintosh DNA Strider and MacDraw software programs (Apple, Cupertino, CA). Protein sequences were aligned by using the FastDB program (IntelliGenetics). In the R. W. Davis strain collection, Bluescript plasmids bearing the *HAT* cDNAs are referred to as follows: pNN476 (*HAT4*), pNN477 (*HAT5*), pNN478 (*HAT22*), and pNN479 (*HAT24*).

Genomic and cDNA Libraries. The genomic library was prepared by using DNA (16) from aseptically grown *Arabidopsis thaliana* (Columbia) seedlings. The DNA, which ranged in size from 5 to 50 kilobases (kb), was rendered blunt with T4 DNA polymerase, ligated to *Eco*RI adaptors (pGAGCTCGAG/CTCGAGCTCTTAA), isolated as 10- to 20-kb inserts by agarose gel electrophoresis, phosphorylated with T4 polynucleotide kinase, and ligated to an equimolar mix of LambdaGEM-11 (Promega) cut with *Eco*RI and *Xba* I. Approximately 200,000 clones were packaged *in vitro*, amplified, and stored in 7% (vol/vol) dimethyl sulfoxide at −70°C. The cDNA library was prepared by using poly(A)-purified mRNA (17) derived from the aerial parts of 6-week-old soil-grown *Arabidopsis thaliana* (Columbia). The mRNA was reverse transcribed by using avian myeloblastosis virus reverse transcriptase as described (17) and inserted into the *Eco*RI site of λYES (17). Approximately 10^6 recombinants were packaged *in vitro*, amplified, and stored in 7% dimethyl sulfoxide at −70°C.

Library Screening. The *Escherichia coli* strain LE392 (BNN127) containing plasmid pMC9 (17) was infected with phage from the genomic and cDNA libraries and was plated on ten 15-cm plates per library at a density of 5000 plaques per plate. Plaques were transferred to Colony/PlaqueScreen (DuPont) filters as recommended by the manufacturer and were hybridized and washed essentially as described by Bürglin *et al.* (18). Briefly, filters were prehybridized for 16 hr at 37°C in 500 ml of 6× SSC/5× Denhardt's solution/0.05% sodium pyrophosphate/1% SDS and *E. coli* tRNA at 0.1 mg/ml (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7; 1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone). Hybridization was performed for 20 hr at 42°C in 375 ml of prehybridization solution supplemented with dextran sulfate at 100 mg/ml and 750 ng of an equimolar mix of HB-1, PRD-1, and PRD-2 (Fig. 1*B*) end-labeled with [γ-^{32}P]ATP to a final

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M90394 for *HAT4*, M90416 for *HAT5*, M90417 for *HAT22*, and M90418 for *HAT24*).

specific activity of 10^6 cpm/ml. Filters were rinsed three times for 15 min each at room temperature with 500 ml of 6× SSC/0.05% sodium pyrophosphate, then with three 100 ml volumes of 3.0 M tetramethylammonium chloride/50 mM Tris·HCl, pH 8.0/0.2 mM EDTA. Filters were then washed twice for 15 min each at 54°C in 350 ml of 3.0 M tetramethylammonium chloride/50 mM Tris·HCl, pH 8.0/0.2 mM EDTA to remove oligonucleotides with more than two mismatches (18).

cDNA Screening. The cDNA inserts were liberated from the λYES vectors (17) with *Eco*RI, following conversion of the λ clones to plasmid DNA. Digests were subjected to agarose gel electrophoresis, transferred to GeneScreen hybridization membrane (DuPont), and probed with the HB-1/PRD-1/PRD-2 oligonucleotides, using the conditions employed to screen the genomic and cDNA libraries.

RNA Analysis. Ten micrograms of total RNA (19) was fractionated by agarose gel electrophoresis, transferred to GeneScreen membrane, probed with *HAT4, -5, -22*, and *-24* *Eco*RI inserts (see Fig. 1*C*) and autoradiographed for 7 days with two intensifying screens. The cDNA inserts were labeled with ^{32}P by using a random primer labeling kit (Stratagene). Electrophoresis and hybridization conditions were those recommended by the GeneScreen manufacturer.

RESULTS

To isolate homeobox genes from *Arabidopsis*, a strategy developed by Bürglin *et al.* (18) was used in which degenerate oligonucleotide probes were employed to identify homeobox genes from *Caenorhabditis elegans*. Three pools of degenerate oligonucleotides (Fig. 1*A*; HB-1, PRD-1, and PRD-2) complementary to sequences encoding the highly conserved homeodomain helix three were used to screen *Arabidopsis* genomic and cDNA libraries. Under high-stringency conditions, approximately 1% of the 10,000 to 20,000-base-pair (bp) genomic inserts scored positive (Fig. 1*B*), suggesting that homeoboxes in the *Arabidopsis* genome occur once every 1–2 million bp. Assuming a genome size of 70 million bp (21), *Arabidopsis* probably contains between 35 and 70 homeobox genes.

From a total of 50,000 cDNA plaques screened, 48 clones (≈0.1% of the total) scored positive (Fig. 1*B*). The 10-fold lower frequency of positive clones in the cDNA library compared with the genomic library is attributed to the facts that cDNAs were prepared solely to mRNA isolated from the above-ground parts of adult plants, homeobox genes are typically expressed at low levels, and most inserts in the genomic library probably contain more than one gene. Of the 48 cDNA plaques identified, 41 clones continued to score positive through two additional rounds of purification. Hybridization analysis of the cDNA inserts revealed a spectrum of intensities from inserts that ranged in size from 0.5 to 2.2 kb (Fig. 1*C*).

To characterize a subset of the cDNAs in more detail, the nucleotide sequences of eight inserts that hybridized strongly to HB-1 (Fig. 1*C*; clones 4, 5, 10, 13, 22, 24, 25, and 47) were determined. Of the eight clones analyzed, all eight contain open reading frames that exhibit considerable amino acid homology to known homeodomains (Fig. 2*A*), and four of the eight unequivocally correspond to distinct genes. These genes are designated *HAT4, -5, -22*, and *-24* (for homeobox from *Arabidopsis thaliana*), in accordance with the recommended *Arabidopsis* nomenclature.

The nucleotide sequence of clone 13 (Fig. 1*C*) shares extensive identity with *HAT4* and presumably represents an overlapping cDNA. Clones 10, 25, and 47 (Fig. 1*C*) are identical to *HAT5* within the homeodomain, though nonidentity at both the 5′ and 3′ ends of the cDNAs as well as the presence of single nucleotide differences outside the homeobox suggests that these sequences may have arisen from

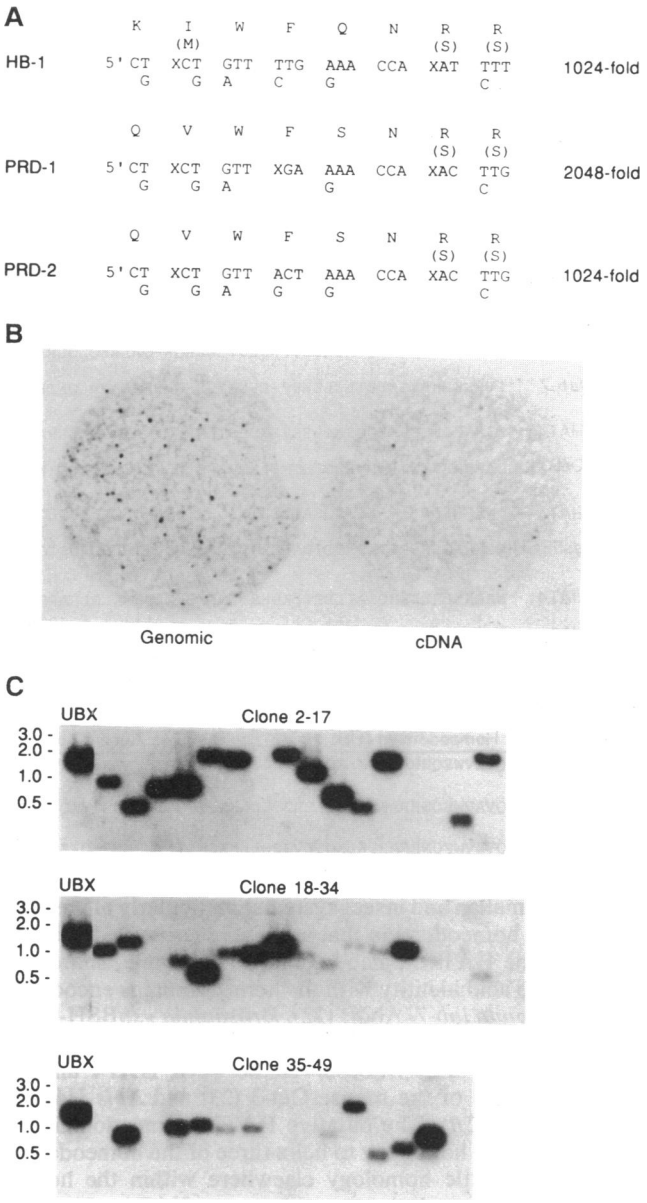

FIG. 1. Homeobox screen of *Arabidopsis* genomic and cDNA libraries. (*A*) Oligonucleotides used as homeobox probes. Shown is the sequence of three 1024- to 2048-fold degenerate oligonucleotides complementary to the coding strand of homeodomain helix three (18). Residue symbols in parentheses denote amino acids that have not been previously observed at these positions, and X denotes positions in which all four nucleotides were included in the synthesis of the oligonucleotides. (*B*) Genomic and cDNA library screen using homeobox probes. Shown are filter lifts of plaques from *Arabidopsis* genomic and cDNA libraries probed with an equimolar mix of HB-1, PRD-1, and PRD-2 (see *A*) and washed under conditions that allow two or fewer mismatches. (*C*) cDNA clone screen using homeobox probes. Shown is a DNA blot of 48 cDNAs (clones 2–49) probed and washed as in Fig. 1*B*. An equimolar amount of plasmid DNA containing the *Drosophila* Ultrabithorax (UBX) homeobox (20) was included as a hybridization standard. Insert sizes were estimated by using DNA markers (BRL) whose migration positions (0.5–3.0 kb) are shown to the left of each blot.

mRNAs encoded by distinct cellular genes; nonetheless, these clones have not been given a *HAT* number, as it has not been determined unequivocally that these cDNAs correspond to distinct genes.

Alignment of the homeodomains of HAT4, -5, and -22 reveal that these proteins exhibit considerable similarity to each other and to previously characterized homeodomains

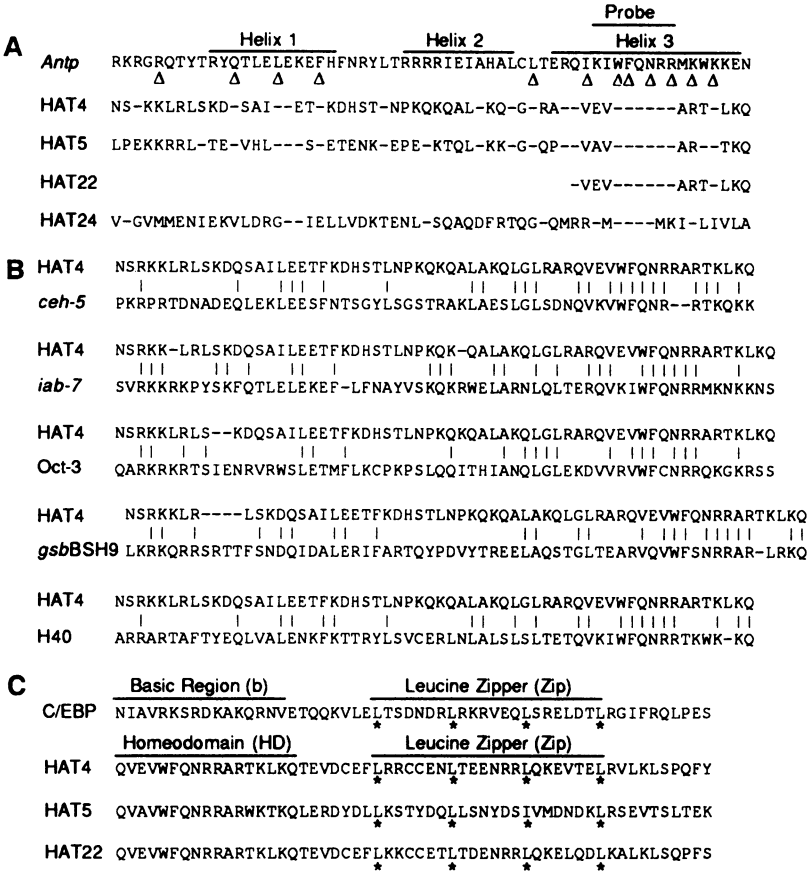

FIG. 2. Homeodomain and leucine zipper motifs of *Arabidopsis* HAT proteins. (*A*) Sequence of the HAT homeodomains. Shown are the HAT4, -5, -22, and -24 homeodomains aligned with the Antennapedia (*Antp*) homeodomain from *Drosophila* (3, 5). The three α-helices of *Antp* (3, 5), as well as helix three residues targeted by the degenerate oligonucleotide probes (see Fig. 1*A*) are indicated by solid bars. Residues that are highly conserved among all homeodomain proteins (3, 5) are indicated by open triangles, and identities between the HAT proteins and *Antp* are indicated by dashes. The convention for genetic nomenclature recommended by the committee at the Third International *Arabidopsis* Meeting (East Lansing, MI, April 1987) requires wild-type gene names to be italicized, capitalized, and composed of three letters. In this paper, designations for genes and gene products are identical except that the latter are not in italics. (*B*) Comparison of the HAT4 homeodomain with homeodomains from animal systems. The HAT4 homeodomain is aligned with *C. elegans* ceh-5 product (18), *Drosophila iab-7*/AbdB (22), mouse Oct-3 (23), *Drosophila gsb*BSH-9 (24), and *Apis* H40 (25) homeodomains. Amino acid identities and gaps inserted into the sequences to optimize alignments are indicated by vertical bars and dashes, respectively. (*C*) Sequence of the HAT leucine zippers. Shown are the putative leucine zippers of HAT4, -5, and -22 aligned with the mammalian C/EBP protein (26–28). Solid bars indicate the basic region and leucine zipper of C/EBP, as well as the homeodomains and putative leucine zippers of the HAT proteins. Asterisks denote leucine and isoleucine residues that define the heptad repeats.

from mammalian and insect systems, particularly at positions within the homeodomain that are highly conserved (Fig. 2 *A* and *B*). The HAT4 homeodomain, for example, shares 40–45% amino acid identity with the homeodomains encoded by the *Drosophila iab-7*/AbdB (22), *Drosophila gsb*BSH-9 (24), and *C. elegans ceh-5* (18) genes (Fig. 2*B*); somewhat lesser similarity (35–40%) is observed between HAT4 and the homeodomains of the mouse Oct-3 (23) and *Apis* H40 (25) proteins (Fig. 2*B*). The putative HAT24 homeodomain displays sequence homology to helix three of the homeodomain but exhibits little homology elsewhere within the homeodomain and bears little resemblance to HAT4, -5, or -22; furthermore, the presence of a glycine residue in presumptive helix one raises some question about whether HAT24 contains a bona fide homeodomain (Fig. 2*A*).

Additional sequence analysis of the cDNAs revealed a second motif common to HAT4, -5, and -22, namely the presence of a leucine zipper motif tightly linked to each of the homeodomains (Fig. 2*C*). The presence of four leucine residues repeated at exactly seven-amino acid intervals is reminiscent of the leucine zipper dimerization motif previously described in C/EBP (26, 27); importantly, the spacing of the putative leucine zippers relative to the HAT homeodomains is identical to the spacing between the leucine zipper and DNA-binding domain of C/EBP (26–28). It is interesting to note that the putative leucine zippers of HAT4 and HAT22 are nearly identical at the amino acid level, despite the fact that the corresponding nucleotide sequence of these regions is highly divergent (Fig. 2*C* and data not shown). The HAT24 protein does not contain a leucine zipper motif.

Hybridization analysis revealed that each of the *HAT* genes encodes transcripts of low abundance ranging in size from 1.2 to 1.8 kb (Fig. 3). On the basis of hybridization intensity, it is estimated that *HAT4*, *-5*, and *-24* constitute about 0.012%, 0.05%, and 0.004% of the mRNA in adult tissue, respectively (Fig. 3 and data not shown). The transcript encoded by *HAT22* constitutes about 0.002% of the mRNA in adult plants and has not been detected in seedlings (Fig. 3 and data not shown).

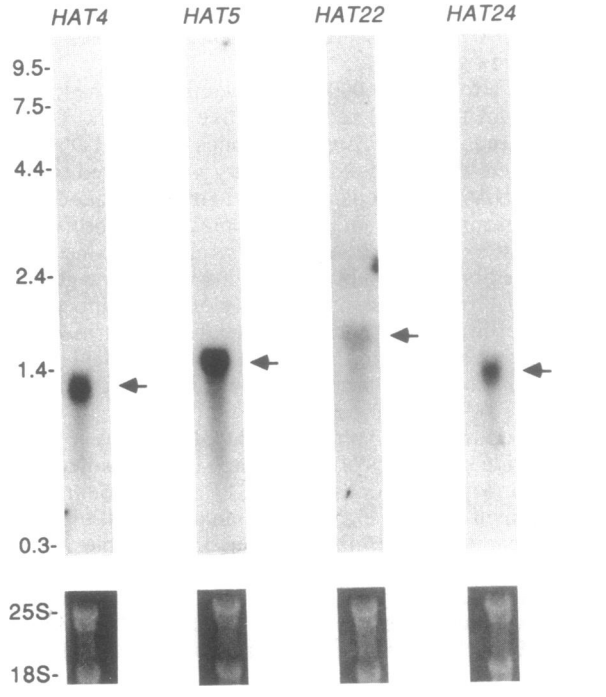

FIG. 3. Identification of *HAT* transcripts. Shown is a hybridization (Northern) analysis, using *HAT* cDNAs as probes, of adult *Arabidopsis thaliana* (Columbia) RNA. Arrows denote the position of each transcript relative to RNA molecular weight standards (BRL) which ranged in size from 0.3 to 9.5 kb. At the bottom are photographs of the ethidium-stained nuclear ribosomal RNAs (25S and 18S).

```
ctt caa ctg ctg cgt cta gct gtc ctt cct tct tca ttt gct tct tct tct cta gct cag   60

cgc gga tcg ctg cag tag tac cct gac gta gcc ttt ctt ctt cct tta ctt tct cat ctt  120

cta tct ctc aaa aga aaa gca gac aac ttt att tgc aaa aac aga gtt ttt ttt tct tat  180

ctt gag aaa gtt caa cag aag ATG ATG TTC GAG AAA GAC GAT CTG GGT CTA AGC TTA GGC  240
                           Met Met Phe Glu Lys Asp Asp Leu Gly Leu Ser Leu Gly   13

TTG AAT TTT CCA AAG AAA CAG ATC AAT CTC AAA TCA AAT CCA TCT GTT TCT GTT ACT CCT  300
Leu Asn Phe Pro Lys Lys Gln Ile Asn Leu Lys Ser Asn Pro Ser Val Ser Val Thr Pro   33

TCT TCT TCT TCT TTT GGA TTA TTC AGA AGA TCT TCA TGG AAC GAG AGT TTT ACT TCT TCA  360
Ser Ser Ser Ser Phe Gly Leu Phe Arg Arg Ser Ser Trp Asn Glu Ser Phe Thr Ser Ser   53

GTT CCA AAC TCA GAT TCG TCA CAA AAA GAA ACA AGA ACT TTC ATC CGA GGA ATC GAC GTG  420
Val Pro Asn Ser Asp Ser Ser Gln Lys Glu Thr Arg Thr Phe Ile Arg Gly Ile Asp Val   73

AAC AGA CCA CCG TCT ACA GCG GAA TAC GGC GAC GAA GAC GCT GGA GTA TCT TCA CCT AAC  480
Asn Arg Pro Pro Ser Thr Ala Glu Tyr Gly Asp Glu Asp Ala Gly Val Ser Ser Pro Asn   93

AGT ACA GTC TCA AGC TCT ACA GGG AAA AGA AGC GAG AGA GAA GAA GAC ACA GAT CCA CAA  540
Ser Thr Val Ser Ser Ser Thr Gly Lys Arg Ser Glu Arg Glu Glu Asp Thr Asp Pro Gln  113
                                                    *   *   *       *

GGC TCA AGA GGA ATC AGT GAC GAT GAA GAT GGT GAT AAC TCC AGG AAA AAG CTT AGA CTT  600
Gly Ser Arg Gly Ile Ser Asp Asp Glu Asp Gly Asp Asn Ser Arg Lys Lys Leu Arg Leu  133
                         *   *   *   *       *

TCC AAA GAT CAA TCT GCT ATT CTT GAA GAG ACC TTC AAA GAT CAC AGT ACT CTC AAT CCG  660
Ser Lys Asp Gln Ser Ala Ile Leu Glu Glu Thr Phe Lys Asp His Ser Thr Leu Asn Pro  153

AAG CAG AAG CAA GCA TTG GCT AAA CAA TTA GGG TTA CGA GCA AGA CAA GTG GAA GTT TGG  720
Lys Gln Lys Gln Ala Leu Ala Lys Gln Leu Gly Leu Arg Ala Arg Gln Val Glu Val Trp  173

TTT CAG AAC AGA CGA GCA AGA ACA AAG CTG AAG CAA ACG GAG GTA GAC TGC GAG TTC TTA  780
Phe Gln Asn Arg Arg Ala Arg Thr Lys Leu Lys Gln Thr Glu Val Asp Cys Glu Phe Leu  193

CGG AGA TGC TGC GAG AAT CTA ACG GAA GAG AAC CGT CGG CTA CAA AAA GAA GTA ACG GAA  840
Arg Arg Cys Cys Glu Asn Leu Thr Glu Glu Asn Arg Arg Leu Gln Lys Glu Val Thr Glu  213

TTG AGA GTA CTT AAG CTC TCT CCT CAG TTC TAC ATG CAC ATG AGC CCA CCC ACT ACT TTG  900
Leu Arg Val Leu Lys Leu Ser Pro Gln Phe Tyr Met His Met Ser Pro Pro Thr Thr Leu  233

ACC ATG TGC CCT TCA TGT GAA CAC GTG TCG GTC CCG CCA CCA CAA CCT CAG GCT GCT ACG  960
Thr Met Cys Pro Ser Cys Glu His Val Ser Val Pro Pro Pro Gln Pro Gln Ala Ala Thr  253

TCA GCG CAC CAC CGG TCG TTG CCG GTC AAT GCG TGG GCT CCT GCT ACG AGG ATA TCT CAC 1020
Ser Ala His His Arg Ser Leu Pro Val Asn Ala Trp Ala Pro Ala Thr Arg Ile Ser His  273

GGC TTG ACT TTT GAC GCT CTT CGT CCT AGG TCC TAA gtc ttt tta ctt gca acc aaa ggg 1080
Gly Leu Thr Phe Asp Ala Leu Arg Pro Arg Ser OCH                                  284

cat ttt ggt cgt ttt tta agt ttc atg gac cag ata tgc atg tag ttg tta aca tgt atg 1140

tat ttt ctt aga aag aaa gaa aaa cag att aat att aaa aaa aaa aaa aaa aaa aaa aaa    1197
```

FIG. 4. Nucleotide sequence of the *HAT4* cDNA. Shown is the 1197-bp sequence of the *HAT4* cDNA and the deduced 284-amino acid protein encoded by the largest open reading frame. The homeodomain, leucines of the presumptive leucine zipper, and acidic region are indicated by solid lines, boxes, and asterisks, respectively.

To investigate whether the HAT proteins contain additional sequence motifs common to transcription factors, the complete nucleotide sequence of *HAT4* cDNA was determined. Analysis of the 1197-bp cDNA revealed an open reading frame encoding a 284-amino acid protein with a predicted molecular mass of approximately 32 kilodaltons (Fig. 4). A cluster of negatively charged amino acids found immediately upstream of the HAT4 homeodomain (Fig. 4) is reminiscent of the acidic activation domains of well-characterized transcriptional activators (29). The size of the intact *HAT4* cDNA (1197 bp) agrees well with the estimated size of the *HAT4* transcript (1.2 kb) and thus probably represents a full-length cDNA (Figs. 3 and 4).

DISCUSSION

These results indicate that *Arabidopsis thaliana* contains a large number of homeobox genes, confirming and extending two reports that appeared recently documenting the existence of a small number of homeobox genes in maize (30) and *Arabidopsis* (31). On the basis of the frequency with which clones were isolated from the genomic library and assuming a genome size of 70 million bp (21), it is estimated that *Arabidopsis* contains upwards of 50 homeobox genes. Thus, similar to *Drosophila* (32) and *Caenorhabditis* (18), it appears that *Arabidopsis* utilizes the homeodomain extensively as a functional motif despite the vast evolutionary distance separating organisms of the plant and animal kingdoms. These results support further the idea that basic mechanisms of transcriptional control have been highly conserved in eukaryotes (29, 33–35).

What is the functional role of homeobox genes in *Arabidopsis?* It is tempting to speculate that, as in animal systems (4), plant homeobox genes encode transcriptional regulators that mediate important developmental processes. The presence of a homeodomain, putative leucine zipper, and potential acidic activation domain in several of the HAT proteins strongly suggests that these factors function as transcriptional regulators in *Arabidopsis*. Though the developmental role of these genes remains to be investigated, it is encouraging to note that mutations in the maize homeobox gene *Knotted-1* cause alterations in maize leaf development (30).

It is interesting that most of the plant homeodomain proteins that we have characterized contain a leucine zipper motif. We propose that the juxtaposition of the homeodomain and a putative leucine zipper as seen in HAT4, -5, and 22 should be referred to as an HD-Zip motif to highlight their distinction from the b-Zip proteins from animal systems, which include C/EBP, GCN4, FOS, and JUN (26, 27, 36, 37). The spacing of the homeodomain and the putative leucine zipper in *Arabidopsis* HD-Zip proteins is identical to the distance between the DNA-binding domain and the leucine zipper in b-Zip proteins (26–28, 36, 37); moreover, the presence of characteristic hydrophobic and charged residues within the heptad repeats is analogous to the b-Zip class (26–28, 38, 39). These similarities suggest that HD-Zip pro-

teins may, like members of the b-Zip class, utilize the leucine zipper motif as a dimerization domain. An exciting possibility is that, analogous to FOS and JUN interactions (36, 37), leucine zipper interactions between members of the HD-Zip family might mediate the formation of functional heterodimers.

It is striking that HD-Zip proteins have not been described in animals, despite characterization of more than 100 homeodomain proteins from animal systems (3). What is the significance of the apparent uniqueness of HD-Zip proteins to higher plants? Given that profound differences exist between plant and animal development (11), it is tempting to speculate that HD-Zip proteins mediate aspects of development that are unique to plants, such as the coupling of development to environmental signals (40). Assuming that HD-Zip proteins are indeed determinants of *Arabidopsis* development, the ability of these proteins to interact as heterodimers might provide a means by which to alter a particular developmental pathway in plants simply by modulating the ratios of the HD-Zip proteins *in vivo*. The answers to these and other basic questions regarding plant homeobox gene function are expected to be complex, particularly if large numbers of these gene products interact during plant development. The isolation, mapping, and sequencing of these genes, coupled with reverse genetics, should greatly aid in the illustration of their function.

We thank the members of the PAN Facility for oligonucleotide synthesis and DNA sequencing, and J. Mulligan for preparation of the genomic and cDNA libraries. We also thank members of the Krasnow and Davis laboratories for helpful discussions and members of the Hake and Zambryski laboratories for communication of results. We thank M. Scott, B. Johnson, and M. Walberg for critical comments on the manuscript. This work was supported by grants from the National Institutes of Health (8 R37 HG00198-17) and National Science Foundation (DMR-9106011) to R.W.D.; M.S. holds a National Science Foundation Postdoctoral Fellowship in plant biology.

1. Scott, M. P. & Weiner, A. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4115–4119.
2. McGinnis, W., Levine, M. S., Hafen, E., Kuroiwa, A. & Gehring, W. J. (1984) *Nature (London)* **308**, 428–433.
3. Scott, M. P., Tamkun, J. W. & Hartzell, G. W., III (1989) *Biochim. Biophys. Acta* **989**, 25–48.
4. Hayashi, S. & Scott, M. P. (1990) *Cell* **63**, 883–894.
5. Gehring, W. J., Müller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y. Q., Otting, G. & Wüthrich, K. (1990) *Trends Genet.* **6**, 323–329.
6. Harrison, S. C. (1991) *Nature (London)* **353**, 715–719.
7. Hanes, S. D. & Brent, R. (1989) *Cell* **57**, 1275–1283.
8. Hanes, S. D. & Brent, R. (1991) *Science* **251**, 426–430.
9. Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990) *Cell* **63**, 579–590.
10. Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991) *Cell* **67**, 517–528.
11. Steeves, T. A. & Sussex, I. M. (1990) *Patterns in Plant Development* (Cambridge Univ. Press, New York), 2nd Ed., pp. 1–5.
12. Ma, J., Przibilla, E., Hu, J., Bogorad, L. & Ptashne, M. (1988) *Nature (London)* **334**, 631–633.
13. Hilson, P., de Froidmont, D., Lejour, C., Hirai, S.-Y., Jaquemin, J.-M. & Yaniv, M. (1990) *Plant Cell* **1**, 651–658.
14. Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R. G. & Chua, N.-H. (1990) *Nature (London)* **346**, 390–394.
15. Schena, M., Lloyd, A. M. & Davis, R. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10421–10425.
16. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D. & Smith, J. A., eds. (1988) *Current Protocols in Molecular Biology* (Greene & Wiley-Interscience, New York), pp. 2.3.1–3.
17. Elledge, S. J., Mulligan, J. T., Ramer, S. W., Spottswood, M. & Davis, R. W. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1731–1735.
18. Bürglin, T. R., Finney, M., Coulson, A. & Ruvkun, G. (1989) *Nature (London)* **341**, 239–243.
19. Sharrock, R. A. & Quail, P. H. (1989) *Genes Dev.* **3**, 1745–1757.
20. Kornfeld, K., Saint, R. B., Beachy, P. A., Harte, P. J., Peattie, D. A. & Hogness, D. S. (1989) *Genes Dev.* **3**, 243–258.
21. Leutwiler, L. S., Hough-Evans, B. R. & Meyerowitz, E. M. (1984) *Mol. Gen. Genet.* **194**, 15–23.
22. Celniker, S. E., Keelan, D. J. & Lewis, E. B. (1989) *Genes Dev.* **3**, 1424–1436.
23. Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W. J. & Staudt, L. M. (1990) *Nature (London)* **345**, 686–692.
24. Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. & Noll, M. (1986) *Cell* **47**, 1033–1040.
25. Walldorf, U., Fleig, R., Gehring, W. J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9971–9975.
26. Landschulz, W. H., Johnson, P. F. & McKnight, S. L. (1988) *Science* **240**, 1759–1764.
27. Agre, P., Johnson, P. F. & McKnight, S. L. (1989) *Science* **246**, 922–926.
28. Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J. & McKnight, S. L. (1988) *Genes Dev.* **2**, 786–800.
29. Ptashne, M. & Gann, A. A. F. (1990) *Nature (London)* **346**, 329–331.
30. Vollbrecht, E., Veit, B., Sinha, N. & Hake, S. (1991) *Nature (London)* **350**, 241–243.
31. Ruberti, I., Sessa, G., Lucchetti, S. & Morelli, G. (1991) *EMBO J.* **10**, 1787–1791.
32. Ashburner, M., Glover, D. M., Saunders, R. D. C., Duncan, I., Hartl, D., Merriam, J., Lee, G., Johnsen, J., Kafatos, F. C., Sidén-Kiamos, I., Louis, C. & Savakis, C. (1991) *Science* **254**, 247–262.
33. Schena, M. & Yamamoto, K. R. (1988) *Science* **241**, 965–967.
34. Schena, M., Freedman, L. P. & Yamamoto, K. R. (1989) *Genes Dev.* **3**, 1590–1601.
35. Schena, M. (1989) *Experientia* **45**, 972–983.
36. Struhl, K. (1989) *Trends Biochem. Sci.* **14**, 137–140.
37. Jones, N. (1990) *Cell* **61**, 9–11.
38. O'Shea, E. K., Rutkowski, R. & Kim, P. S. (1989) *Science* **243**, 538–542.
39. O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991) *Science* **254**, 539–544.
40. Braam, J. & Davis, R. W. (1990) *Cell* **60**, 357–364.