

Identification of a homeobox-containing gene with enhanced expression during soybean (*Glycine max* L.) somatic embryo development

Hongchang Ma¹, Michael D. McMullen^{1,2,3} and John J. Finer^{1,*}

¹Department of Agronomy, The Ohio State University (*author for correspondence); ²Corn and Soybean Research Unit, USDA-Agricultural Research Service, Wooster, OH 44691, USA; ³Present address: USDA-ARS, Room 301 Curtis Hall, University of Missouri, Columbia, MO 65211, USA

Received 31 May 1993; accepted in revised form 16 November 1993

Key words: embryogenesis, *Glycine max*, helix-turn-helix motif, homeobox, homeodomain

Abstract

Homeotic genes are key 'switches' that control developmental processes. Homeotic genes containing the consensus 'homeobox' domain have been identified from a number of organisms including *Drosophila melanogaster*, *Caenorhabditis elegans*, *Homo sapiens*, and *Zea mays*. Although homeotic genes have been demonstrated to be important in embryo development of some insects, amphibians, and mammals, there are no reports of their involvement in plant embryogenesis. Here, we report the isolation and characterization of a cDNA clone for a homeobox-containing gene expressed in somatic embryos of soybean. The cDNA (*Sbhl* for soybean homeobox-containing gene) was isolated using maize *Knotted-1* (*Kn1*) cDNA as a heterologous probe. The *Sbhl* cDNA clone is 1515 bp long which is the approximate size of its transcript. Within the homeodomain, the amino acid sequence of a helix-turn-helix structure, and invariant and conserved residues were identified. The deduced SBH1 protein shares a high amino acid identity with KN1 protein (47.0% overall and 87.5% for the homeodomain). Southern hybridization analysis indicated that *Sbhl* is a member of a small gene family. The expression of *Sbhl* is development- and tissue-specific. The transcript of *Sbhl* was present in early-stage somatic embryos, increased prior to cotyledon formation and decreased thereafter. *Sbhl* was weakly expressed in soybean stems and hypocotyl but was not detected in other plant tissues and nonembryogenic materials. The enhanced expression during embryogenesis, the homology with maize *Kn1* gene, and the regulatory nature of homeodomain proteins suggest that the SBH1 protein plays an important role in plant embryo development.

Introduction

Homeotic genes control pattern formation and morphological structure determination in multi-

cellular eukaryotes [10, 11, 32]. Initially identified in *Drosophila* [21, 33] as single gene mutations capable of altering the identity of complex morphological structures, homeotic genes have been

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number L13663.

isolated and characterized from a wide range of animal and plant species [5, 10]. A common feature of many homeotic genes of both plants and animals is a conserved nucleotide sequence known as the homeobox which encodes the homeodomain. The secondary structure of the homeodomain proteins and the interaction between the homeodomain and target DNA have been confirmed by NMR spectroscopy [27] and the homeodomain-DNA complex has been identified [16]. The homeodomain consists of about 60 amino acids with a helix-turn-helix structure conferring a specific DNA binding function.

Although homeotic genes containing homeobox sequences have been shown to be important in embryo development of insects, amphibians and mammals [10, 32], there are no reports of their involvement in plant embryogenesis. Here, we report the isolation of a cDNA for a homeobox-containing gene that is preferentially expressed in the early stages of somatic embryo development of soybean.

Materials and methods

Plant materials

Embryogenic suspension cultures of soybean were initiated and maintained according to Finer and Nagasawa [9]. Suspensions were subcultured monthly in a liquid medium containing 23 μ M 2,4-dichlorophenoxyacetic acid and 5 mM asparagine. For embryo development, embryogenic tissue was transferred to a hormone-free solid medium at 23 °C [8]. Developing embryos were harvested 0, 3, 7, 14, and 28 days after transfer to the embryo development medium.

cDNA library construction and screening

Total RNA was extracted from proliferating (day 0) embryogenic suspension culture material using guanidinium thiocyanate [6] and poly(A)⁺ RNA was obtained by one passage on an oligo (dT)-cellulose column [1]. A cDNA library was con-

structed from the poly(A)⁺ RNA in λ gt10 according to the protocols in the cDNA Synthesis System Plus and cDNA Cloning System- λ gt10 Kits (Amersham, Arlington Heights, IL). *Escherichia coli* NM514 was infected with phage and plated on 100 mm plates at density of 500 pfu/plate. Three replica filters were prepared from each plate [3]. After baking under vacuum for 2 h at 80 °C, the filters were prehybridized in 5 \times SSC, 5 \times Denhardt's solution, 0.1% SDS, and 100 μ g/ml denatured salmon sperm DNA at 55 °C for 6 h. *Knl* cDNA ([37], kindly provided by S. Hake, USDA-ARS, Albany, CA) probe was prepared by random priming [7]. Hybridization was performed for 2 days using the same solution as the prehybridization plus probe. Replica filters were washed with either 4 \times , 2 \times , 1 \times SSC plus 0.1% SDS at 55 °C. After autoradiography, plaques hybridizing to the *Knl* cDNA probe were further purified by two rounds of plaque purification.

Subcloning and sequencing

The phage DNA was digested with *Hind* III and *Bgl* II to release the intact cDNA fragment. The fragment was purified using GeneClean (BIO 101, La Jolla, CA) and subcloned into pBluscript KS- and SK- in both orientations. Ordered deletions were created using exonuclease III as described by Sambrook *et al.* [30]. Single-strand DNAs were produced [35] and used for sequencing with the dideoxy chain termination method [30]. Sequence data were analyzed using IntellGenetics (Mountain View, CA) sequence analysis programs and NCBI BLAST E-Mail Server.

Southern hybridization analysis

Genomic DNA was isolated from leaf tissue of soybean cv. 'Fayette' according to Saghai-Marooft *et al.* [29]. DNAs were digested with restriction enzymes, electrophoresed, and transferred onto GeneScreen Plus membrane by the 'dry blot'

method [15]. The membranes were prehybridized 6 h in $5 \times$ SSC, $5 \times$ Denhardt's solution, 0.1% SDS, 50 mM Tris pH 8.0, 10 mM EDTA, and 100 μ g/ml denatured salmon sperm DNA and then hybridized for 2 days with random primer-labelled [7] *Sbh1* cDNA in the above solution plus 10% dextran sulfate. Membranes were washed either at low stringency ($1 \times$ SSC, 0.1% SDS, 65 °C) or at high stringency ($0.1 \times$ SSC, 0.1% SDS, 65 °C). The membranes were exposed on Kodak film with intensifying screens for 2–3 days.

Northern hybridization analysis

Total RNA and poly(A)⁺ RNA were isolated as described above. One μ g of poly(A)⁺ RNA or 10 μ g of total RNA were used for gel electrophoresis and northern hybridization analysis as described by Thomas [34] except a modified 'dry blot' procedure [15] was used for RNA transfer. The modified 'dry blot' was performed as follows. After electrophoresis, the gel was laid on a single sheet of Whatman 3M paper saturated with 10 mM sodium phosphate buffer (pH 6.5). A GeneScreen Plus membrane (Du Pont) was moistened with 10 mM sodium phosphate buffer and then placed on the top of gel. The membrane was covered with one sheet of Whatman 3M paper soaked in 10 mM sodium phosphate buffer followed by four dry sheets of Whatman 3M paper and a 5 cm stack of paper towels. A 0.5 to 1 kg weight was placed on top of the paper towels and the transfer proceeded for 1–2 h. After hybridization with *Sbh1* cDNA, membranes were washed at high stringency ($0.1 \times$ SSC, 0.1% SDS, 65 °C) and exposed on Kodak film with intensifying screens for 2–3 days.

Results and discussion

Embryo development

Development of somatic embryos from the clumps of suspension culture tissue progressed

rapidly [9]. As early as 3 days after plating, greening of the apical portion of the embryos was observed. By day 7, somatic embryos had enlarged and by 14 days after plating, elongation of the embryo was evident. Well-developed embryos with clearly defined cotyledons and hypocotyls were formed by 28 days after plating (Fig. 1).

Isolation of a cDNA clone encoding a homeodomain protein

About 6000 plaques from the proliferating soybean somatic embryo cDNA library were screened using the maize *Kn1* cDNA [37] as a probe. Ten plaques initially appeared to hybridize to *Kn1* cDNA. From these 10 clones, one clone, which gave strong signals on three replica filters washed at different stringency ($4 \times$, $2 \times$, $1 \times$ SSC, with 0.1% SDS) was selected for further characterization. We designated the cDNA clone *Sbh1* for soybean homeobox-containing gene 1). The other nine original clones were discarded due to the lack of hybridization in subsequent screening with *Kn1* cDNA.

Structures of *Sbh1* clone and deduced protein

The DNA sequence of the *Sbh1* cDNA is 1515 bp long (Fig. 2). The predicted SBH1 protein is 379 amino acids long with a translation start codon AUG at position 156 and stop codon UAG at nucleotide 1293. The overall amino acid identity between SBH1 and KN1 is 47.0% (87.5% identity with KN1 homeodomain) (Fig. 3). The SBH1 homeodomain, located near the carboxyl terminus, is 64 amino acids long and contains the four invariant residues in the recognition helix (helix III) and six of the eight highly conserved residues or similar substitutions (Fig. 4). Although a high sequence similarity was observed between the homeodomain of SBH1 and homeodomains of three maize proteins (KN1, ZMH1, ZMH2), the homology between the homeodomain of SBH1 and other homeodomain proteins was limited to the most conserved residues (Fig. 4).

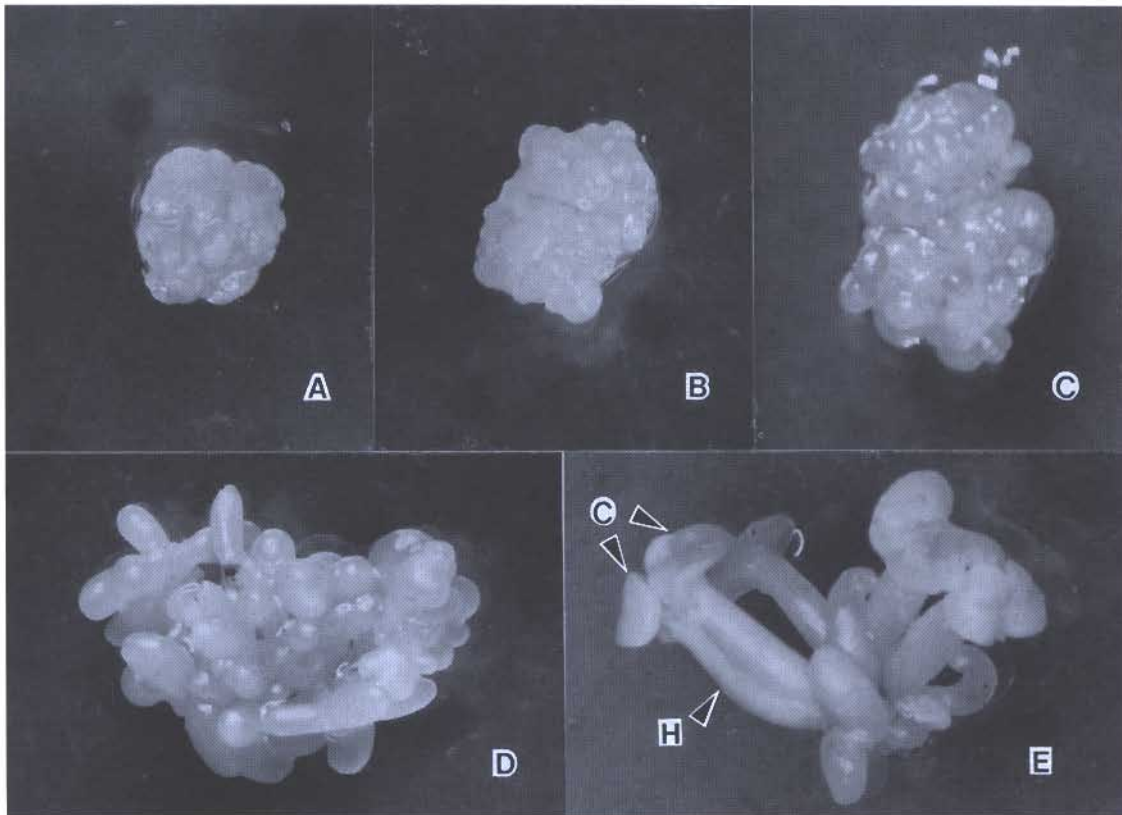


Fig. 1. The development of soybean somatic embryos. A. A clump of embryogenic soybean tissue from proliferation medium (referred as 0 day in the text). B. Three days after transfer to development medium, C: After 14 days, E: After 28 days. The letter C indicates cotyledons, H indicates hypocotyl.

The ELK (glutamic acid, leucine, and lysine) region found in other plant homeodomain proteins [36] is also present in the SBH1 protein (Fig. 3). The 'invariant' glutamic acid is replaced by an aspartic acid and the periodicity of the hydrophobic residues is conserved. The PYP (proline, tyrosine, proline) sequence between helices 1 and 2, conserved in all five maize homeodomain proteins [2] is also present in SBH1 protein (Fig. 4). An acidic region, characteristic of transcription activating domains [26], is present at the amino side of the SBH1 homeodomain (Fig. 2). This region includes 110 amino acids with a net charge of -15 . A leucine zipper region that is found in certain homeobox-containing genes from *Arabidopsis thaliana* [20, 28, 31] is not present in either SBH1 or KN1 proteins.

Although high homology is shown between SBH1 and KN1 proteins, several unique features are present in *Sbhl* cDNA and SBH1 protein. There is a longer leader sequence prior to the translation start site of the predicted SBH1 protein (Fig. 2). Within this leader sequence, there is a small open reading frame at position 108–141 which potentially encodes the peptide MCVVCVCVCC. Comparison of this sequence with the Protein Database did not reveal similarity with any other leader peptides. We have not yet determined whether this leader peptide is present *in vivo*. If it is present, it may be involved in translational control similar to other leader peptides of some regulatory proteins such as the three ORFs in *Opaque-2* of maize [19] and the four short ORFs in *GCN-4* of yeast [12].

Although most eukaryotic proteins initiate at

AATAAGAGAATTGTGTGTCGTGTTTGTGTTTTGTTGGTTTGTGTAAGGTTAGCTAGTG	59
AGTATTCTACACAAGGTTGGTGGTAGGGCAAAAAGGATAAGACAGTGAATGTGTGTGTGT	119
GTGTGTGTGTGTGTGTGTGTGTTGACAAGCAAAAGCTATGGAGGGTGGTAGTAGTAGCTCT	179
M E G G S S S S	
AATGGCACTTCTTATCTGTGGCTTTTGGAGAAAACAACAGTGGTGGGCTATGCCCAATG	239
N G T S Y L L A F G E N N S G G L C P M	
ACGATGATGCCTTTGGTGACTTCCCATCAGCTGGTCATCATCCAATAAATCCTAGTAAT	299
T M M P L V T S H H A G H H P I N P S N	
AATAATAATGTAACACAACTGTCTCTTCATTCCCAACTGCAGTAACAGTACTGGAAC	359
N N N V N T N C L F I P N C S N S T G T	
CCTTCTATCATGCTCCACAATAATCACAACAACAACAAACTGATGATGATAACAAC	419
P S I M L H N N H N N N K T D D D D N N	
AACAACACTGGGTTAGGTTACTATTTTCATGGAGAGTGACCACCACCACATCACCACGGC	479
N N T G L G Y Y F M E S D H H H H H G	
AACAACAACAACAATGGAAGCTCCTCCTCCTCCTCTCTGCTGTCAAGGCCAAGATC	539
N N N N N G S S S S S S S A V K A K I	
ATGGTCTATCCTCACTATACCGTCTCTGGCAGCTTACGTCAATTGTCAGAAGGTTGGG	599
M A H P H Y H R L L A A Y V N C Q K V G	
GCCCGCCTGAAGTGGTGGCAAGTTAGAAGAAGCATGTGCTTCTGCAGCGACAATGGCT	659
<u>A P P E V V A R L E E A C A S A A T M A</u>	
GGTGGTGTGCAGCAGCTGGATCAAGCTGCATAGGTGAAGATCCAGCTTTGGATCAGTTC	719
<u>G G D A A A G S S C I G E D P A L D Q F</u>	
ATGGAGGCTTACTGTGAGATGCTCACAAGTATGAGCAAGAAGTCTCCAAACCCTTAAAG	779
<u>M E A Y C E M L T K Y E Q E L S K P L K</u>	
GAAGCCATGCTCTTCTCAAAGGATCGAGTGCAGTTCAAAATCTTACAATTTCTTCC	839
<u>E A M L F L Q R I E C Q F K N L T I S S</u>	
TCCGACTTTGCTAGCAATGAGGCTGGTATAGGAATGGATCGTCTGAAGAGGATGTTGAT	899
<u>S D F A S N E G G D R N G S S E E D V D</u>	
CTACACAACATGATAGATCCCCAGGCAGAGGACAGGGATTTAAAGGGTCAGCTTTTGCGC	959
<u>L H N M I D P Q A E D R D L K G Q L L R</u>	
AAGTATAGCGGATACTTGGGCAGTCTGAAGCAAGAATTCATGAAGAAGAGGAAGAAAGGA	1019
<u>K Y S G Y L G S L K Q E F M K K R K K G</u>	
AAGTACTCTAAAGAAGCAAGGCAACAATTAATGAAATGGTGGAAACAGACATTACAAATGG	1079
<u>K L P K E A R Q Q L L E W W N R H Y K W</u>	
CCTTACCCATCCGAATCCAGAAGCTGGCTCTTGCAGAGTCCGACAGGCTGGATCAGAAG	1139
<u>P Y P S E S Q K L A L A E S T G L D Q K</u>	
CAAATCAACAACCTGGTTTATTAATCAAAGGAAACGGCACTGGAAGCCTTCAGAGGACATG	1199
<u>Q I N N W F I N Q R K R H W K P S E D M</u>	
CAGTTTGTGGTGTGATGGATCCAAGCCATCCACACTATTACATGGATAATGTTCTAGGCAAT	1259
<u>Q F V V M D P S H P H Y Y M D N V L G N</u>	
CCATTTCCCATGGATCTTTCCCATCCCATGCTCTAGAAAATTATCCCTCGTTTGTGGGCT	1319
<u>P F P M D L S H P M L *</u>	
GCTGATAATAGATTCAAACTCGTGTGCTGCTACTTATTAACCTTACAATTATTAATAT	1379
<u>TAATTAATATGCATTCTAAGAAATCCTGATGCTATACTATAATATAGTACGCAGGTGT</u>	1439
<u>ATCCCTTGCTAGCTTTTGTAGACGGTCTTGTGTGGATCATCTAGTTGAAGGAGTTATGA</u>	1499
<u>ATAAATAAAATCCAT</u>	1515

Fig. 2. Nucleic acid sequence of *Sbhl* and the deduced amino-acid sequence of SBH1. The homeobox nucleic acid sequence and the homeodomain protein are boxed. Putative polyadenylation signals are underlined. The suggested acidic region is noted by double underline. The nucleic acid sequence that possibly encodes the short leader peptide is in bold type.

the first AUG, about 5–10% deviation exists from the 'first-AUG rule' [17, 18]. In the case of SBH1, the predicted protein starts at the second AUG instead of the first (Fig. 2). When comparing the sequence context of the first AUG (CAGUGAAUGU) and second AUG (AAAGCUAUGG) with the consensus sequence for initiation in higher eukaryotes GCC(A/G)CCAUGG, the purine in position -3 and G in position +4 are present in the second AUG con-

text indicating that the second AUG is more favorable for protein initiation.

Sbhl is a member of a small gene family

When leaf tissue DNA of soybean was digested with several restriction enzymes, the recognition sequences of which were not present in the *Sbhl* cDNA sequence, and hybridized with *Sbhl*

```

MEGSSSSNGTSYLLAFGENNSGGLCPMTMPLVTSHHAGHHPINPSNNNNVNTNCLFIPNCNSSTGTPTS
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
MEEITQHFQ---VGASSHGCHGQHHHHHHHPWASSLSAVV---APLPPQPPSAGLP-

IMLHNNHNNKTDNNTGLGYYFMESDHHHHHHGNNNNNGSSSSSS-----SSAVKAKIMAHPHYH
--LTLN-----TVAATGN-SGGSGNPVQLANGGGLLDACVKAKEPSSSSPYAGDVEAIKAKIISHPHYH

RLLAAYVNCQKVGAPPEVVARL-EEACASAA---TMAGGDAAGSSCIGEDPALDQFMEAYCEMLTKYEQ
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
SLLTAYLECNKVGAPPEVSARLTEIAQEVEARQRTALGGLAAA-----TEPELDQFMEAYHEMLVKFRE

ELS KPLKEAMLFQRIEQPKNLTISSSDFASNEGGDRNGSSEED-----VDLHNMIDPQAE
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
ELTRPLQEFAMEFMRVESQLNSLSIS---GRSLRNILSSGSSEEDQEGSGGETELPEVDAH-----GVD

RDLKGQLLRKYSGLKQEFMKRKRKGLPKPEARQQLLEWNNRHYKWPYPSESOKLALAEESTGLDQKQ
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QELKHHLLKKYSGLSLSLQELSSKKKKGLPKPEARQQLLSWWDQHYKWPYPSETQKVALAEESTGLDLKQ

INNWFINQKRHRHWKPSEDMQFVVMDPHSH--PHYMDNVLGNPFPMDLSHPML
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
INNWFINQKRHRHWKPSEEMHLMMDGYHTTNAFYMD---GHFINDGGLYRL

```

Fig. 3. The comparison between the predicted SBH1 protein (top line) and KN1 (bottom line). The residues identical to those of SBH1 are indicated by hyphens and the ELK region is underlined. The conserved and hydrophobic residues in the ELK region are in bold type.

cDNA, 2–3 fragments were clearly detected (Fig. 5). The size and intensity of the fragments suggest that more than one copy is present in the soybean genome. This is supported by the isolation

of other homeobox-containing genes from the same cDNA library using the homeobox region from *Sbh1* as a probe (data not shown). At least one of these other clones is different from

	Helix I	Helix II turn	Helix III	Reference
Sbh1	MKKRKKGLPKPEARQQLLEWNNRHYKWPYPSESOKLALAEESTGLDQKQ	INNWFINQKRHRHWKPS		<i>Glycine</i>
Kn1	S--K-----S--DQ-----T--V-----L-----			<i>Zea</i> (37)
Zmh1	LR--RA----GDTTTSI-KQ--QE-S-----T-DD-AK-V-E---QL-----N-HNN			<i>Zea</i> (37)
Zmh2	LR--RA----GDTAST-KA--QA-S-----T-ED-AR-VQE---QL-----N-HNN			<i>Zea</i> (37)
Zmhox1a	NSTARKGHFGPVINQKLHEHFQTKQ---R-V-ES---EL--TFR-V-K--ETR-HSARVA-			<i>Zea</i> (2)
Zmhox1b	NI-DRKGFHFGPVISQKLHEHFQTKQ---R-L-ES---EL--TFH-V-R--E-R-HFARLA-			<i>Zea</i> (2)
Athb1	QLPE-KRRLTTE-VHLLKESFETENKLEPER-TQ--KKL--QPR-VAV--Q-R-A-WKTKQ			<i>Arabidopsis</i> (28)
Athb2	DNSR-KLRLSKD-SAILEETFKDHSTLNPK--Q---KQL--RAR-VEV--Q-R-A-TKLKQ			<i>Arabidopsis</i> (28)
Athb3	MLGE-KKRLNLE-VRALEKSFELGNKLEPER-MQ--KAL--QPR--AI--Q-R-A-WKTKQ			<i>Arabidopsis</i> (20)
HAT4	NSR-KLRLSKD-SAILEETFKDHSTLNPK--Q---KQL--RAR-VEV--Q-R-A-TKLKQ			<i>Arabidopsis</i> (31)
HAT5	LPE-KRRLTTE-VHLLKESFETENKLEPER-TQ--KKL--QPR-VAV--Q-R-A-WKTKQ			<i>Arabidopsis</i> (31)
PR1	AR--RRNFN-Q-TEI-N-YFYS-LSN-----EA-EE--KKC-ITVS-VS---G-K-I-YK-NI			<i>Homo</i> (25)
PBX2	AR--RRNFS-Q-TEV-N-YFYS-LSN-----EA-EE--KKC-ITVS-VS---G-K-I-YK-NI			<i>Homo</i> (23)
MATPI	MTTVR-QCS-CTKPH-MR-LLL--DN-----N-EFYD-SAATG-TRTQLRN--S-R-R			<i>Saccharomyces</i> (14)
MATa2	T-PYRGHRFT--NVRI-ES-FAKNIEN--LDTKGLN-MKN-S-SRI--K--VS-R-RKEKTIT			<i>Saccharomyces</i> (24)
CEH-5	P--PRTDNAD-QLEK-E-SF-...TSG-L-G-TRAK-----L--SDN-VKV--Q-R-TKQK-ID			<i>Caenorhabditis</i> (4)
Antp	ER--GRQTYTRYQTLE-EKEF...HFNR-LTRRRRIEI-HALC-TER--KI--Q-R-MKWK-EN			<i>Drosophila</i> (21, 33)
En1	ED--PRTAFTA-QL-R-KAEF...QANR-ITEQRQT--QELS-NES--KI--Q-K-AKIK-AT			<i>Mus</i> (13)
Consensus	R Q L F		L I/VWF N R K K	

Fig. 4. The amino acid sequence of SBH1 homeodomain aligned with several homeodomains from other organisms. The residues identical to those of SBH1 are indicated by hyphens. The helix-turn-helix structure is positioned according to the motif of KN1. The 4 invariant residues in the recognition helix are marked with an asterisk (*) and the eight highly conserved residues are marked with a period (·).

Sbhl based on restriction patterns and the intensity of the hybridization signals (data not shown).

Expression of the *Sbhl* clone

Soybean embryogenic suspension cultures proliferate as clumps of globular embryos (Fig. 1A). After transfer to the hormone-free solid development medium, the globular embryos developed to the heart, torpedo, and cotyledon stages (Fig. 1). Hybridization of the *Sbhl* cDNA to poly(A)+ RNAs isolated from soybean somatic embryos at different stages of development detected two transcripts (Fig. 6A). The larger transcript (about 1.9 kb) did not vary in intensity up to day 14 and might represent unprocessed *Sbhl* transcript or transcript from a different homeobox-containing gene. The smaller transcript (about 1.6 kb) of similar size to the *Sbhl* cDNA was present during the embryo proliferation stage, increased during

early somatic embryo development (7 days after transfer to hormone-free medium), and decreased thereafter (14 after post-transfer). Both the larger and smaller transcripts were present at very low levels at 28 day after transfer to the development medium. Additional northern hybridization analysis using poly(A)+ RNAs from another soybean embryogenic suspension culture line indicated a similar expression pattern although the decrease in *Sbhl* expression at the 28 day time point was not as pronounced (data not shown). *Sbhl* transcript was not detected in nonembryogenic soybean cultures, roots or leaves but occurred at a low level in stems (Fig. 6B) and hypocotyl (data not shown).

Sbhl mRNA levels were highest from 7 to 14 days after transfer to hormone-free medium. In soybean somatic embryogenesis, this corresponds to embryo elongation and cotyledon initiation (formation of heart-shaped embryos) (Fig. 1). The globular-heart transition is significant in that the cells that give rise to the cotyledons and prevascular tissue are determined [22]. Because *Sbhl* is

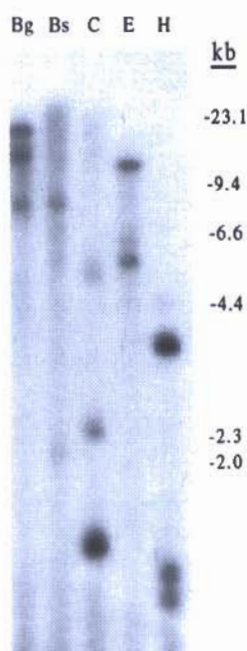


Fig. 5. Southern hybridization analysis of soybean DNA using *Sbhl* cDNA as a probe. Ten μ g soybean leaf DNA were digested with *Bgl* II (Bg), *Bst* EII (Bs), *Cla* I (C), *Eco* RV (E), and *Hind* III (H). The membrane was washed at high stringency ($0.1 \times$ SSC, 65° C).

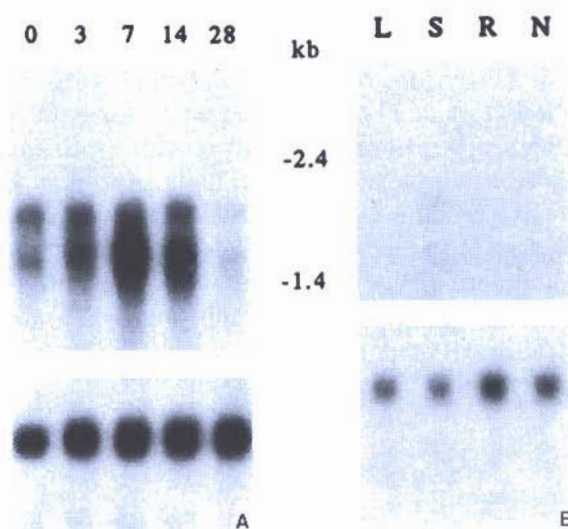


Fig. 6. Northern hybridization analysis of the *Sbhl* transcript. A. One μ g of poly(A)+ RNA from soybean somatic embryos collected 0, 3, 7, 14, and 28 days after transfer to the development medium was loaded per lane. B. Ten μ g of total RNA from leaf (L), stem (S), root (R), and nonembryogenic callus (N) were loaded per lane. The RNAs on each membrane were first hybridized with *Sbhl* cDNA (upper panels) and then re-hybridized with soybean ubiquitin cDNA (lower panels).

expressed at the highest level just at the globular-heart transition stage and SBH1 shares a high homology with KN1 protein, this suggests that SBH1 protein may act as a transcription factor controlling cell differentiation associated with vascular tissue formation. Efforts are currently underway to understand the function of the SBH1 protein by observing alterations in phenotype in soybean embryos and plants transformed with sense and antisense constructions of *Sbhl* and by defining the localization of *Sbhl* expression via *in situ* hybridization. Further analysis of *Sbhl* and other homeobox-containing genes associated with embryogenesis will lead to a better understanding of the molecular mechanism underlying embryogenesis in plants.

Acknowledgements

The authors would like to thank Sarah Hake for providing *Kn1* cDNA clone and Desh-Pal Verma for supplying the soybean ubiquitin clone. We thank Mark Jones, Barbara Norris, and Brenda Schult for technical assistance. Thanks is also extended to John Streeter, Kimberly Campbell and Kevin Simcox for critical reading of the manuscript and to Mark Lagrimini for help in analyzing sequence data. This research was supported by State and Federal funds appropriated to OARDC and USDA-ARS. Mention of trademark or proprietary products does not constitute a warranty of the product by OSU/OARDC or USDA, and also does not imply approval to the exclusion of other products that may also be suitable. OARDC Journal Article 88-93.

References

1. Aviv H, Leder P: Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci USA* 69: 1408-1412 (1972).
2. Bellmann R, Werr W: *Zmhox1a*, the product of a novel maize homeobox gene, interacts with the *Shrunken* 26 bp *feedback* control element. *EMBO J* 11: 3367-3374 (1992).
3. Benton WD, Davis RW: Screening λ gt recombinant clones by hybridization to single plaques *in situ*. *Science* 196: 180-182 (1977).
4. Bürglin TR, Finney M, Coulson A, Ruvkun G: *Caenorhabditis elegans* has scores of homeobox-containing genes. *Nature* 341: 239-243 (1989).
5. Chasan R: In the realm of the homeodomain. *Plant Cell* 4: 237-240 (1992).
6. Davis LG, Dibner MD, Batty JF: *Basic Methods in Molecular Biology*. Elsevier Science Publishing, New York (1986).
7. Feinberg AP, Vogelstein B: A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 132: 6-13 (1983).
8. Finer JJ, McMullen MD: Transformation of soybean via particle bombardment of embryogenic suspension culture tissue. *In Vitro Cell Devel Biol* 27: 175-182 (1991).
9. Finer JJ, Nagasawa A: Development of an embryogenic suspension culture of soybean [*Glycine max* (L.) Merrill.]. *Plant Cell Tissue Organ Cult* 15: 125-136 (1988).
10. Gehring WJ: The homeobox in perspective. *Trends Biotechnol Sci* 17: 277-280 (1992).
11. Gehring WJ, Hiromi Y: Homeotic genes and the Homeobox. *Annu Rev Genet* 20: 147-173 (1986).
12. Hinnebusch AG: Mechanisms of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae*. *Microbiol Rev* 52: 248-273 (1988).
13. Joyner AL, Martin GR: *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila* engrailed gene: expression during embryogenesis. *Genes Devel* 1: 29-38 (1987).
14. Kelly M, Burke J, Smith M, Klar A, Beach D: Four mating-type genes control sexual differentiation in the fission yeast. *EMBO J* 7: 1537-1547 (1988).
15. Kempter B, Lippa P, Neumeier D: A short procedure for Southern blotting on neutral and anionic membranes. *Trends Genet* 7: 109-110 (1991).
16. Kissinger CR, Liu B, Martin-Blanco E, Kornberg TB, Pabo CO: Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* 63: 579-590 (1990).
17. Kozak M: An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl Acids Res* 15: 8125-8148 (1987).
18. Kozak M: The scanning model for translation: An update. *Cell Biol* 108: 229-241 (1989).
19. Lohmer S, Maddaloni M, Motto M, Salamini F, Thompson RD: Translation of the mRNA of the maize transcriptional activator *Opaque-2* is inhibited by upstream open reading frames present in the leader sequence. *Plant Cell* 5: 65-73 (1993).
20. Mattsson J, Söderman E, Svenson M, Borkird C, Engström P: A new homeobox-leucine zipper gene from *Arabidopsis thaliana*. *Plant Mol Biol* 18: 1019-1022 (1992).

21. McGinnis W, Levine MS, Hafen E, Kuroiwa A, Gehring WJ: A conserved DNA sequence in homeotic genes of the *Drosophila* antennapedia and biothorax complexes. *Nature* 308: 428–433 (1984).
22. Meinke DW: Embryonic mutants of *Arabidopsis thaliana*. *Devel Genet* 12: 382–392 (1991).
23. Monica K, Galili N, Nourse J, Saltman D, Cleary ML: PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-oncogene PBX1. *Mol Cell Biol* 11: 6149–6157 (1991).
24. Nasmyth KA, Tatchell K, Hall BD, Astell C, Smith M: A position effect in the control of transcription at yeast mating type loci. *Nature* 289: 244–250 (1981).
25. Nourse J, Mellentin JD, Galili N, Wilkinson J, Stanbridge E, Smith SD, Cleary ML: Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell* 60: 535–545 (1990).
26. Ptashne M: How eukaryotic transcriptional activators work. *Nature* 335: 683–689 (1988).
27. Qian YQ, Billeter M, Otting G, Müller M, Gehring WJ, Wüthrich K: The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. *Cell* 59: 573–580 (1989).
28. Ruberti I, Sessa G, Lucchetti S, Morelli G: A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. *EMBO J* 10: 1787–1791 (1991).
29. Saghai-Marroof MA, Soliman KM, Jorgensen RA, Allard RW: Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 81: 8014–8018 (1984).
30. Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
31. Schena M, Davis RW: HD-Zip proteins: Members of an *Arabidopsis* homeodomain protein superfamily. *Proc Natl Acad Sci USA* 89: 3894–3898 (1992).
32. Scott MP, Tamkun JW, Hartzell III GW: The structure and function of the homeodomain. *Biochim Biophys Acta* 989: 25–48 (1989).
33. Scott MP, Weiner AJ: Structural relationship among genes that control development: Sequence homology between the *Antennapedia*, *Ultrabithorax*, and *fushi tarazu* loci of *Drosophila*. *Proc Natl Acad Sci USA* 81: 4115–4119 (1984).
34. Thomas PS: Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. *Meth Enzymol* 100: 255–266 (1983).
35. Vieira J, Messing J: Production of single-stranded plasmid DNA. *Meth Enzymol* 153: 3–11 (1987).
36. Vollbrecht E, Kerstetter R, Lowe B, Veit B, Hake S: Homeobox genes in plant development: mutational and molecular analysis. In Spradling AC (ed), *Evolutionary Conservation of Developmental Mechanisms*, pp. 111–123. Wiley-Liss, New York (1993).
37. Vollbrecht E, Veit B, Sinha N, Hake S: The developmental gene *Knotted-1* is a member of a maize homeobox gene family. *Nature* 350: 241–243 (1991).