Analysis and Mapping of Gene Families Encoding β-1,3-Glucanases of Soybean

Wei Jin,*[,1] Harry T. Horner,* Reid G. Palmer[†] and Randy C. Shoemaker[†]

**Interdepartmental Plant Physiology Program and Department of Botany, [†]Departments of Agronomy, and Zoology/Genetics and USDA ARS CICG Research Unit, Iowa State University, Ames, Iowa 50011*

Manuscript received October 8, 1998
Accepted for publication June 2, 1999

ABSTRACT

Oligonucleotide primers designed for conserved sequences from coding regions of β-1,3-glucanase genes from different species were used to amplify related sequences from soybean [*Glycine max* (L.) Merr.]. Sequencing and cross-hybridization of amplification products indicated that at least 12 classes of β-1,3-glucanase genes exist in the soybean. Members of classes mapped to 34 loci on five different linkage groups using an F_2 population of 56 individuals. β-1,3-Glucanase genes are clustered onto regions of five linkage groups. Data suggest that more closely related genes are clustered together on one linkage group or on duplicated regions of linkage groups. Northern blot analyses performed on total RNA from root, stem, leaf, pod, flower bud, and hypocotyl using DNA probes for the different classes of β-1,3-glucanase genes revealed that the mRNA levels of all classes were low in young leaves. *SGlu2*, *SGlu4*, *SGlu7*, and *SGlu12* mRNA were highly accumulated in young roots and hypocotyls. *SGlu7* mRNA also accumulated in pods and flower buds.

THE β-1,3-glucanases (EC 3.2.1.39) are hormonally and developmentally regulated plant hydrolytic enzymes found during anther and coleoptile development, pollen tube growth, in endosperm cells, and in the end walls of sieve elements. They also are induced upon pathogen infection or by environmental stresses (Abeles *et al.* 1970; Boller 1987; Mauch *et al.* 1988; Castresana *et al.* 1990; Takeuchi *et al.* 1990; Brederode *et al.* 1991; Dong *et al.* 1991; Leah *et al.* 1991; Ward *et al.* 1991; Schröder *et al.* 1992; Beerhues and Kombrink 1994; Beffa *et al.* 1996). Five distinct classes and a total of 12 β-1,3-glucanase genes have been identified in tobacco (Felix and Meins 1986; Van den Bulcke *et al.* 1989; Linthorst *et al.* 1990; Ori *et al.* 1990; Payne *et al.* 1990; Bucciaglia and Smith 1994). The best characterized are the isoforms that are induced by either pathogen infection or ethylene treatment. They exhibit developmental regulation (Felix and Meins 1986) and appear to be localized primarily in vacuoles (Van den Bulcke *et al.* 1989). The second class, which includes the pathogenesis-related proteins PR-2, PR-N, PR-O, and PR-35 (Kauffmann *et al.* 1987; Van den Bulcke *et al.* 1989), also are induced by pathogens and appear to be localized in the extracellular spaces. The third class, which includes the pathogenesis-related protein PR-Q′ (Payne *et al.* 1990), also is induced by pathogens and seems to be localized in extracellular spaces. The members of the fourth class, including sp41a and sp41b, accumulate to high levels in the transmitting tract of the style and are not pathogen inducible (Ori *et al.* 1990). The fifth class is a secreted β-1,3-glucanase that is expressed in the anther tapetum (Bucciaglia and Smith 1994) and is involved in tetrad callose wall dissolution. Related monocot β-1,3;1,4-glucanases encode proteins similar to β-1,3-glucanases. The function of these latter proteins is to catalyze the hydrolysis of β-1,3;1,4-glucans, which are major components of endosperm cell walls (Fincher *et al.* 1986).

Although many physiological studies have been conducted, little work has been done on the genetic analysis of β-1,3-glucanase genes. Seven β-1,3-glucanase genes were located on the long arm of chromosome 3 in barley (Li *et al.* 1996). Six were clustered in a region <20 cM in length. The mode of regulation of these genes is unknown. Genes on the same chromosome seemed to be closely related to one another but substantially different from those on other chromosomes (Muthukrishnan *et al.* 1984).

Characterization of the β-1,3-glucanase genes of soybean will provide information about these gene families and their evolution in a diploidized polyploid. We report the cloning of β-1,3-glucanase genes, the characterization of mRNA accumulation patterns of these genes, and their chromosomal localization in the genome. The linkage relationships of β-1,3-glucanase genes accounting for mRNAs accumulating in different organs, sequence relationships among the genes found at various loci, and the genomic organization of these loci suggest mechanisms underlying β-1,3-glucanase gene regulation, evolution, and possible functions of multiple gene loci.

Corresponding author: Harry T. Horner, Department of Botany, Iowa State University, Ames, IA 50011-1020. E-mail: hth@iastate.edu

[1]*Present address*: Developmental Cell and Molecular Biology Group, Duke University, Durham, NC 27708.

MATERIALS AND METHODS

Plant material and nucleic acid manipulations: Plant tissues for nucleic acid isolation were collected from the soybean cultivar Minsoy (PI 27890) grown in the greenhouse or growth chamber. Material for RNA extraction was immediately frozen in liquid nitrogen and stored at −80° until used. Soybean genomic DNA was extracted from young leaves (Keim *et al.* 1988). Total DNA (10 μg) was digested and subjected to Southern blotting (Sambrook *et al.* 1989). Total RNA was extracted using guanidinium isothiocyanate extraction (Chomczynski and Sacchi 1987), electrophoresed on formaldehyde-agarose gels, and blotted onto Zeta Probe nylon membranes (Bio-Rad, Richmond, CA; Sambrook *et al.* 1989). Equal amount of RNA loading was verified by probing blots with 18S rRNA (data not shown). Blots were prehybridized for 3 hr at 65° in 5× SSC, 2% SDS, 5× Denhardt's solution, and 0.1 mg/ml herring sperm DNA. Hybridization was carried out overnight at 65° in the same solution. Blots were washed at 60° with 2× SSC + 0.4% SDS and 1× SSC + 0.4% SDS at 60° before autoradiography exposure.

Genetic mapping of β-1,3-glucanase genes: Mapping was conducted using a *Glycine max* (L.) Merr. × *G. soja* (Zieb. & Zucc.) population containing 56 individuals (Keim *et al.* 1990). Parental DNA was digested with 14 restriction endonucleases (*Acc*I, *Alu*I, *Bcl*I, *Bam*HI, *Eco*RI, *Eco*RV, *Dra*I, *Hae*III, *Hha*I, *Hin*dIII, *Hin*fI, *Rsa*I, *Ssp*I, and *Taq*I). Each polymorphism was mapped using MapMaker (Lander *et al.* 1987). The Kosambi (Kosambi 1944) mapping function was selected, and a minimum LOD score of 3 was required for a two-point linkage.

PCR amplification, cloning, and sequence analysis: Sequences encoding β-1,3-glucanase were amplified from soybean genomic DNA and cDNA (prepared from flower bud RNA) using primers for conserved sequences of the β-1,3-glucanases of other species (Simmons 1994). To obtain cDNA, 0.1 μg mRNA was prepared from flower buds. The mRNA was incubated with 5 μm random hexamers; 1 mm each of dATP, dCTP, dGTP, and dTTP; and 2 units/μl Rnase inhibitor (Promega, Madison, WI) in 10 mm Tris-HCl (pH 8.3), 50 mm KCl, and 2.5 mm $MgCl_2$ for 30 min at 37°, and then for an additional 45 min at 37° after the addition of 20 units/μl superscript reverse transcriptase (BRL Life Technology, Rockville, MD). After heating at 65° for 3 min, cDNA was used in the PCR experiments. Each 25-μl reaction contained 60 ng of soybean genomic DNA or 20 ng of flower cDNA; 10 mm Tris-HCl (pH 8.3); 50 mm KCl; 3 mm $MgCl_2$; 200 μm (each) dATP, dCTP, dGTP, and dTTP; 2.5 units of Taq polymerase (BRL Life Technology); and 2 μm (each) PCR primer. PCR amplifications were performed under the following conditions: 96° for 1 min, 44° for 1 min, and 72° for 1 min for 4 cycles, followed by 94° for 1 min, 55° for 1 min, and 72° for 1 min for 30 cycles. The 5′ primer was 5′ CGCGG*N*GT*N*TG*Y*TA*Y*GG 3′; the 3′ primer was 5′ CGCGGCCA*N*CC*NSWY*TC 3′ (where *N* is A, C, G, T; *Y* is C, T; *S* is C, G; and *W* is A, T). The regions used correspond to amino acids 37–41 and 276–282 of the mature protein (Simmons 1994). An aliquot of each PCR product was analyzed by agarose gel electrophoresis. A 700-bp fragment was cut from the gel and subjected to another round of PCR using identical conditions. Aliquots of these PCR products were digested with restriction enzymes recognizing 4-bp sites, and the digestion products were compared with undigested DNA on agarose gels. The PCR products were cloned into the pGEM-T vector (Promega), and ~280 clones were tested by dot blot analysis (Sambrook *et al.* 1989) to determine if they cross-hybridized.

Clones were sequenced on both strands using an automated sequencer. Two to five sequences were obtained for each class. DNA sequence analysis was carried out with the DNAsis (Hitachi, San Francisco, CA) and GCG (Genetics Computer Group, Madison, WI) sequence analysis packages. Alignment of sequences was done using CLUSTAL (Thompson *et al.* 1994). Phylogenetic analysis of amino acid sequences was performed using PAUP version 3.0 (Swofford 1991).

BAC library screening: The soybean bacterial artificial chromosomal (BAC) library (Marek and Shoemaker 1996) was replicated onto nylon membranes (Zeta Probe GT, Bio-Rad). The membranes were screened using two β-1,3-glucanase gene probes, *SGlu2* and *SGlu5*.

The sequence data presented in this article have been submitted to the GenBank Data Libraries under the accession nos. AF034106, AF034107, AF034108, AF034109, AF034110, AF034111, AF034112, AF034113, AF0341114, AF034115, AF034116, and AF034117.

RESULTS

Isolation of soybean β-1,3-glucanase genes and sequence analysis: PCR amplification of soybean genomic DNA using degenerate primers designed from conserved regions of other β-1,3-glucanases resulted in the production of an ~700-bp band. The band was excised from the gel and subjected to a further round of PCR. We then asked whether the PCR product consisted of a mixture of DNA sequences, which would be consistent with the amplification of multigene families. The PCR product was digested with *Hae*III, *Alu*I, *Rsa*I, *Mse*I, and *Sau*3AI. Each restriction digestion yielded a series of fragments whose molecular weights summed to a value greater than that of the original PCR product (data not shown). The presence of a heterogeneous PCR product suggested the involvement of multigene families.

The PCR product was cloned, and ~280 clones were analyzed. The clones were grouped into 12 classes on the basis of cross-hybridization results under stringent hybridization conditions (0.1× SSC/0.1× SDS/60° wash).

PCR experiments using cDNA as a template also generated an ~700-bp band, indicating that the targeted genomic coding regions of the members of the multigene families may not be interrupted by introns.

BAC library screening identified 15 BACs representing *Sglu5* and 5 BACs representing *Sglu2*. The copy number of β-1,3-glucanase gene sequences within each BAC was estimated by digesting the BACs with restriction enzymes not having recognition sites within the β-1,3-glucanase probe sequences and by hybridizing with *Sglu2* and *Sglu5* probes (results not shown). *Sglu2* BACs each contained two copies of the sequence. This result agreed with the prediction of two to four copies based on genomic Southern hybridization patterns (results not shown). The class 5 BACs each appeared to have one to two copies of the class 5 β-1,3-glucanase sequence. This result agreed with the prediction of two copies based on genomic Southern hybridization.

Three to five clones from each class were sequenced on both strands, and the deduced amino acid sequences of the representative clones from each class are shown

(Figure 1). Among the sequenced clones, classes 2 and 7 showed heterogeneity (two subclasses). The deduced protein contains 30 amino acids that are identical among all 12 soybean classes and classes I, II, III, IV, and V of tobacco β-1,3-glucanases (Figure 1). These include conserved acidic amino acids, glutamic acid residues 96 and 245, and tryptophan residues 74 and 248. Glutamic acid 245 and tryptophan 248 are surrounded by highly conserved amino acids that are similarly conserved in plant and yeast glucanases (MacGregor and Ballance 1991; Varghese *et al.* 1994).

Sglu9 is probably a pseudogene because it contains multiple stop codons; however, it did show strong similarity with *Sglu8*. *Sglu8* and *Sglu9* clones differ at only 6 positions in 700 bases of overlapping sequence. Pairwise comparisons between different classes revealed that amino acid identities ranged from 11.1 to 81.8%; similarities ranged from 35.2 to 99.9%. Subclasses 7a and 7b differed at only six positions at the DNA sequence level and showed 98.3% amino acid identity.

Phylogenetic analysis of β-1,3-glucanase-coding regions: To determine the relationships of the 12 classes of soybean β-1,3-glucanases to the five classes of β-1,3-glucanase genes from tobacco, a parsimony analysis was conducted. Amino acid sequences were aligned using CLUSTAL (Thompson *et al.* 1994) and then analyzed using PAUP version 3.0 (Swofford 1991). Figure 2 shows the unrooted consensus tree (from 100 bootstrap replicates) that groups four soybean classes (SGlu1, SGlu8, SGlu4, and SGlu12) as a branch with tobacco class I; five classes (SGlu3, SGlu5, SGlu6, SGlu10, and SGlu11) are grouped closely to acidic β-1,3-glucanases (class II) and stylar acidic glucanases (class IV). The structurally similar, pathogen-induced acidic glucanases (class II) and stylar acidic glucanases (class IV) are grouped into one branch. This result agrees with a previous report (Bucciaglia and Smith 1994). SGlu2 is grouped closely to class III. SGlu7a and 7b and Tag1 (class V) are grouped together, suggesting that SGlu7 could be an anther-specific β-1,3-glucanase.

Mapping of 12 classes of β-1,3-glucanase genes: An F_2 population derived from an interspecific cross was used to determine the genomic location of 12 soybean classes of β-1,3-glucanase genes. Clones representing separate classes were hybridized to Southern blots of soybean genomic DNA digested with 14 restriction enzymes to identify polymorphisms and to estimate gene family copy number. Coding regions of genes hybridized to two to seven fragments, indicating that most β-1,3-glucanase class probes are members of a supergene family with classes that range in size from two members (*SGlu7*) to as many as seven members (*SGlu6*). The average number of fragments detected suggests the presence of >40 genes.

The segregation analysis of 56 F_2 individuals of the F_2 population placed the 11 classes into 5 of 26 linkage groups, B1, J, K, L, and N1 (Figure 3). All these linkage groups contained more than two loci of β-1,3-glucanases. Two classes, *SGlu4* and *SGlu12*, cosegregated with the pea β-1,3-glucanase gene involved with fungal resistance (Chang *et al.* 1992).

Some class-specific probes (*SGlu2*, *SGlu3*, *SGlu5*, and *SGlu9*) detected multiple polymorphic fragments. These mapped to different loci, indicating that some gene families are dispersed. *SGlu7a* and *b* were not mapped because no polymorphisms were detected between parental DNA digested with any of the 14 restriction enzymes.

In the soybean, large domains of different linkage groups seem to have been derived from the same ancestral linkage group through duplication (Shoemaker *et al.* 1996a,b). The different domains contain homologous members of the same gene families. Some of the β-1,3-glucanase gene loci (*e.g.*, *SGlu2*) lie within these duplicated regions (Figure 3). For example, two markers (pB162 and PEG488) detected loci on linkage groups L and N1. On both linkage groups, these markers also are linked to multiple *SGlu2* loci (Figure 3). This result suggests that these multiple β-1,3-glucanase gene loci have arisen during evolution via duplications of large chromosomal regions in which linkage relationships between β-1,3-glucanase genes and the other genes have been maintained.

Nucleotide sequences, phylogenetic relationships, and chromosomal locations: Twelve β-1,3-glucanase genes used in the mapping project also were sequenced and analyzed, and the extent to which β-1,3-glucanase genes mapped to individual loci is correlated to their sequence relationships (Figure 3). The sequenced region covers ~60% of the coding region, including the proposed catalytic sites. The 12 β-1,3-glucanase sequences show an average nucleotide sequence identity of 64.8% (range 21.4–95%), an average amino acid sequence identity of 36.2% (range 11.1–81.8%), and an average amino acid sequence similarity of 70.2% (range 35.2–99.9%). Those β-1,3-glucanases that mapped together always showed >88% nucleotide sequence identity. For example, *Sglu5* and *Sglu11* both mapped to a locus on linkage group B. As Figures 2 and 3 show, sequence-similar genes are clustered on one linkage group or on duplicated regions of linkage groups; *e.g.*, *Sglu1*, *4*, *8*, and *12* group into tobacco class I and are all mapped on linkage group K.

mRNA accumulation patterns of different classes: To study mRNA accumulation patterns of 12 classes, the presence of mRNAs corresponding to each class was analyzed by RNA blot analyses using class-specific probes under stringent conditions (Figure 4). The mRNA levels of all classes were quite low in young leaves. However, *SGlu2*, *SGlu4*, *SGlu7*, and *SGlu12* mRNA were highly accumulated in young roots and hypocotyls. *SGlu7* mRNA also was accumulated in the pod and flower bud. We were unable to detect mRNA from *SGlu1*, *SGlu3*, *SGlu8*, *SGlu9* (pseudogenes), or *Sglu11* genes on RNA

```
          ------>                                                                  Δ 77
SGlu8     GSGVCYGMMGDNLPPANEVVSLYKSNDIMRMRIYNPDQAALQALGISGIEILG-VLHQDLQGLATN-ASTAQQWVQS
SGlu1     GSGVCYGMMGDNLPPANEVVSLYKSNDIMRMRIYNPDQAALQALGNSGIELILGVLHQDLQGLATN-ASTAQQWVQS
SGlu4     GSGVCYGMLGNNLPSANEVIGLYRSNNIRRMRLYDPNQAALEALRNSGIELILGVPNSDLQGLATN-PDTSRQWVQK
SGlu12    --GVCYGMLGNNLPSANDVIGLYRSNNIKRMRLYDPNQAALEALRNSGIELILGVPNSDLQGLATN-PDTSRQWVQK
ClassI    SIGVCYGMLGNNLPNHWEVIQLYKSRNIGRLRLYDPNHGALQALKGS---NILGLPNSDVKHIASG-MEHARWWVQK
SGlu2     GSGVCYGRLGNNLPTPQEVVALYNQANIRRMRIYGPSPEVLEALRGSNIELLLDIPNDNLRNLASS-QDNANKWVQD
ClassIII  QAGVCYGRQGNGLPSPADVVSLCNRNNIRRMRIYDPDQPTLEALRGSNIELMLGVPNPDLENVAAS-QANADTWVQN
SGlu3     GSGVCYGVLGNNLPSRQEVVDLYKTNGIGRMRIYYPDEEALQALRGSGIELIMDVAKETLQSMTD--PNAATDWVNK
SGlu6     IREYVTGVLGNNLPSRQEVVDLYKTNGIGRMRIYYPDEEALQALRGSGIELIMDVAKETLQSMTD--PNAATDWVNK
SGlu10    GSGVCYGVIGDNLPSRQEVVDLYKTNGIGRMRIYYPDEEALQALRGSGIELIMDVAKETLQSLTD--SNAATDWVNK
SGlu5     GSGVCYGGNGNNLPTKQAVVDLYKSNRIGKIRLYYPDEGVLQALRGSNIEVILGVPNDQLQSLTN--AGAATNWVNK
SGlu11    GSGVCYGGNGNNLPTKQAVVDLYKSNRIGKIRLYYPDEGVLQALRGSNIEVILGVPNDQLHSLTN--AGAATNWVNK
ClassII   -IGVCYGKHANNLPSDQDVINLYNANGIRKMRIYNPDTNVFNALRGSNIEIILDVPLQDLQSLTD--PSRANGWVQD
ClassIV   NIGVCYGKIANNLPSEQDVINLYKANGIRKMRIYNSDTNIFKSLNGSNIEIILDVPNQDLEALAN--SSIANGWVQD
SGlu7a    GIRVCYGRSADDLPTPDKVAQLVQLHKIKYVRIYDSNIQVLKAFANTGIELMIGVPNSDLLSFSQF-QSNADSWLKN
SGlu7b    GSGVCYGRSADDLPTPDKVAQLVQLHKIKYVRIYDSNIQVLKAFANTGIELMIGVPNSDLLSFSQF-QSNADSWLKN
ClassV    AVGVCYGRVGTNLPPPSEAINLIKSIGVSRIRLFNPDPEALQPFAGTGIELLVGVPNEILPTLANSPVTISMEWLQT
            .:: *    .** .   : * ::: ..  *::.:  :    :  :    :  .:   :   :  ::  :  *

                           ---                                                    154
SGlu8     NVLNFWPSVKIKHVVVGNEINPVGSSSEFAQYVLPAIQNIYQAIRAQGLQDLIKVTTAIDMTLLGN-SYPPSQSYFR
SGlu1     NVLNFWPSVKIKHVVVGNEINPVGSSSEFAQYVLPAIQNIYQAIRAQGLQDLIKVTTAIDMTLLGN-SYPPSQSYFR
SGlu4     NVLNFWPSVKIKYVAVGNELSPVGRSSSVAQYVLPAIQNVYQAIRAQGLHDQIKVSTSIDMTLIGN-SFPPPQGSFR
SGlu12    NVXNXWPSVKIKYVAVGNEVSPVGGSSSVAQYVLPAIQNVYQAIXAQGLXDQIKVSTXIDMTLIGN-SFPPSQGSFR
ClassI    NVKDFWPDVKIKYIAVGNEISPVTGTSYLTSFLTPAMVNIYKAIGEAGLGNNIKVSTSVDMTLIGN-SYPPSQGSFR
SGlu2     NIKNYANNVRFRYVSVGNEVKPEHS---FAQFLVPALENIQRAISNAGLGNQVKVSTAIDTGALAE-SFPPSKGSFK
ClassIII  NVRNYG-NVKFRYIAVGNEVSPLNENSKYVPVLLNAMRNIQTAISGAGLGNQIKVSTAIETGLTTD-TSPPSNGRFK
SGlu3     YVTAYSQDVNFKYI-VGNEIHPNTN---EAQYILSAMTNIQNAISSANLQ--IKVSTAIDSTFIAPPSYPPNDAVFT
SGlu6     YVTAYSQDVNFKYIAVGNEIHPNTN---EAQYILSAMTNIQNAISSANLQ--IKVSTAIDSTFIAPPSYPPNDAVFT
SGlu10    YVTPYSQDVNFKYIAVGNEIHPNTN---EAQYILSAMTNIQNAISSANLQ--IKVSTAIDSTLITN-SYPPNDGVFT
SGlu5     YVKAYSQNVKFKYIAVGNEIHPGDS---LAGSVLPALENIQKAISAANLQGQMKVSTAIDTTLLGN-SYPPKDGVFS
SGlu11    YVKAYSQNVKFKYIAVGNEIHPGDS---LAGSVLPALENIQKAISAANLQGQMKVSTAIDTTLLGN-SYPPKDGVFS
ClassII   NIINHFPDVKFKYIAVGNEVSPGNNG-QYAPFVAPAMQNVYNALAAAGLQDQIKVSTATYSGILAN-TYPPKDSIFR
ClassIV   NIRSHFPYVKFKYISIGNEVSPSNNG-QYSQFLLHAMENVYNALAAAGLQDKIKVTTATYSGLLAN-TYPPKDSIFR
SGlu7a    SVLPYYPATKIAYITVGAEVTESPNN--ASSFVVPAMTNVLTALKKLGLHKKIKVSSTHSLGVLSR-SFPPSAGAFN
SGlu7b    SVLPYYPATKIAYITVGAEVTESPNN--ASSFVVPAMTNVLTALKKLGLHKKIKVSSTHSLGVLSR-SFPPSAGAFN
ClassV    NIFAHVSPPQVKYLAVGNEIFLKDP--FYSPHIVPAISNLYQALQTLGLATTIKLSSSHASTILSN-SYPPSSGVFN
            .. :: :        * *:            :  *: *:  *:   .*   :*:::          : **  . *

                                                                                 231
SGlu8     TDVR-SYLDPIIGYLVYANAPLLANVLPYFSYSDNP-IDISLSYALFNSTNVVV-WDGQYGYQNLFDAMLDAVHVAI
SGlu1     TDVR-SYLDPIIGYLVYANAPLLANVLPYFSYSNNP-IDISLSYALFNSTNVVV-WDGQYGYQNLFDAMLDAVHVAI
SGlu4     GDV--SYLDPIIGYLVYANAPLLVNVYPYFSYTGNP-RDISLPYALFTAPNVVV-WDGQYGYQNLFDAMLDSVHAAI
SGlu12    GDVR-SYIDPIIGYLVYANAPLLVNVYPYFSYTGNP-RDISLPYALFTAPNVVV-WDGQYGYQNLFDAMLDSVHAAI
ClassI    NDAR-WFTDPIVGFLRDTRAPLLVNIYPYFSYSGNP-GQISLPYSLFTAPNVVV-QDGSRQYRNLFDAMLDSVYAAL
SGlu2     SDYRGAYLDGVIRFLVNNNAPLMVNVYSYFAYTANP-KDISLDYALFRSPSVVV-QDGSLGYRNLFDASVDAVYAAL
ClassIII  DDVR-QFIEPIINFLVTNRAPLLVNLYPYFAIANN--ADIKLEYALFTSSEVVV-NDNGRGYRNLFDAILDATYSAL
SGlu3     SDAE-PYVKPIIDFLVRNEAPLLANVYPYFAYANDQQNSIPLAYALFTQQ-----GNNDAGYQNLFDAMLDSIYAAV
SGlu6     SDAE-PYVKPIIDFLVRNEAPLLANVYPYFAYANDQQNSIPLAYALYTQQ-----GNNDAGYQNLFDAMLDSIYAAV
SGlu10    SDAE-PYIKPIINFLVSNGAPILANVYPYFAYANDQ--SIPLAYALFTQQ-----GNNDVGYQNLFDAMLDSIYAAL
SGlu5     SSAS-SYIRPIVNFLARNGAPLLANVYPYFAYVNNQQ-SIGLDYALFTKH-----GNNEVGYQNLFDALLDSLYAAL
SGlu11    SSAS-SYIRPIVNFLARNGAPLLANVYPYFAYVNNQQ-SIGLDYALFTKH-----GNNEVGYQNLFDALLDSLYAAL
ClassII   GEFN-SFINPIIQFLVQHNLPLLANVYPYFGHIFNT-ADVPLSYALFTQQ-----EANPAGYQNLFDALLDSMYFAV
ClassIV   EEFK-SFINPIIEFLARNNLPLLANIYPYFGHIYNT-VDVPLSYALFNQQ-----GTNSTGYQNLFDALLDSIYFAV
SGlu7a    SSHA-HFLKPMLEFLAENQSPFMIDIYPYYAHRDSR-SKVSLDYALFDASSEVIDPNTGLLYTNMFDAQIDAIYFAL
SGlu7b    SSHA-HFLKPMLEFLAENQSPFMIDIYPYYAYRDSR-SKVSLDYALFDASSEVIDPNTGLLYTNMFDAQIDAIYFAL
ClassV    STIR-PFLLPFLQFLRHTSSPLMVNVYPFFAYINNP-QYVSLDHAVFRSS--YVEYDQNLAYDNMFDASIDAFVYAM
                :   .: :*     *:: :: .::.   .    : * ::::             * *:*** :*:   *:
```

Figure 1.—Alignment of deduced amino acid sequences of β-1,3-glucanases from soybean. Arrows indicate locations of PCR primers used to amplify β-1,3-glucanase sequences. Dashed regions indicate gaps in sequences introduced to maximize alignment. Asterisks represent conserved residues. Conserved tryptophans are marked by triangles. Identical or functionally conserved amino acids surrounding acidic residues are marked with hyphens on the top. Periods and colons represent different degrees of conservative substitution. Tobacco classes I, II, III, IV, and V amino acid sequences were included for comparison. GenBank database accession numbers for these genes are as follows: class I, M20619; class II, M60462; class III, X54456; class IV, X54430; and class V, Z28697.

```
                                  Δ 251
                                -   --
SGlu8          DNTGIGYVEVVVSEWGWPNS
SGlu1          DNTGIGYVEVVVSERGWPNS
SGlu4          DNTKIGYVEVVVSESGWPN-
SGlu12         DNTKIGYVEVVVSESGWPNS
ClassI         ERSGGASVGIVVSESGWPSA
SGlu2          EKAGGGSLNIVVSEWGWPNS
ClassIII       EKASGSSLEIVVSESGWPSA
SGlu3          EKVGASNLQIVVSESGWPNS
SGlu6          EKVGASQFADSGFEKRWPNS
SGlu10         EKVGASNLQIVVSEG-WPNS
SGlu5          EKVGAPNVKVVVSECGWPNS
SGlu11         EKVGAPNVKVVVSECGWPNS
ClassII        EKAGGQNVEIIVSESGWPSE
ClassIV        EKAGGPNVEIIVSESGWPSE
SGlu7a         MALDFRTIKVMVTECGWPNS
SGlu7b         MALNFRTIKVMVTECGWPNS
ClassV         EKEGFEGIPVMVTETGWPTA
                       .     *  **.
                          <---------
```

Figure 1.—*Continued.*

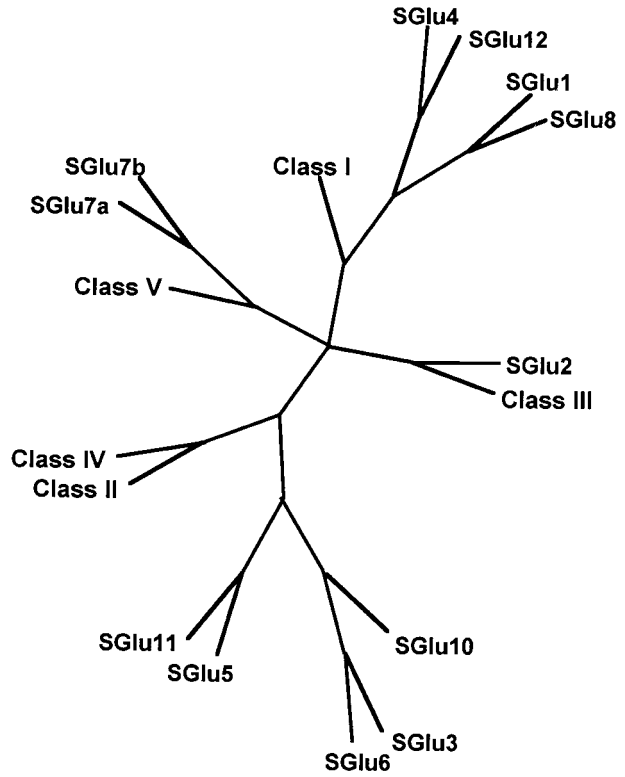

Figure 2.—Phylogenetic analysis of selected plant β-1,3-glucanases. Deduced amino acid sequences were aligned, and an unrooted tree was established using parsimony analysis. Tobacco classes I, II, III, IV, and V are included for comparison.

gel blots. The lack of detectable mRNA corresponding to these four classes suggests that these genes were not expressed, expressed at low levels, or expressed at a developmental stage or condition not tested in this study. In addition, it appears that *SGlu5*, *SGlu6*, and *Sglu10* probes hybridize to mRNA of different sizes. This may indicate multiple starts or stops of transcription, differential processing of transcripts, or transcripts from multiple genes.

DISCUSSION

The goal of this study was an analysis of β-1,3-glucanase gene families in the soybean. Through exhaustive cross-hybridizations and sequence analyses, we identified and classified β-1,3-glucanase gene families of moderate size into 12 classes. We then analyzed sequence relationships among these genes and determined their genomic locations.

Using an F_2 population, 45 restriction fragment length polymorphisms were mapped with 14 restriction enzymes and 12 β-1,3-glucanase probes. We resolved 34 distinct β-1,3-glucanase loci on five different linkage groups. Some loci contained clusters of β-1,3-glucanase genes. It is likely that at least some genes will be tandemly arranged in a cluster, as observed with other multigene families (Yamaguchi-Shinozaki *et al.* 1989; Sutliff *et al.* 1991; Kanazin *et al.* 1996; Sullivan *et al.* 1996). Tandem arrays provide a template for recombination events, including unequal crossing over and gene conversion, which can lead to expansion and further diversification of the sort that is apparent among the classes we have cloned (for review see Maeda and Smithies 1986; Clegg *et al.* 1997).

The distribution of β-1,3-glucanase genes may protect against catastrophic losses of β-1,3-glucanase genes by unequal crossing over or gene slippage that might occur if β-1,3-glucanase genes were clustered at a single locus (Dover 1993). The existence of multiple β-1,3-glucanase genes might also promote the development and maintenance of structural and functional diversity in the β-1,3-glucanase gene families.

Our data show that some β-1,3-glucanase genes are organized into duplicated regions in the genome (Figure 3). These gene loci are probably related evolutionarily and probably have arisen during evolution via duplications of large chromosomal regions. However, retrotransposition and duplication of individual genes also might have occurred during the evolution of β-1,3-glucanase gene families since *SGlu9* mapped as a single locus on linkage groups B1 and J.

In higher plants, β-1,3-glucanases are encoded by gene families of considerable complexity (Linthorst *et al.* 1990; Xu *et al.* 1992). The significance of the gene multiplicity is unclear. Such multiplicity could provide more flexibility for spatial and temporal regulation of diverse functions of β-1,3-glucanase: microsporogenesis, pollen tube growth, senescence, and disease resistance described previously, or "pathogenesis-related" β-1,3-glucanases might be expressed either constitutively or inducibly, and there may be requirements for different

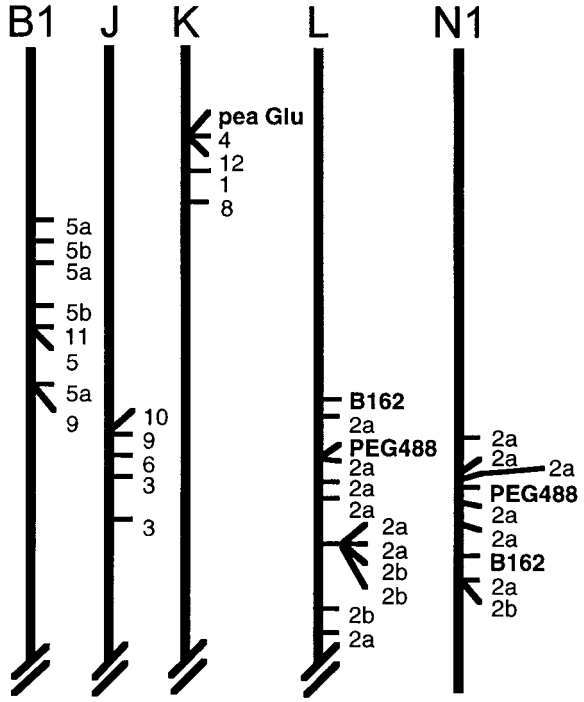

Figure 3.—Distribution of β-1,3-glucanases on soybean genetic map. *SGlu* loci indicated by family designators (1–12) are shown on right of each linkage group. B162 and PEG488 are previously mapped RFLP markers. B1, J, K, L, and N1 are the names of the linkage groups (Shoemaker *et al.* 1996b).

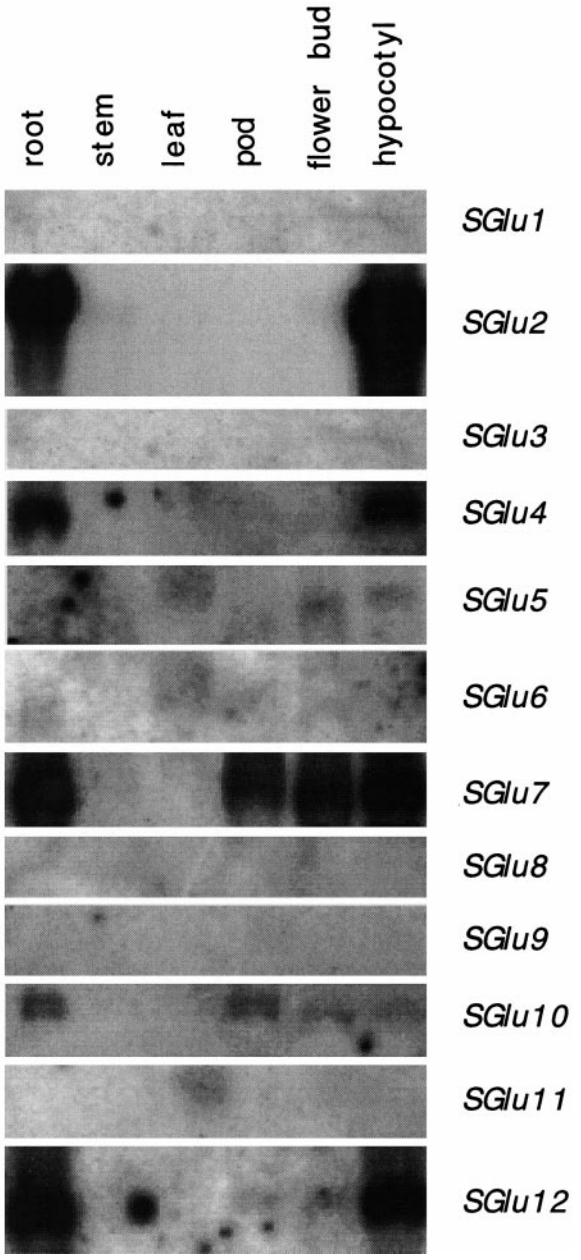

Figure 4.—mRNA accumulation patterns from 12 classes of β-1,3-glucanase genes. RNA gel blot hybridization analysis illustrates the pattern of β-1,3-glucanase mRNA accumulation in young roots, stems, leaves, pods, flower buds, and hypocotyls. mRNA transcripts are ~1.2 kb.

responses in different tissues. It may also reflect functional differences between closely related proteins. Multiplicity of β-1,3-glucanase functions might confer advantages to plants by providing several lines of defense against invading microorganisms. Also, the diversity of β-1,3-glucanases, as well as their organ specificity and developmental and differential expression patterns, may indicate that this enzyme has additional, as yet unidentified, biological functions in plant growth and development.

In an effort to group 12 classes of soybean β-1,3-glucanases into previously described glucanase classes (Shinshi *et al.* 1988; Linthorst *et al.* 1990; Ori *et al.* 1990; Payne *et al.* 1990; Bucciaglia and Smith 1994), we used parsimony analysis to determine their relationship (Figure 2). Our results are consistent with classifications reported previously (classes I–V; Figure 2) on the basis of sequence similarity. Whether this classification accurately reflects functionally distinct classes of β-glucanases in soybean remains to be demonstrated.

Sequence and phylogenetic analyses of the β-1,3-glucanase genes showed that members of the same β-1,3-glucanase class, as well as closely related classes, often mapped to the same linkage group or to duplicated regions of linkage groups. All class I β-1,3-glucanases mapped on linkage group K, all class III β-1,3-glucanases mapped to paralogous regions on linkage groups L and N1, and all class II and IV β-1,3-glucanases mapped on linkage groups B1 and J. Other examples of genes with similar functions and expression patterns that reside on the same linkage group can be found among gene families involved in the regulation of floral identity and cell differentiation (Pickett and Meeks-Wagner 1995) and disease-resistance genes (Kanazin *et al.* 1996).

The multigene families encoding the β-1,3-glucanases are large; all the member genes clearly have a common ancestral origin but have undergone considerable divergence such that individual genes encoding proteins share 11.1–81.5% amino acid identity. Three classes of

tobacco β-1,3-glucanase cDNA clones differ at only 18 positions in 1055 bases of overlapping sequence (Shinshi *et al.* 1988). The high homology level of the tobacco β-1,3-glucanases is maintained by intergenomic DNA exchange between genes (Sperisen *et al.* 1991). In the soybean, 2 classes of β-1,3-glucanase genes (*SGlu8* and *SGlu9*) differ at only 6 positions in 700 bases. How this high level of homology is maintained is unknown.

Our results indicate that in the soybean, pathogen invasion is not a prerequisite for expression of some of the β-1,3-glucanase genes. *SGlu2*, *SGlu4*, *SGlu7*, and *SGlu12* mRNA were accumulated in young roots and hypocotyl. These genes could represent a form of defense against possible threats from rhizosphere microbes (Memelink *et al.* 1990). These results do not preclude a role for these β-1,3-glucanases in defense against microorganism infection. Preemptive or proactive expression of the enzymes (independent of microbial attack) could provide some measure of insurance against infection during the crucial phases of the life cycle in which plant tissues are susceptible to pathogen attack (Knogge *et al.* 1987).

The 12 classes of β-1,3-glucanase genes are distributed in clusters, and two of them (*Sglu 4* and *Sglu 12*; Figure 3) mapped at the same location as a known pathogen-induced gene (the β-1,3-glucanase gene expressed in pea pods challenged with pathogen *Fusarium solani* f. sp. mapped to this location; Chang *et al.* 1992). Our findings demonstrate that β-1,3-glucanase genes with similar sequences are clustered. Duplications of large chromosomal domains followed by extensive gene duplication and divergence and duplication of individual genes are involved in the evolution of the β-1,3-glucanase gene families.

We thank Kim Lewers for help with mapping and C. Baldwin for help with screening the BACs library; D. Grant and Laura Marek for helpful discussions during the course of this work; L. Hadwiger for pea β-1,3-glucanase cDNA clone; Y. Takeuchi for PEG488 cDNA clone from soybean; and J. T. Colbert, A. G. Smith, and M. P. Timko for critically reviewing the manuscript. This is Journal Paper J-17722 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa (Projects 3236 and 3352), and from the U.S. Department of Agriculture (USDA), Agricultural Research Service, Corn Insect and Crop Genetics Research Unit. The mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by Iowa State University or the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

LITERATURE CITED

Abeles, F. B., R. P. Bosshart, L. E. Forrence and W. H. Habig, 1970 Preparation and purification of glucanase and chitinase from bean leaves. Plant Physiol. **47:** 129–134.

Beerhues, L., and E. Kombrink, 1994 Primary structure and expression of mRNAs encoding basic chitinase and 1,3-β-glucanase in potato. Plant Mol. Biol. **24:** 353–367.

Beffa, R. S., R.-M. Hofer, M. Thomas and F. Meins, Jr., 1996 Decreased susceptibility to viral disease of β-1,3-glucanase-deficient plants generated by antisense transformation. Plant Cell **8:** 1001–1011.

Boller, T., 1987 Hydrolytic enzymes in plant disease resistance, pp. 385–413 in *Plant-Microbe Interactions: Molecular and Genetic Perspectives*, Vol. 2, edited by T. Kosuge and E. W. Nester. Macmillan, New York.

Brederode, F. T., H. J. M. Linthorst and J. F. Bol, 1991 Differential induction of acquired resistance and PR gene expression in tobacco by virus infection, ethephon treatment, UV light and wounding. Plant Mol. Biol. **17:** 1117–1125.

Bucciaglia, P. A., and A. G. Smith, 1994 Cloning and characterization of *Tag1*, a tobacco anther β-1,3-glucanase expressed during tetrad dissolution. Plant Mol. Biol. **24:** 903–914.

Castresana, C., F. de Carvalho, G. Gheysen, M. Habets, D. Inzé *et al.*, 1990 Tissue-specific and pathogen-induced regulation of a *Nicotiana plumbaginifolia* β-1,3-glucanase gene. Plant Cell **2:** 1131–1143.

Chang, M. M., L. A. Hadwiger and D. Horovitz, 1992 Molecular characterization of a pea β-1,3-glucanase induced by *Fusarium solani* and chitosan challenge. Plant Mol. Biol. **20:** 609–618.

Chomczynski, P., and N. Sacchi, 1987 Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162:** 156–159.

Clegg, M. T., P. M. Cummings and M. L. Durbin, 1997 The evolution of plant nuclear genes. Proc. Natl. Acad. Sci. USA **94:** 7791–7798.

Dong, X., M. Mindrinos, K. R. Davis and F. M. Ausubel, 1991 Induction of *Arabidopsis* defense genes by virulent and avirulent *Pseudomonas syringae* strains and by a cloned avirulence gene. Plant Cell **3:** 61–72.

Dover, G. A., 1993 Evolution of genetic redundancy for advanced players. Curr. Opin. Genet. Dev. **3:** 902–910.

Felix, G., and F. Meins, Jr., 1986 Developmental and hormonal regulation of β-1,3-glucanase in tobacco. Planta **167:** 206–211.

Fincher, G. B., P. A. Lock, M. M. Morgan, K. Lingelbach, R. E. H. Wettenhall *et al.*, 1986 Primary structure of the (1,3; 1,4)-β-d-glucan 4-glucohydrolase from barley aleurone. Proc. Natl. Acad. Sci. USA **83:** 2081–2085.

Kanazin, V., L. F. Marek and R. C. Shoemaker, 1996 Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA **93:** 11746–11750.

Kauffmann, S., M. Legrand, P. Geoffroy and B. Fritig, 1987 Biological function of 'pathogenesis-related' proteins: four PR proteins of tobacco have 1,3-β-glucanase activity. EMBO J. **6:** 3209–3212.

Keim, P., T. C. Olson and R. C. Shoemaker, 1988 A rapid protocol for isolating soybean DNA. Soybean Genet. Newslett. **15:** 150–152.

Keim, P., B. W. Diers, T. C. Olson and R. C. Shoemaker, 1990 RFLP mapping in soybean: association between marker loci and variation in quantitative traits. Genetics **126:** 735–742.

Kosambi, D. D., 1944 The estimation of map distance from recombination values. Ann. Eugen. **12:** 172–175.

Knogge, W., E. Kombrink, E. Schmelzer and K. Hahlbrock, 1987 Occurrence of phytoalexins and other putative defense-related substances in uninfected parsley plants. Planta **171:** 279–287.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer program for constructing genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Leah, R., H. Tommerup, I. Svendsen and J. Mundy, 1991 Biochemical and molecular characterization of three barley seed proteins with antifungal properties. J. Biol. Chem. **266:** 1564–1573.

Li, C.-D., P. Langridge, R. C. M. Lance, P. Xu and G. B. Fincher, 1996 Seven members of the (1-3)-β-glucanase gene family in barley (*Hordeum vulgare*) are clustered on the long arm of chromosome 3 (3HL). Theor. Appl. Genet. **92:** 791–796.

Linthorst, H. J. M., L. S. Melchers, A. Mayer, J. S. C. Van Roekel, B. J. C. Cornelissen *et al.*, 1990 Analysis of gene families encoding acidic and basic β-1,3-glucanases of tobacco. Proc. Natl. Acad. Sci. USA **87:** 8756–8760.

Maeda, N., and O. Smithies, 1986 The evolution of multigene families: human haptoglobin genes. Annu. Rev. Genet. **20:** 81–108.

MacGregor, E. A., and G. M. Ballance, 1991 Possible secondary structure in plant and yeast β-glucanases. Biochem. J. **274:** 41–44.

Marek, L. F., and R. C. Shoemaker, 1996 Construction and size characterization of a bacterial artificial chromosomal (BAC) library from soybean. Soybean Genet. Newslett. **23:** 126–129.

Mauch, F., L. A. Hadwiger and T. Boller, 1988 Antifungal hy-

drolases in pea tissue. I. Purification and characterization of two chitinases and two β-1,3-glucanases differentially regulated during development and in response to fungal infection. Plant Physiol. **87:** 325–333.

Memelink, J., H. J. M. Linthorst, R. A. Schilperoort and J. H. C. Hoge, 1990 Tobacco genes encoding acidic and basic isoforms of pathogenesis-related proteins display different expression patterns. Plant Mol. Biol. **14:** 119–126.

Muthukrishnan, S., B. S. Gill, M. Swegle and G. R. Chandra, 1984 Structural genes for α-amylases are located on barley chromosomes 1 and 6. J. Biol. Chem. **259:** 13637–13639.

Ori, N., G. Sessa, T. Lotan, S. Himmelhoch and R. Fluhr, 1990 A major stylar polypeptide (sp41) is a member of the pathogenesis-related proteins superclass. EMBO J. **9:** 3429–3436.

Payne, G., E. Ward, T. Gaffney, P. Ahl Goy, M. Moyer *et al.*, 1990 Evidence for a third structural class of β-1,3-glucanase in tobacco. Plant Mol. Biol. **15:** 797–808.

Pickett, F. B., and D. R. Meeks-Wagner, 1995 Seeing double: appreciating genetic redundancy. Plant Cell **7:** 1347–1356.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schröder, M., K. Hahlbrock and E. Kombrink, 1992 Temporal and spatial patterns of 1,3-β-glucanase and chitinase induction in potato leaves infected by *Phytophthora infestans.* Plant J. **2:** 161–172.

Shinshi, H., H. Wenzler, J.-M. Neuhaus, G. Felix, J. Hofsteenge *et al.*, 1988 Evidence for N- and C- terminal processing of a plant defense-related enzyme: primary structure of tobacco prepro-β-1,3-glucanase. Proc. Natl. Acad. Sci. USA **85:** 5541–5545.

Shoemaker, R. C., K. Polzin, J. Labate, J. Specht, E. C. Brummer *et al.*, 1996a Genome duplication in soybean (*Glycine* subgenus *soja*). Genetics **144:** 329–338.

Shoemaker, R. C., K. M. Polzin, L. L. Lorenzen and J. E. Specht, 1996b Molecular genetic mapping of soybean, pp. 37–56 in *Soybean: Genetics, Molecular Biology and Biotechnology*, edited by D. P. S. Verma and R. C. Shoemaker. CAB International, Wallingford, UK.

Simmons, C. R., 1994 The physiology and molecular biology of plant 1,3-β-d-glucanases and 1,3;1,4-β-d-glucanases. Crit. Rev. Plant Sci. **13:** 325–387.

Sperisen, C., J. Ryals and F. Meins, Jr., 1991 Comparison of cloned genes provides evidence for intergenomic exchange of DNA in the evolution of a tobacco glucan endo-1,3-β-glucosidase gene family. Proc. Natl. Acad. Sci. USA **88:** 1820–1824.

Sullivan, S. L., M. C. Adamson, K. J. Ressler, C. A. Kozak and L. B. Buck, 1996 The chromosomal distribution of mouse odorant receptor genes. Proc. Natl. Acad. Sci. USA **93:** 884–888.

Sutliff, T. D., N. Huang, J. C. Litts and R. L. Rodriguez, 1991 Characterization of an α-amylase multigene cluster in rice. Plant Mol. Biol. **16:** 579–591.

Swofford, D. L., 1991 PAUP: phylogenetic analysis using parsimony, Version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign, IL.

Takeuchi, Y., M. Yoshikawa, G. Takeba, K. Tanaka, D. Shibata *et al.*, 1990 Molecular cloning and ethylene induction of mRNA encoding a phytoalexin elicitor-releasing factor, β-1,3-endoglucanase, in soybean. Plant Physiol. **93:** 673–682.

Thompson, J. D., D. G. Higgins and T. J. Gibson, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:** 4673–4680.

Van den Bulcke, M., G. Bauw, C. Castresana, M. Van Montagu and J. Vandekerckhove, 1989 Characterization of vacuolar and extracellular β-(1,3)-glucanases of tobacco: evidence for a strictly compartmentalized plant defense system. Proc. Natl. Acad. Sci. USA **86:** 2673–2677.

Varghese, J. N., T. P. J. Garrett, P. M. Colman, L. Chen, P. B. Høj *et al.*, 1994 Three-dimensional structures of two plant β-glucan endohydrolases with distinct substrate specificities. Proc. Natl. Acad. Sci. USA **91:** 2785–2789.

Ward, E. R., G. B. Payne, M. B. Moyer, S. C. Williams, S. S. Dincher *et al.*, 1991 Differential regulation of β-1,3-glucanase messenger RNAs in response to pathogen infection. Plant Physiol. **96:** 390–397.

Xu, P., J. Wang and G. B. Fincher, 1992 Evolution and differential expression of the (1-3)-β-glucan endohydrolase-encoding gene family in barley, *Hordeum vulgare.* Gene **120:** 157–165.

Yamaguchi-Shinozaki, K., J. Mundy and N. H. Chua, 1989 Four tightly linked *rab* genes are differentially expressed in rice. Plant Mol. Biol. **14:** 29–39.

Communicating editor: V. Sundaresan