Theor Appl Genet (1998) 96: 1186–1192

E. Grosselindemann · M. Robertson
J. A. Wilmer · P. M. Chandler

Genetic variation in pea (*Pisum*) dehydrins: sequence elements responsible for length differences between dehydrin alleles and between dehydrin loci in *Pisum sativum* L.

Received: 12 June 1997 / Accepted: 29 October 1997

Abstract The electrophoretic patterns of dehydrins extracted from mature seeds of a range of pea (*Pisum*) species revealed extensive variation in dehydrin polypeptide mobility. Variation was also observed among lines of *P. sativum*. Crosses between lines with different dehydrin electrophoretic patterns produced F_1 seeds with additive patterns, and segregation in the F_2 generation was consistent with a 1:2:1 ratio, indicating allelic variation at each of two dehydrin loci (*Dhn2*, *Dhn3*). Genetic linkage was observed between *Dhn2* and *Dhn3*, and the segregation ratios indicated preferential transmission of one allele at the *Dhn3* locus. Dehydrin cDNA clones were characterised that encoded the allelic variants at *Dhn2* and *Dhn3*. Their deduced amino-acid sequences were very similar to each other as well as to the product of the *Dhn1* locus reported previously. Comparisons were made between the sequences of allelic variants at a single locus, and between the products of different loci. Differences in the electrophoretic mobilities between allelic variants at *Dhn2* and *Dhn3* were associated with differences in polypeptide length resulting principally from tandem duplications of 21 (*Dhn2*) or 24 (*Dhn3*) amino-acid residues. These duplications accounted for much of the difference in length between dehydrins encoded by the different loci. The conserved core of one of the duplicated regions varied in copy number, and small insertions/deletions of amino acids near this core also contributed to length variation both between allelic forms and between loci. Dehydrins possess characteristic highly conserved amino-acid sequence motifs, yet vary considerably in length. Mechanisms involving sequence duplication appear to be responsible for generating the length differences observed between allelic variants as well as between the products of different loci.

Key words Allelic variation · Genetic variation · Dehydrin · Pea (*Pisum sativum* L.)

Communicated by P. Langridge

E. Grosselindemann · M. Robertson (✉)
J. A. Wilmer[1] · P. M. Chandler
CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia
Fax: 61 2 6246 5000
E-mail: masumi.robertson@pi.csiro.au

Present address:
[1] DLO-Research Institute for Agrobiology and Soil Fertility, PO Box 14, 6700 AA Wageningen, The Netherlands

Introduction

Dehydrins are proteins that are characterised by unique conserved amino-acid sequence motifs, and by their accumulation in vegetative parts of plants experiencing water deficit (Close et al. 1989). In many plant species they also accumulate in seeds during development and maturation, as one sequence class among the group-II LEA (*l*ate *e*mbryogenesis *a*bundant) proteins (Dure et al. 1989). Abscisic acid (ABA) levels are elevated in plants experiencing water deficit, and there is evidence in several species that ABA is an important regulator of dehydrin gene expression (Robertson and Chandler 1992; Bray 1993). It has been postulated that dehydrin proteins might be involved in stress tolerance, although direct evidence for such a role is still lacking (Skriver and Mundy 1990; Bray 1993; Chandler and Robertson 1994). However, one other member of the LEA protein family has recently been shown to confer osmotic-stress tolerance and dehydration tolerance in transgenic rice plants (Xu et al. 1996).

We have previously characterised a dehydrin in pea (*Pisum sativum* L.) corresponding to a M_r-23 000 polypeptide extracted from mature seeds (Robertson and Chandler 1992). Immunoblot analysis using a maize

dehydrin antiserum revealed two additional polypeptides with different electrophoretic mobilities, possibly representing the products of other dehydrin loci. While investigating dehydrin polypeptides in a collection of wild *Pisum* species, as well as in commercially available lines of *P. sativum*, we encountered extensive variation in the electrophoretic mobility of different dehydrin polypeptides. In barley (*Hordeum vulgare* L. 'Himalaya'), length variation between different dehydrins was due to differences in the copy number of a semi-conserved tandemly repeated motif with a core sequence YGQ (Close et al. 1989). However, this sequence is absent from pea dehydrin, and so it was of interest to determine the basis for length variation among dehydrins encoded by different loci. During the course of this investigation it became clear that some of the heterogeneity in dehydrin patterns reflected variation between alleles, rather than between loci. Cloning and sequencing studies provided a basis for understanding this genetic variation.

Materials and methods

Plant material and growth conditions

Seeds of a range of *Pisum* species (*P. abyssinicum*, *P. arvense*, *P. asiaticum*, *P. cinereum*, *P. elatius*, *P. fulvum*, *P. humili*, *P. jomardii*, *P. nepalensis*, *P. tibetanicum*, *P. zeylanicum*), as well as *P. sativum* cv Greenfeast, selection 086, were obtained from Hart Schroeder CSIRO Plant Industry, Canberra (Schroeder 1982). Eleven lines of *P. sativum* were obtained from commercial sources. Subsequently, two lines were studied in detail, 086 from cv Greenfeast, and one of the commercially available lines, Hortico snow pea [cv Mammoth Melting, from Hortico (Aust.) Pty. Ltd., Milperra, NSW 2214]. For convenience these lines will be called 086 and MM, respectively. Plants were grown in a naturally lit glasshouse.

Western-blot analysis

Extraction of polypeptides from mature seeds, electrophoresis and transfer to membranes, and immunoblotting using maize dehydrin antiserum was carried out as previously described (Robertson and Chandler 1992).

RNA extraction, cDNA cloning and DNA sequencing

RNA extraction from developing cotyledons of 086, cDNA library preparation and immuno-screening for dehydrin cDNA clones using the maize dehydrin antiserum has been previously described (Robertson and Chandler 1992). A λZAP cDNA library (Stratagene, LaJolla, Calif.) was prepared by the same method using RNA extracted from cotyledons of dehydration-stressed MM seedlings. MM seeds were germinated and seedlings grown on moist filter paper for 7 days, prior to dehydration treatment for 10 days as previously described (Robertson and Chandler 1992). At this stage the cotyledons contained high levels of dehydrin mRNAs. The MM library was screened with a pea dehydrin cDNA insert (pPsB210) obtained from 086 and with the maize dehydrin antiserum (Robertson and Chandler 1992). Plasmid DNA was prepared and sequenced by the dideoxy chain-termination method using a Deaza sequencing kit (Pharmacia, North Ryde, New South Wales). All sequences were confirmed by reading both cDNA strands except for clone T123. Part of this sequence (nucleotide positions 694–853) could only be read from the 3′ end because of closely related repeated sequences. The same sequence for this region was obtained using three different primers. DNA sequences were analysed by the University of Wisconsin Genetics Computer Group programs (Devereux et al. 1984).

Identification of cDNA clones

Polypeptides encoded by cDNA clones obtained from the 086 library were analysed by coupled in vitro transcription/translation of the clones. Following the addition of a 086 cotyledon-protein extract, the radio-labelled products were fractionated by two-dimensional gel electrophoresis and transferred to a membrane. The positions of 086 dehydrin marker-polypeptides were revealed by reaction with dehydrin antiserum, and the positions of the in vitro translation products by autoradiography (Robertson and Chandler 1992).

Polypeptides encoded by cDNA clones obtained from the MM library were expressed following the procedure of Sambrook et al. (1989). Cultures containing the appropriate plasmid DNA were incubated overnight, with aeration, in LB medium containing ampicillin (50 μg/ml) at 37°C, sub-cultured and incubated until D_{600} = 1.1. Isopropylthiogalactoside was added to a final concentration of 1 mM, and the cultures were further incubated at 37°C for 2.5 h. Cells were harvested by centrifugation, lysed by a freeze-thaw cycle, and polypeptides in the lysate were suspended in 0.5 M NaCl, 20 mM TES-KOH pH 8.0. Heat-stable polypeptides were prepared as described (Close et al. 1989) and analysed by gel electrophoresis and immunoblotting.

Results and discussion

Dehydrin polypeptide patterns in seed extracts of different *Pisum* species

Extracts from mature seeds of a collection of *Pisum* species were electrophoresed, blotted to a membrane, and dehydrin polypeptides visualised following reaction with maize dehydrin antiserum (Fig. 1a). Generally, each species had a distinct pattern consisting of two or three major dehydrin polypeptides with M_r values ranging between 20 000 and 40 000. These data indicate extensive variation in the electrophoretic mobilities of dehydrins from different *Pisum* species.

Pea species are inbreeding, and are expected to have a high degree of homozygosity. The variation in dehydrin polypeptides observed in Fig. 1a is therefore likely to reflect differences between the products of different loci. To study such variation in greater detail, dehydrins could be cloned from the different wild species, and their deduced amino-acid sequences compared. However, previous studies (Robertson and Chandler 1992; M. Robertson, unpublished data) had indicated that within *P. sativum* line 086 the three dehydrin loci were very closely related to each other in sequence. This observation means that difficulties might be encountered in an analysis of dehydrin variation among the wild species based solely on cDNA cloning and sequencing. In comparing two dehydrin sequences from

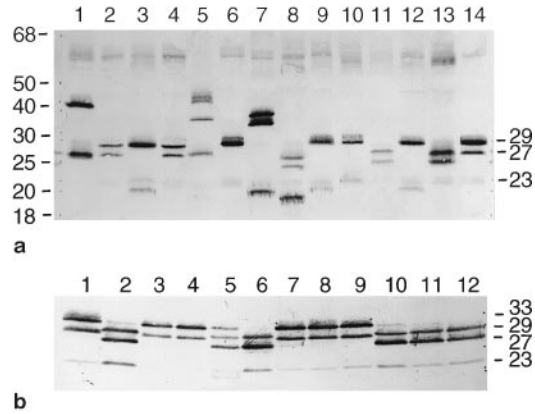

Fig. 1a, b Dehydrin polypeptides in mature seeds of different *Pisum* species (**a**) and of 12 lines of *P. sativum* (**b**) following SDS-PAGE and detected by immunoblotting using a maize dehydrin antiserum. *Lanes* in **a** refer to (*1*) *P. abyssinicum*; (*2*) *P. arvense*; (*3*) *P. asiaticum*; (*4*) *P. cinereum*; (*5*, *6*) *P. elatius*; (*7*) *P. fulvum*; (*8*) *P. humili*; (*9*) *P. jomardii*; (*10*) *P. nepalensis*; (*11*) *P. sativum*; (*12*, *13*) *P. tibetanicum*; (*14*) *P. zeylanicum*. *Lanes* in **b** refer to 11 commercially available lines of "snow" peas, "snap" peas, and "sugar" peas (*lanes 1–11*; line MM = *lane 9*), and line 086 of cv 'Greenfeast' (*lane 12*). M_r values ($\times 10^{-3}$) of dehydrin polypeptides are indicated on the right hand side of both panels, and of marker polypeptides on the left hand side of **a**

different species it may not be possible to determine whether the two sequences correspond to allelic variants at a single *Dhn* locus, or to the products of different *Dhn* loci. This problem will be of particular importance in comparisons between species which have dehydrins of similar size and/or mobility. For example, most *Pisum* species analysed in Fig. 1a had either a single polypeptide, or two polypeptides with similar mobility, at $M_r \approx 30\,000$. A single polypeptide at M_r 29 000 was present in the *P. abyssinicum* extract (lane *1*), whereas two polypeptides of similar M_r were observed in extracts of *P. arvense* (lane *2*); which one, therefore, represents an allelic variant of the *P. arvense* polypeptide, and which one represents the product of a different locus?

To begin an analysis of dehydrin variation in *Pisum* we chose to first examine variation within a single species, *P. sativum*, since by crossing and inheritance studies it should be possible to differentiate between variants at a single locus and variants at different loci. Proteins were extracted from mature seeds of 12 lines of *P. sativum* and immunoblotted as described above. The dehydrin polypeptides formed two major patterns (Fig. 1b). The first pattern, represented by 086 (lane 12 in Fig. 1b) and four other varieties (lanes 2, 6, 10, 11), had three polypeptides with an M_r of 23 000, 27 000 and 29 000 respectively. A different pattern was observed in six other varieties; polypeptides of M_r 23 000 and 29 000 were present, but the M_r-27 000 polypeptide was absent and a new polypeptide of M_r 33 000 was present. The one remaining variety (lane 5) had a pattern which appeared to be a mixture of the two major patterns. There are several explanations that could account for variation in the M_r-27 000 and -33 000 polypeptides: they might represent (1) the products of identical genes but with differences in post-translational modification (either during seed development or during protein extraction) arising from different maternal backgrounds, (2) allelic variants at a single locus, or (3) products of different dehydrin loci.

Inheritance studies of dehydrin patterns

Crossing studies were undertaken to distinguish between the possible sources of genetic variation described above. Reciprocal crosses carried out between lines 086 and MM yielded F_1 seeds with additive dehydrin patterns (data not shown, but see Fig. 2 below), indicating that the M_r-27 000 and -33 000 polypeptides represent distinct gene products, rather than variants due to post-translational modification. Extracts prepared from individual F_2 seeds showed segregation for M_r-27 000/33 000 dehydrin electrophoretic patterns (Fig. 2), and the observed numbers in each category (31, 59, and 20 with the 086, hybrid, and MM patterns respectively) were consistent with a 1 : 2 : 1 ratio ($\chi^2_2 = 2.78$, $0.1 < P < 0.5$). We conclude that the M_r-27 000 and -33 000 polypeptides represent allelic variants at a new dehydrin locus (*Dhn2*, see Table 1), rather than the products of two different loci.

Heterogeneity of the MM dehydrin pattern

In the course of further analysis of dehydrins in line MM we noted differences in electrophoretic patterns between the original seed stock and a new seed stock that was derived from a single plant grown from the original stock. The original extracts (Fig. 1b) were made on ten seeds from the commercial supplier, and showed a dehydrin polypeptide at M_r 29 000. Extracts of the new seed stock lacked the M_r-29 000 dehydrin polypeptide, but instead possessed a new polypeptide

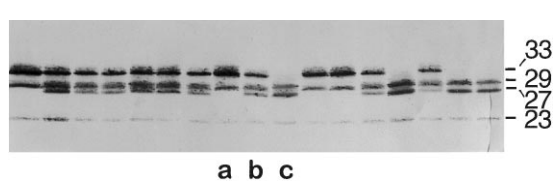

Fig. 2 Segregation of dehydrin electrophoretic patterns in a sample of F_2 seeds from the cross 086 × MM. The dehydrin patterns of individual F_2 seeds were characterised following electrophoresis and immunoblotting using a maize dehydrin antiserum. Three patterns were observed, represented by *lanes a* (MM parent), *b* (additive or F_1), *c* (086 parent). M_r values ($\times 10^{-3}$) of dehydrin polypeptides are indicated

at M_r-22 000 (data not shown, but see Fig. 3 below). Later extracts prepared from individual seeds of the original stock revealed that it was heterogeneous, with eight out of nine seeds having a M_r-23 000/29 000/33 000 dehydrin profile, and one out of nine seeds having a M_r-22 000/23 000/33 000 dehydrin profile (data not shown). We conclude that the new MM seed stock represented a minor variant present in the initial population.

These observations suggested that the M_r-22 000 and -29 000 polypeptides might also represent allelic variants at a third dehydrin locus. A cross between two MM plants, one derived from a seed with a M_r-29 000 pattern and the other from a seed with a M_r-22 000 pattern, produced F_1 seeds with an additive pattern (data not shown, but see Fig. 3 below). In individual F_2 seeds three types of dehydrin pattern were observed; all seeds had the M_r-23 000 and -33 000 polypeptides, but differences were seen in whether seeds had the M_r-22 000 polypeptide, both M_r-22000 and -29 000 polypeptides, or the M_r-29 000 polypeptide (Fig. 3). In a sample of 117 seeds we observed these patterns in 37, 62, and 18 individuals respectively. There was a considerable bias towards one of the original parental types (M_r 22 000), the distribution borders being significantly different from 1:2:1 ($\chi^2_2 = 6.59$, $P = 0.04$). The most likely explanation for any significant deviation from a 1:2:1 ratio is preferential transmission of the M_r-22 000 allele. We conclude that the M_r-22 000 and M_r-29 000 polypeptides represent allelic variants at a further dehydrin locus (*Dhn3*, see Table 1). We note in Fig. 3 that seeds homozygous for the M_r-22 000 dehydrin variant tend to show greater staining in the M_r-23 000 region, but the reason for this is unknown.

Genetic linkage between the *Dhn2* and *Dhn3* loci

Evidence presented above suggests that there is allelic variation at two distinct dehydrin loci (*Dhn2* and *Dhn3*, Table 1) in *P. sativum*, allowing linkage between these loci to be assessed. A cross was made between line 086 and an individual of line MM possessing the M_r-22 000 variant at the *Dhn3* locus. Protein extracted from individual F_1 seeds contained allelic alternatives at *Dhn2* (M_r-27 000 and -33 000) and at *Dhn3* (M_r-22 000 and -29 000) in addition to the M_r-23 000 dehydrin polypeptide (encoded by *Dhn1*) present in both parents (data not shown, but see Fig. 4). Analysis of individual F_2 seeds (Fig. 4) revealed the presence of both parental patterns, the additive (F_1) pattern, but an absence of non-parental associations. In 28 F_2 seeds we observed three seeds with an 086 pattern, ten seeds with an additive pattern, and 15 seeds with the MM pattern (M_r-22 000 variant). This lack of non-parental types leads us to conclude that *Dhn2* and *Dhn3* are closely linked, although there was a marked distortion from the 1:2:1 segregation ratio expected for closely linked loci ($\chi^2_2 = 12.6$, $P < 0.002$), which was again consistent with preferential transmission of the M_r-22 000 allelic variant at *Dhn3*. Using the method of Brown (1980) we estimate with a 95% level of confidence that there is less than 5% recombination between *Dhn2* and *Dhn3*.

Pea dehydrin cDNA clones

Cloning studies were carried out to analyse the allelic variation described above, as well as to analyse variation between loci. The cDNA library from line 086 was screened with a maize dehydrin antiserum. Plaques which expressed polypeptides reactive with the antiserum represented 0.4% of the library, and included three discrete sequence categories. One of these categories, represented by the original clone pPsB12 (Robertson and Chandler 1992), was designated as the *Dhn1* locus. The other two categories were closely related (87–90%) in nucleotide sequence to B12, and were represented by clones pPsB210 and pPsB213 (subsequently referred to as B210 and B213). The MM cDNA library (derived from MM plants expressing M_r-22 000/23 000/33 000 dehydrin polypeptides) was screened both with the maize dehydrin antiserum and with clone B210. About 0.3% of the cDNA clones reacted with either antiserum or clone. Two clones were

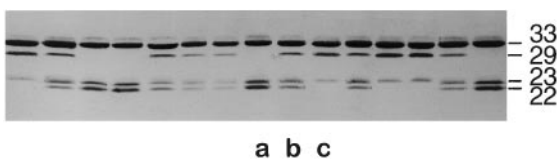

Fig. 3 Segregation of dehydrin electrophoretic patterns in a sample of F_2 seeds from the cross MM (M_r-22 000 variant) × MM (M_r-29 000 variant). The dehydrin patterns of individual F_2 seeds were characterised following electrophoresis and immunoblotting using a maize dehydrin antiserum. Three patterns were observed, represented by *lanes a* (M_r-22 000 parent), *b* (additive or F_1), and *c* (M_r-29 000 parent). M_r values ($\times 10^{-3}$) of dehydrin polypeptides are indicated

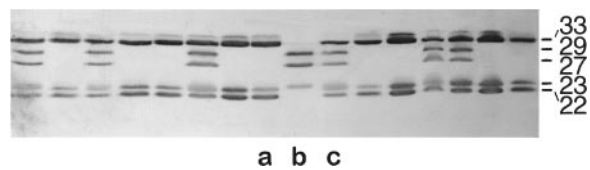

Fig. 4 Segregation of dehydrin electrophoretic patterns in a sample of F_2 seeds from the cross 086 × MM (M_r-22 000 variant). The dehydrin patterns of individual F_2 seeds were characterised following electrophoresis and immunoblotting using a maize dehydrin antiserum. Three patterns were present, represented by *lanes a* (MM parent), *b* (086 parent), and *c* (additive or F_1). M_r values ($\times 10^{-3}$) of dehydrin polypeptides are indicated

characterised, pPsT41 and pPsT123 (referred to subsequently as T41 and T123); these were closely related to each other in nucleotide sequence (90% identity) and to the cDNA clones from the 086 library (81–99% identity).

Relationship of dehydrin clones to dehydrin polypeptides

To relate the four new dehydrin cDNA clones to individual dehydrin polypeptides two approaches were adopted. For clones B210 and B213, radioactively labelled polypeptide products were synthesized in a coupled in vitro transcription/translation reaction, mixed with non-radioactive polypeptides extracted from mature 086 cotyledons, and analysed by two-dimensional gel electrophoresis, immunoblotting and then autoradiography (Fig. 5). First, for each filter, the positions of the M_r-29 000, -27 000 and -23 000 dehydrins from mature cotyledons were revealed using maize dehydrin antiserum (Fig. 5a, b). Then the positions of any radioactive gene products derived from a clone were revealed by autoradiography of the filter, and related to the marked positions of the dehydrin polypeptides. The locations of the B210 (Fig. 5c) and B213 (Fig. 5d) gene products were identical to those of the M_r-27 000 and -29 000 dehydrins respectively, indicating that B210 encodes the M_r-27 000 dehydrin and B213 encodes the M_r-29 000 dehydrin. Clone B12 has previously been shown to encode the M_r-23 000 dehydrin polypeptide (Robertson and Chandler 1992). We conclude that the dehydrins of line 086 undergo no post-translational processing, since the in vitro products co-migrated with in vivo products. The three major dehydrin polypeptides have unique amino-acid sequences, and presumably correspond to three different genetic loci.

For clones T41 and T123, polypeptides were synthesised using an in vivo approach (see Materials and methods), and analysed by SDS-PAGE and immunoblotting. The T41-encoded immunoreactive polypeptides included one with identical mobility to the new M_r-22 000 dehydrin, as well as two products with higher mobility (Fig. 6, lane 1). One of these might result from an alternative translation start at an ATG codon 13 amino-acid residues from the initiator codon, but the nature of the other is unknown. There was also a faint band at M_r 34 000. Clone T123 encoded an immunoreactive polypeptide with the same mobility as the M_r-33 000 dehydrin (Fig. 6, lane 2). There was also a faint band at M_r 34 000, presumably the same as that in lane 1, and probably corresponding to an *Escherichia coli* polypeptide that shows slight cross-reaction with the antiserum. The immunoreactive M_r-22 000 and -33 000 polypeptides, encoded by clones T41 and T123 respectively, were shown by two-dimensional electrophoresis to have identical isoelectric points (pI) to the corresponding dehydrin polypeptides in the MM standard (data not shown). We conclude that clone T41 encodes the M_r-22 000 dehydrin and clone T123 the M_r-33 000 dehydrin.

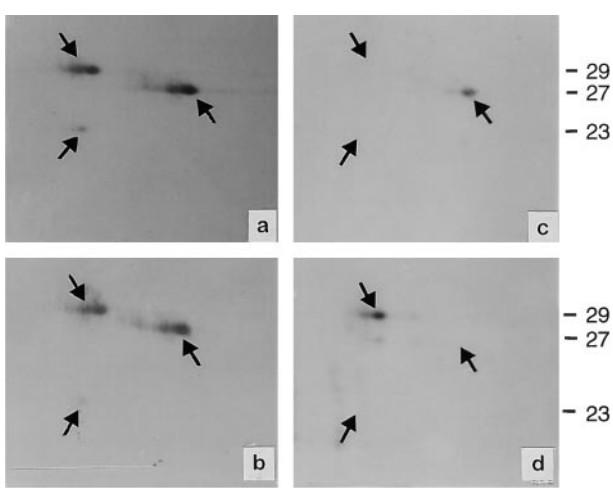

Fig. 5a–d Identification of polypeptides encoded by cDNA clones B210 and B213. Two-dimensional gel electrophoresis and immunoblotting of polypeptides from mature 086 cotyledons mixed with radioactively labelled in vitro transcription/translation products of the cDNA clones. The positions of non-radioactive dehydrin polypeptides on the filters were visualised following reaction with maize dehydrin antiserum (clone B210, panel **a**; clone B213, panel **b**). The locations of radioactive clone products were determined by autoradiography of the same filters (clone B210, panel **c**; clone B213, panel **d**). Only a small portion of the original gel and filter is shown. *Arrows* indicate positions of dehydrin polypeptides detected by the antiserum. M_r values ($\times 10^{-3}$) of 086 dehydrin polypeptides are indicated

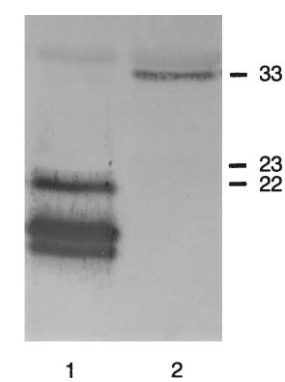

Fig. 6 Identification of dehydrin polypeptides encoded by cDNA clones T41 and T123. Polypeptides were synthesised by *E. coli* cells harboring the appropriate cDNA clone, heated, fractionated by SDS-PAGE and immunoblotted using maize dehydrin antiserum. *Lane 1* polypeptides encoded by clone T41; *lane 2* polypeptides encoded by clone T123. M_r values of MM dehydrin polypeptides are indicated

Table 1 Characteristics of pea dehydrin cDNA clones, the corresponding genetic loci, and properties of the encoded dehydrin proteins

Clone	Pea line	Locus	Properties of encoded dehydrin proteins				
			M_r	kDa	Amino acids	pI	EYVREE repeats
pPsB12	086	*Dhn1*	23 000	20.4	197	6.4	2
pPsB210	086	*Dhn2*	27 000	24.5	232	6.9	4
pPsT123	MM	*Dhn2*	33 000	27.1	255	7.0	5
pPsB213	086	*Dhn3*	29 000	23.9	232	6.5	3
pPsT41	MM	*Dhn3*	22 000	20.8	198	6.4	2

Characteristics of clones and their deduced amino-acid sequences

The properties of the clones, the proteins they encode, and their corresponding genetic loci are shown in Table 1. Each clone contains the entire coding region for dehydrins that range in length from 197 to 255 amino acids. The molecular masses of the deduced polypeptides are generally lower than would be estimated on the basis of electrophoretic mobility. For the three dehydrins in line 086, the difference in mobility between the M_r-27 000 and M_r-29 000 polypeptides must result from effects of the amino-acid sequence, since both polypeptides have the same number of residues (232 amino acids), and the "smaller" M_r-27 000 polypeptide actually has a greater mass than the M_r-29 000 polypeptide (Table 1). We can exclude post-translational processing accounting for the different mobilities, because the in vitro synthesised products co-migrated in two-dimensional electrophoresis with the in vivo products (Fig. 5).

In Fig. 7 the deduced amino-acid sequences are aligned, and various features of the sequences are highlighted. As expected, each sequence contains the conserved dehydrin motifs D(E/Q)YGNP (near the NH_2 terminus, and repeated once) and KIKEK(I/L)PG (one copy in the middle of the polypeptide, the other close to the COOH terminus). The sequences also share characteristics common to most dehydrins in being glycine-rich (23.5–28%), possessing a high content of charged amino acids, and lacking tryptophan. However, unlike most other dehydrins, they all lack a conserved stretch of serine residues. The five dehydrin sequences are closely related (86–98%) to each other in amino-acid sequence.

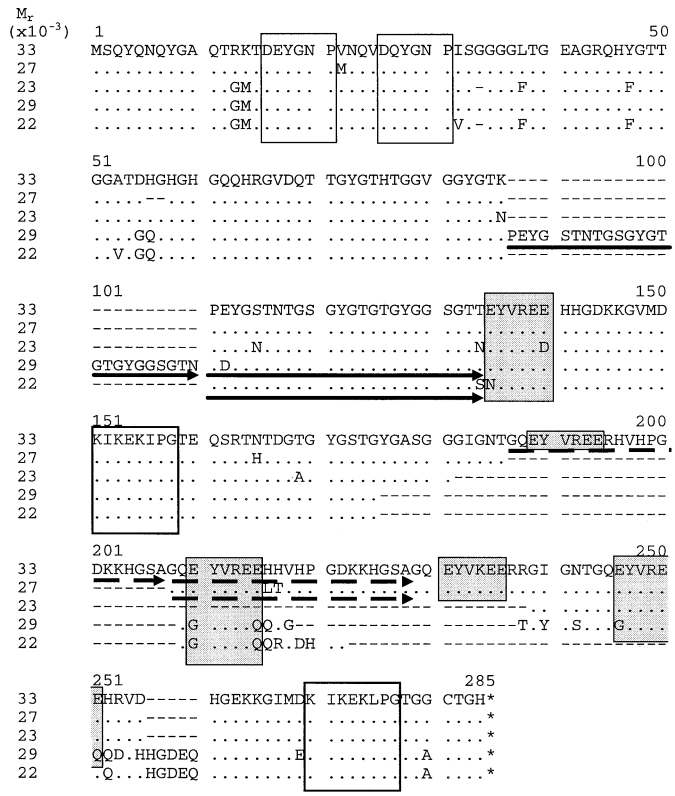

Fig. 7 Alignment of deduced amino-acid sequences from five pea dehydrin cDNA clones M_r values ($\times 10^{-3}$) of encoded polypeptides are indicated. Polypeptides and clone assignments are as follows: M_r 33000 = T123; M_r 27 000 = B210; M_r 23 000 = B12; M_r 29 000 = B213; M_r 22 000 = T41. *Dots* represent identical amino-acid residues; gaps (–) were introduced to improve sequence alignment. *Solid and dashed arrows* indicate the duplicated 24 and 21 amino-acid segments (motif A and motif B respectively). *Boxed segments* indicate conserved dehydrin motifs (DE/QYGNP, KIKEKI/LPG) and the *shaded regions* indicate the conserved core (EYVR/KEE) of motif B

Genetic variation in dehydrin polypeptides

The data in Fig. 7 allow us to compare allelic differences between dehydrins, and infer mechanisms by which differences may have arisen. For allelic variants at *Dhn2* (M_r 27 000 and 33 000), we observe that variation in copy number of a 21 amino-acid segment (motif B) was largely responsible for the difference in length between the polypeptides. In addition, there was variation in the copy number of a repeated $(HG)_n$ motif (position 56) that accounted for a further difference of two amino-acid residues, There were also four amino-acid substitutions between the polypeptide sequences. Similarly, at the *Dhn3* locus the M_r-22 000 and -29 000 allelic variants differed principally in possessing

(respectively) one or two copies of a 24 amino-acid sequence (motif A). In addition, there was length variation at three other sites that accounted for a further net difference in length of ten amino-acid residues, as well as numerous amino-acid substitutions. In summary, the major source of length heterogeneity between allelic variants at both *Dhn2* and *Dhn3* was the difference in copy number of large-sequence motifs (A and B, see Fig. 7), but there were also smaller insertion/deletion events and numerous amino-acid substitutions.

The allelic variation observed in lines 086 and MM is probably a consequence of deliberate outcrossing during breeding, and the residual heterogeneity that persists despite breeder selection and inbreeding. In parallel studies in barley, allelic variation at dehydrin loci was also observed (Lång et al. 1998). In this case the origin of the variation is less certain, and the differences between the two alleles are due to a small (6 bp) insertion/deletion, and six single-base substitutions.

The data in Fig. 7 also allow us to compare dehydrins encoded by three different loci in line 086. To summarise the differences it is convenient to consider DHN2 and DHN3 polypeptides each in relation to DHN1. An additional copy of motif A largely accounts for the difference in length between DHN3 and DHN1. Length differences between DHN2 and DHN1 polypeptides resulted largely from the presence or absence (respectively) of motif B. Thus, differences in copy number of these two motifs explains much of the size variation observed between dehydrin allelic variants (see paragraph above) as well as between the products of different dehydrin loci in line 086. A core portion of motif B, EYV(K/R)EE, also occurred at other positions in the dehydrin molecule, and variation in its copy number, as well as short insertion/deletions of neighbouring sequences, were important contributors to length variation between the different dehydrins (Table 1). This core sequence is unique to pea dehydrins, but in barley similar variation in the copy number of a semi-conserved motif with a core sequence YGQ was responsible for variation in the length of dehydrins representing different loci (Close et al. 1989). Parallel studies in barley have also revealed the existence of allelic variation within an inbreeding line (Lång et al. 1998).

Acknowledgements E.G. was the recipient of a Postdoctoral Fellowship from the Deutsche Forschungsgemeinschaft. M.R. was the recipient of a Yasuko Hiraoka Myer Research Fellowship. We thank Hart Schroeder for his generous provision of wild *Pisum* seeds, and Tony Pryor and Tony Brown for helpful discussions. The skilled technical assistance of Judy Radik is gratefully acknowledged. E.G. and M.R. contributed equally to this study. The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the following accession numbers: X63062 for clone pPsB210, X63063 for clone pPsB213, U91970 for clone T41, and U91969 for clone T123.

References

Bray EA (1993) Molecular responses to water deficit. Plant Physiol 103:1035–1040

Brown AHD (1980) Genetic basis of alcohol dehydrogenase polymorphism in *Hordeum spontaneum*. J Hered 70:127–128

Chandler PM, Robertson M (1994) Gene expression regulated by abscisic acid and its relation to stress tolerance. Annu Rev Plant Physiol Plant Mol Biol 45:113–141

Close TJ, Kort AA, Chandler PM (1989) A cDNA-based comparison of dehydration-induced polypeptides (dehydrins) in barley and corn. Plant Mol Biol 13:95–108

Devereux J, Haeberzi P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Dure L III, Crouch M, Harada J, Ho T-HD, Mundy J, Quatrano R, Thomas T, Sung ZR (1989) Common amino acid sequence domains among the LEA polypeptides of higher plants. Plant Mol Biol 12:475–486

Lång V, Robertson M, Chandler PM (1998) Allelic variation in the dehydrin gene family of 'Himalaya' barley (*Hordeum vulgare* L.). Theoret Appl Genet

Robertson M, Chandler PM (1992) Pea dehydrins: identification, characterization and expression. Plant Mol Biol 19:1031–1044

Sambrook L, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbour Laboratory, Cold Spring Harbour, New York

Schroeder HE (1982) Quantitative studies on the cotyledonary proteins in the genus *Pisum*. J Sci Food Agric 33:623–633

Skriver K, Mundy J (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell 2:503–512

Xu D, Duan X, Wang B, Hong B, Ho TD, Wu R (1996) Expression of a late embryogenesis abundant protein gene, pHVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol 110:249–257