Plant Molecular Biology 11: 147–160 (1988)

Molecular analysis of the alcohol dehydrogenase gene family of barley

Martin Trick,[1]* Elizabeth S. Dennis,[2] Kenneth J. R. Edwards[1] and William J. Peacock[2]
[1]*Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK;* [2]*CSIRO Division of Plant Industry, P.O. Box 1600, Canberra, ACT 2601, Australia;* **Present address: AFRC Institute of Plant Science Research, Cambridge Laboratory, Maris Lane, Trumpington, Cambridge CB2 2LQ UK*

Received 1 March 1988; accepted 19 April 1988

Key words: alcohol dehydrogenase, barley, DNA sequences, gene duplication

Abstract

One partial and two complete genomic clones of the three loci specifying alcohol dehydrogenase (ADH) in barley were isolated by screening libraries with a maize *Adh1* cDNA probe. Each gene is characterised by an intron arrangement similar to that of both maize *Adh1* and *Adh2*, although two genes show an exon fusion. A comparison with the maize coding sequences unambiguously assorts the barley loci into an *Adh1*-like gene and two *Adh2*-like genes, indicating that an ancient gene duplication underlies the widespread occurrence of two *Adh* loci in higher plants. In the barley lineage there has been a further duplication-transposition of a progenitor "*Adh2*" locus to give rise to the extant three-gene system, with gene copies of different ancestry being closely linked. An *Adh1* null-allele, *Adh1-M9*, has been cloned; the available sequence includes an intron with a missing acceptor splice signal. Two independent clones of one of the barley *Adh2*-like genes have an 18 bp in-frame deletion towards the 3′ end of the coding sequence. The barley *Adh2*-like genes are extensively diverged in their 5′ sequences apart from a conserved 15 bp motif in the mRNA leader region and sequences at the start of transcription. A sequence related to the hexanucleotide core of a regulatory element found in maize *Adh1* and in other anaerobically induced plant genes is present in the 5′ region of barley *Adh2*.

Introduction

Alcohol dehydrogenases (EC 1.1.1.1; ADH) are NAD^+-dependent oxidoreductases which catalyse the oxidation of a range of alcohols. It is generally assumed that ADH activity *in vivo* regenerates NAD^+ by the reduction of acetaldehyde to ethanol, facilitating glycolysis under anaerobic stress. Alcohol dehydrogenase (*Adh*) genes are among the best characterised loci in higher plants; both genetically and at the molecular level, the best known system being that of maize *Adh1* [19]. In many species a broadly similar pattern of expression may be recognised in which several ADH isozymes are inducible in tissues (particularly root systems) by some degree of anoxia, whereas usually only one isozyme is expressed "constitutively" in mature, dry seed and in pollen.

The genetics of the ADH isozymes have been thoroughly elucidated in maize [21, 46] where ADH is encoded by two loci; *Adh1* is expressed in dry seed and both *Adh1* and *Adh2* are expressed under anaerobic conditions. The active ADH enzyme is dimeric; the isozymes being formed by homodimers and intergenic heterodimers [18, 21]. In maize, the two *Adh* loci are unlinked, with *Adh1* on chromosome 1 [47] and *Adh2* on chromosome 4 [14].

The maize pattern, in which one gene is very active

in seed (and also in pollen) while both genes can be anaerobically induced, fits a number of other plant species, including Pearl Millet, Sunflower, *Eucalyptus obliqua, Pinus ponderosa* and *Bromus mollis* [for references see 24]. *Lupinus angustifolius* also has two *Adh* genes but both are expressed in germinating seeds while only one is active in pollen and in flooded roots [37]. Initial studies on cultivated barley [27, 28] and *Hordeum spontaneum* [7] suggested that the system in maize also applied to these closely related and interfertile species. More recently it was demonstrated that a third *Adh* locus exists, *Adh3*, which is not linked to *Adh1* and *Adh2* [29], these genes being closely linked on chromosome 4 [32]. *Adh3* is expressed only under extreme anoxic conditions such as an atmosphere of pure nitrogen or prolonged root flooding [26, 29]. Significant polymorphism at the isozyme level has been characterised at the *Adh3* locus [25].

Great progress has been made in the molecular genetics of the *Adh* system in maize with the isolation of both cDNA and genomic clones of each locus [10, 11, 23]. Sequence analysis has revealed that the two genes have an identical arrangement of nine introns, although these differ in sequence and length. The coding sequences are divergent by some 18% at the nucleotide level. It is probable, therefore, that the *Adh1* and *Adh2* loci are the products of an ancestral gene duplication. Analysis of 5′ regions [11] revealed short conserved sequences adjacent to, and upstream of, the TATA box as candidates for transcriptional controls [51] and a conserved motif in the 5′ untranslated regions of the mRNAs which might be associated with the known property of selective translation of *Adh*-specific messages under anaerobiosis [17, 42, 43]. Recently, transient expression experiments on the maize *Adh1* promoter [33, 51] have identified sequences necessary for anaerobic expression. The core of this sequence seems to be conserved in *Adh2* and also in other anaerobically induced genes such as maize aldolase [12]. If *Adh* is to be useful as a selectable marker for plant gene transformation systems [20, 39], it is important to gauge the species specificity of regulatory sequences.

We decided to clone the barley *Adh* genes to directly address the interesting question of gene homology between the barley and maize systems and also to look for conservation of such regulatory sequences. We report the molecular cloning of an allele of each of the barley *Adh* genes, present the nucleotide sequences and comment on their evolutionary relationships. A sequence closely related to the anaerobic "core" has been identified in the one barley gene for which extensive flanking sequence is available.

Materials and methods

Plant material and DNA extraction

Plant lines included *Hordeum vulgare* cv. Proctor and the ADH1 null mutants *Adh1-M9* and *Adh1-M140*, which were selected after sodium azide mutagenesis in a Proctor background [28]. *Hordeum spontaneum* genotypes *Adh1-F* and *Adh1-M* and *Zea mays* genotype *Adh1-S* were used in protein dissociation-reassociation experiments according to procedures previously described [28]. For DNA, shoot tissue from 4-day-old seedlings germinated on agar was ground to a fine powder in liquid nitrogen. 5 ml of 0.1 M NaCl, 50 mM EDTA pH 8.5 was added per gram of tissue and the slurry made 2% sarcosyl, 1 M $NaClO_4$ and 0.1% diethyl pyrocarbonate. An equal volume of phenol/chloroform/isoamyl alcohol (24:24:1) was added and the homogenate stirred at 4 °C for 1 h. The aqueous phase was recovered by centrifugation, re-extracted and ethanol-precipitated. The precipitate was pelleted and resuspended in a small volume of autodigested protease solution (1 mg/ml in 50 mM Tris, 10 mM EDTA pH 8.0) and incubated for 1 h at 37 °C. DNA was then banded on a CsCl density gradient by centrifugation at 45000 rpm for 18 h, butanol-extracted, ethanol-precipitated and finally resuspended in TE buffer.

Southern transfer and hybridisation

10 μg of barley DNA was digested with five-fold excess of restriction enzyme, electrophoresed on 0.8% agarose, and blotted onto nitrocellulose [49] with a

10-minute acid wash [50]. Probe DNA was nick-translated [41] to a specific activity of 5×10^8 cpm μg^{-1}. Filters were hybridised with 2×10^7 cpm, at a probe concentration of 10 ng ml^{-1}, in 5× SSC, 5× Denhardt's solution, 0.5% SDS at 68 °C overnight and were washed in 1× SSC at the same temperature. The probe used in genomic blots was pZML793, a full-length cDNA clone of maize *Adh1* [10].

Genomic cloning

Genomic DNAs from various barley genotypes were digested to completion with *Eco* RI or *Bam* HI and ligated to either *Eco* RI-digested λgt7 arms [9] or *Bam* HI-digested λEMBL4 arms [22]. Ligated DNAs were packaged *in vitro* and plated without amplification on K802 host cells. Library screening was carried out by plaque hybridisation [3] using gel-purified pZML793 insert as a probe. Positive phage were isolated by several rounds of plaque purification and DNAs prepared by standard techniques [36]. Barley DNA inserts were recloned into pUC8 or pUC9 plasmid vectors.

RNA analysis

RNA samples were run on formaldehyde denaturing gels and transferred to nitrocellulose. Northern hybridisations were performed according to standard techniques [36]. S_1 mapping of transcription starts was carried out by hybridisation of single-stranded M13 probes complementary to the 5′ ends of barley *Adh* messages essentially according to [10]. To determine the start of transcription for the *Adh2* mRNA, an appropriate probe was annealed to 5 μg of poly(A)$^+$ RNA extracted from root tissue of anaerobically treated *Adh1-M9* mutant seedlings.

Sequencing

Nucleotide sequences were determined using the dideoxy chain termination method [45]. Bal31 deletion fragments, or restriction fragments made blunt with Klenow, were subcloned into *Sma* I-digested, phosphatased M13mp8 or mp9 vectors. Sequences were determined several times independently from each strand. Specific oligonucleotide primers were synthesised to assist in certain regions. Sequence analyses were performed on the University of Cambridge IBM 3081 computer.

Results

Homology between maize and barley ADH1 polypeptides

Functional homology at the protein level was demonstrated between the maize and barley AHD1 enzymes by *in vitro* reassociation of dissociated protein extracts from mature seeds. Figure 1 shows an ADH activity gel of such an experiment. The ADH1 homodimers of maize and of two electrophoretically marked genotypes of barley migrate to different positions (Fig. 1a, b, e). In each of the two interspecific reassociation tracks (Fig. 1c, f) there appears a novel band of ADH activity with a mobility intermediate to those of the respective homodimers, indicating that the ADH1 polypeptides of maize and barley retain sufficient homology to allow the formation of enzymatically active heterodimers.

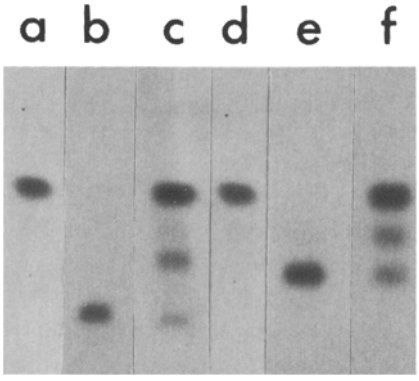


Fig. 1. Starch gel zymogram of *in vitro* barley/maize ADH1 heterodimerisations. Dissociated/reassociated protein extracts from dry seed tissue: (a) maize *Adh1-S*; (b) barley *Adh1-F*; (c) maize *Adh1-S* + barley *Adh1-F*; (d) maize *Adh1-S*; (e) barley *Adh1-M*; (f) maize *Adh1-S* + barley *Adh1-M*. Anodal migration is towards the bottom of the figure.

Detection of Adh *sequences in barley genomic DNAs*

Barley genomic DNAs were digested with *Eco* RI, blotted and probed with nick-translated pZML793, a full-length maize *Adh1* cDNA clone [10]. Filters were washed at a reduced stringency (1× SSC at 68 °C). Three discrete hybridising fragments of 11 kb, 6.6 kb and 3.5 kb were readily detected (Fig. 2); on some blots, an additional faint band of less than 2 kb could be resolved. This sequence complexity indicated the presence of a multigene family in barley, consistent with the gene copy number of three inferred from isozyme studies [29]. Similar patterns of hybridisation were obtained from Proctor, *Adh1-M9* and *Adh1-M140* genotypes, both with *Eco* RI (Fig. 2) and with several other restriction enzymes (data not shown), indicating that the ADH1-null phenotype of these mutants [28] is not accompanied by gross rearrangements at any of the gene loci.

Barley Adh *gene isolation*

Genomic libraries were constructed from DNA prepared from both Proctor and *Adh1-M9* genotypes and screened by plaque hybridisation using a gel-purified pZML793 insert probe. Two clones were isolated from an *Adh1-M9 Eco* RI limit digest library made in the vector λgt7. These recombinants, λM9 15.1 and λM9 1.2, contained 6.6 kb and 11.2 kb inserts respectively, corresponding to the sizes of the two largest *Eco* RI fragments detected by blotting. Subsequently, a *Bam* HI library was constructed in the vector λEMBL4 from Proctor DNA and two more positive phage, λPr 3.1 and λPr 7.1, were recovered with inserts of 10.5 kb and 5 kb respectively. Each barley genomic segment was recloned in a plasmid vector for further analysis.

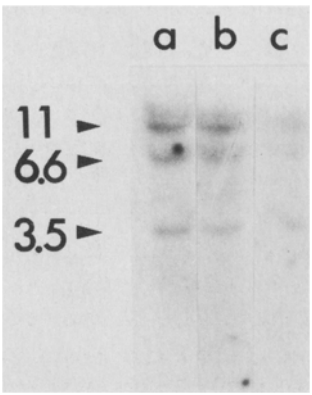


Fig. 2. Genomic organisation of *Adh* sequences in barley. DNAs were digested with *Eco* RI, blotted and probed with a maize *Adh1* cDNA, pZML793. Genotypes: (a) Proctor; (b) *Adh1-M9*; (c) *Adh1-M140*. Approximate sizes of hybridising fragments are given in kb.

Homology relationships between these four clones was investigated by blotting experiments, performed at high stringency (data not shown). Under these conditions the pM9 15.1 probe did not cross-hybridise with the other three clones, while pM9 1.2 and pPr 3.1 behaved identically, cross-hybridising strongly with each other and weakly with pPr 7.1. The probe derived from pPr 7.1 gave the reciprocal effect.

Additionally, each probe was used on Northern blots of poly(A)$^+$ RNA extracted from mature, dry seeds of both Proctor and *Adh1-M9* genotypes (data not shown). Only the pM9 15.1 probe was found to hybridise, producing a $\simeq$1.5 kb band only in Proctor mRNA. Since dry seed expresses the ADH1 isozyme almost exclusively [28], this suggested that the pM9 15.1 insert contained barley *Adh1* sequences and, therefore, the remaining genomic clones originated from other *Adh*-specific loci. Furthermore, the absence of a signal in *Adh1-M9* mRNA indicated that this mutant phenotype arises from a dysfunction in the transcription or processing of *Adh1* message and not a structural lesion in the gene coding sequence.

Gene sequences

Blotting experiments on the four genomic clones using pZML793 and derived 5′ and 3′ probes established the location and transcriptional orientation of the *Adh* genes in each DNA segment (summarised in Fig. 3). Each clone was characterised by a unique restriction map and yet each contained both 5′ and 3′ sequences. Complete nucleotide sequences of the coding regions of pM9 1.2 and pPr 7.1 were determined. A total of approximately 1 kb of sequence, distributed throughout the coding region, was deter-

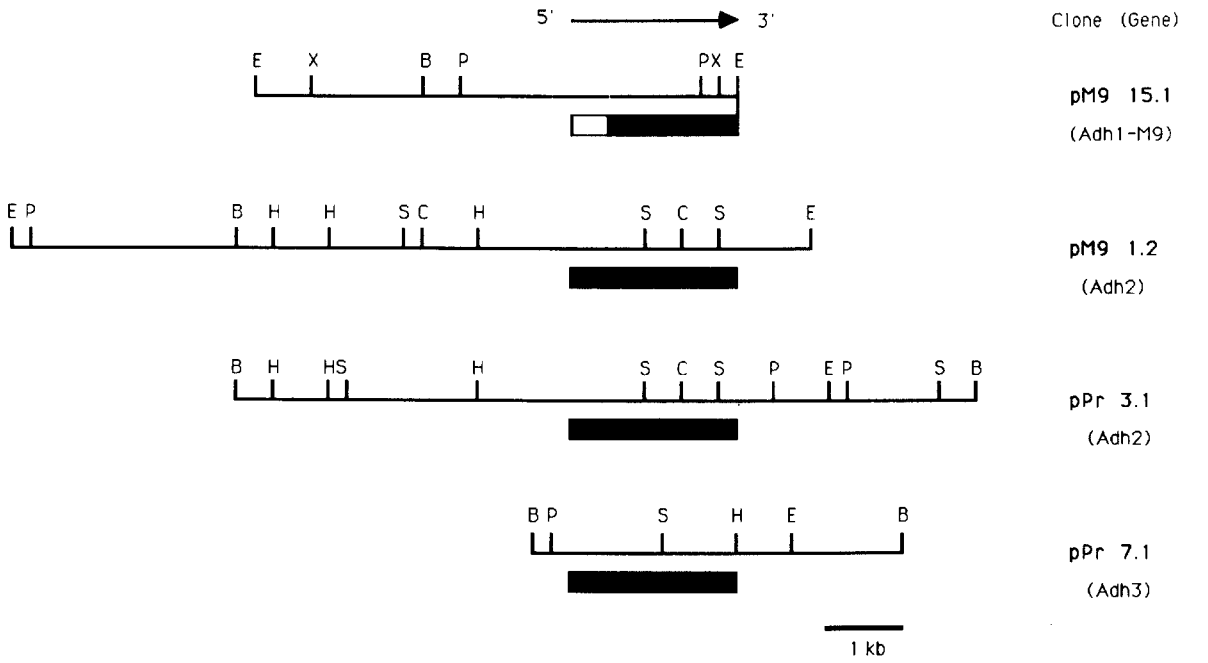


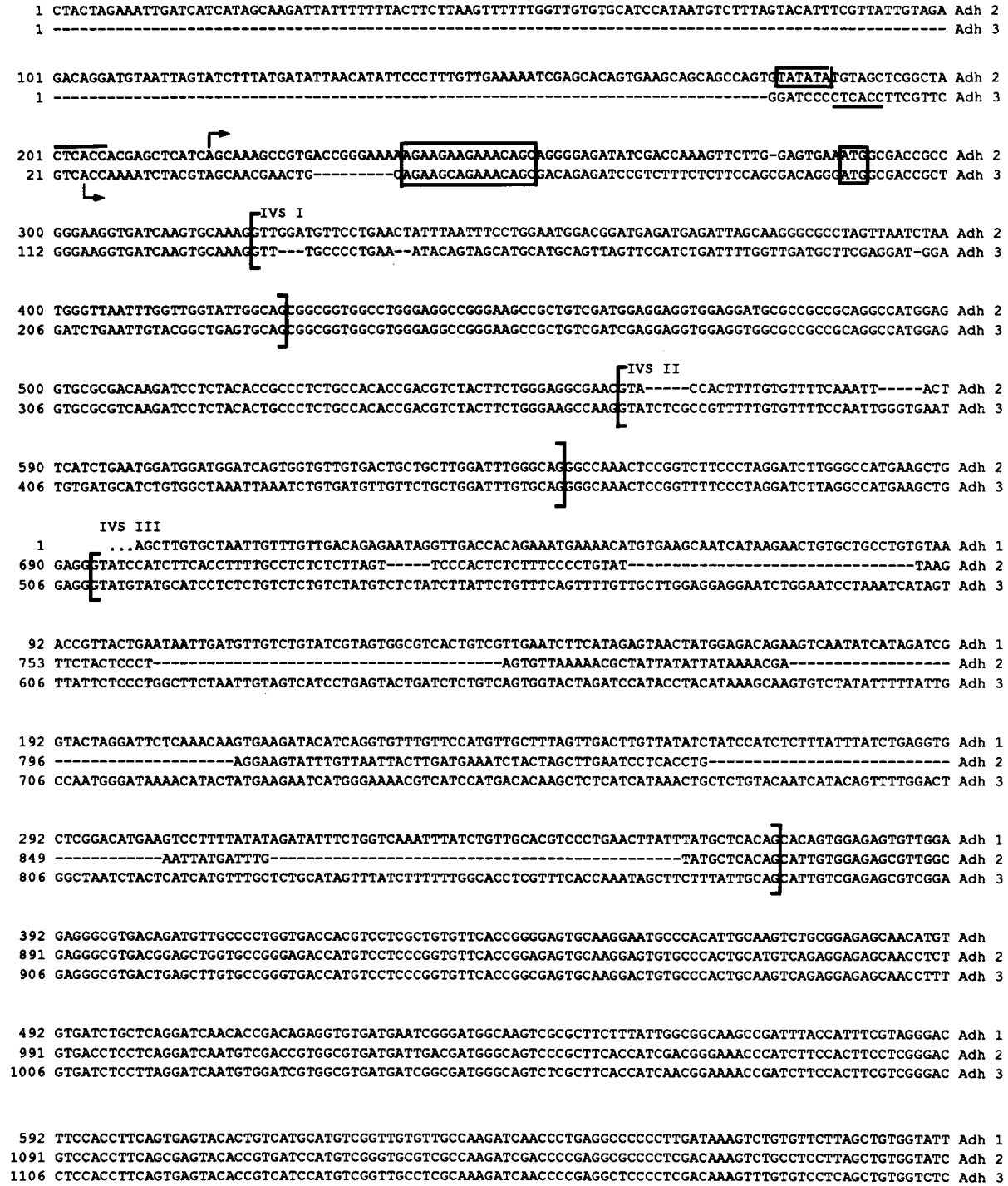
Fig. 3. Restriction maps of barley *Adh* genomic clones. Inferred gene identities are given with the clone designations. Pr and M9 refer to the Proctor and *Adh1-M9* genotypes from which clones were isolated. The locations of the gene coding regions are displayed as solid bars along with the direction of transcription. The *Adh1-M9* gene is truncated in pM9 15.1 by an *Eco* RI cloning site, the open part of the bar indicates the 5′ region of the gene not sequenced. Restriction sites are as follows; B, *Bam* HI; C, *Cla* I; E, *Eco* RI; H, *Hind* III; P, *Pst* I; S, *Sal* I; X, *Xba* I.

mined from pPr 3.1. This proved to be identical to the pM9 1.2 sequence. We concluded that these two clones carried copies of the same gene (see Discussion). In addition, a partial sequence from pM9 15.1 was obtained comprising 1.5 kb extending from within an internal intervening sequence (IVS III) and truncated by the distal *Eco* RI site of the genomic clone (Fig. 3). The sequences (Fig. 4) show extensive homology with the maize *Adh* genes.

The gene coding sequences are interrupted by AT-rich introns at precisely the same locations as in the maize genes except that IVS IX has been eliminated from both the genes represented in pM9 1.2 and pPr 7.1. Each intron, with one exception (see below), has the GT and AG 5′ and 3′ splicing signals characteristic of eukaryotic genes [6]. These intron/exon junction sequences may be expanded into the consensus sequences (A/G)AGGTATG and T(G/T)CAGG(T/G) which are essentially the same as the corresponding ones in the maize genes and similar to the canonical animal sequences.

The sequence for IVS V of the pM9 15.1 gene lacks a 3′ acceptor splice site (Fig. 4). No alternative splice sites in the appropriate reading frame are available in the adjacent downstream sequences. Although a direct comparison with the progenitor allele has not been possible, inspection of the sequence suggests that a small deletion may have removed the splice site; this is the only irregularity found in the available sequence. A splicing mutation could be consistent with the RNA phenotype and this, combined with the gene's homology to maize *Adh1* (see below), strengthens the proposal that pM9 15.1 contains the *Adh1-M9* allele.

One of the other two *Adh* genes cloned shows a major mutation in coding sequence in that the 5′ end of exon 9 in pM9 1.2 is characterised by an 18 bp deletion which conserves the reading frame (Fig. 4). This feature is also observed in the identical copy of this gene independently cloned in pPr 3.1 and so this feature is assumed not to be a cloning or sequencing artefact.


IVS I
IVS II
IVS III

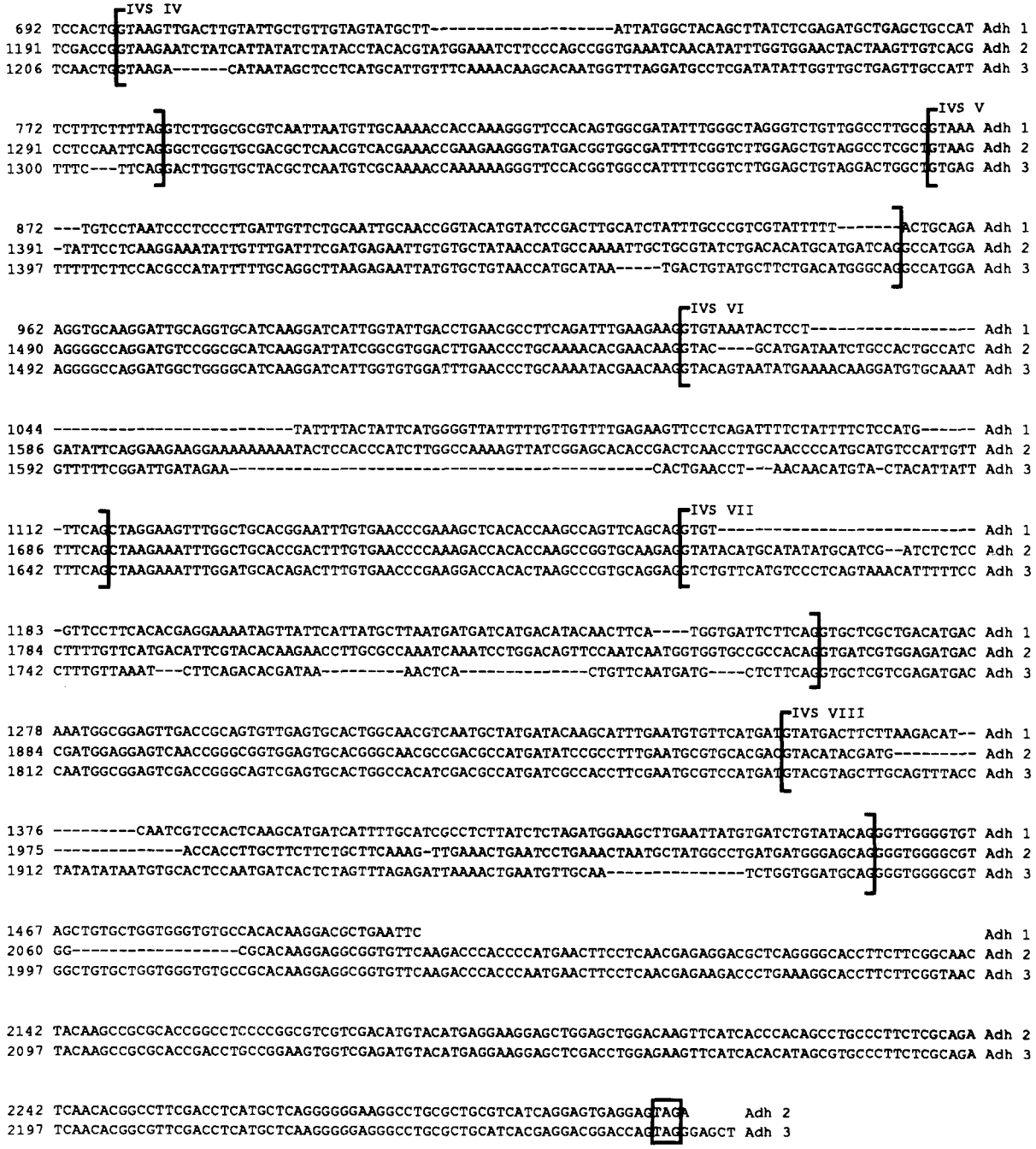


Fig. 4. Nucleotide sequences of the barley *Adh* clones. The partial sequence determined from pM9 15.1 (*Adh1-M9*) is displayed above the aligned sequences of the complete coding regions of pM9 1.2 (*Adh2*) and pPr 7.1 (*Adh3*). Positions of intervening sequences are indicated. Features such as the presumptive TATA box and translational initiation and termination codons are boxed, as is the 5′ leader homology. Sites of transcription initiation, either mapped or presumptive, are indicated by arrows. The 6 bp conserved motif 5′ to these sites is underlined.

The coding sequences were extracted and compared with each other and with the two maize *Adh* genes. Divergences at nucleotide and amino acid levels were calculated; the partial sequence from *Adh1-M9* being included in this analysis. These data are summarised in Table 1. The results of the cross-hybridisation experiments described above are borne out by the barley sequence homologies, with pM9 1.2 and pPr 7.1 being about half as divergent as either is to pM9 15.1. The comparisons with the maize genes show that the latter clone is the most homologous to maize *Adh1*; less than a third of the 15.8% nucleotide divergence recorded results in amino acid substitutions. The genes cloned in pM9 1.2 and pPr 7.1 are more homologous to maize *Adh2*, pPr 7.1 being the most conserved with exactly a third of the divergence being expressed but with pM9 1.2 displaying considerably more amino acid substitution.

The lengths of equivalent introns are generally different between barley and maize and their sequences totally diverged but, amongst the barley genes, small regions of homology can occasionally be detected between equivalent intervening sequences. At the intron-exon junctions, short motifs, including the splicing signals, sometimes show homologies which run counter to the gene relationships already described. For example, 11 nucleotides at the 3′ end of IVS III are perfectly matched between the *Adh1-M9* allele and the pM9 1.2 sequence (Fig. 4) genes which, on the basis of coding sequence homology, have been long diverged. Presumably, these residual homologies reflect fluctuations in the rates of divergence of neutral DNA sequences.

5′ sequences

From the 5′ regions of the two complete genes available, 187 bp of sequence upstream of the initiating ATG has been determined from pM9 1.2, but the equivalent region of the other gene is truncated by the *Bam* HI cloning site only 99 bp 5′ to the start of translation.

S_1 mapping experiments were performed to define the start of transcription of the pM9 1.2-specific message (data not shown). An appropriate 5′ probe was annealed to RNA extracted from moderately induced root tissue of the *Adh1-M9* mutant which expresses only *Adh2*-specific message. A DNA fragment extending 70 nt 5′ from the translational start was protected, giving the probable transcription initiation sequence as TCATCAGCAA. A presumptive TATA box sequence (GTATATATG) is located 31 bp upstream of this transcription start (Fig. 4), conforming to the plant consensus in both sequence

Table 1. Table of % sequence divergence in all pairwise comparisons of barley and maize *Adh* genes. The partial sequence of pM9 15.1 is used for *Adh1*. The 18 bp deletions in pM9 1.2 (*Adh2*) and pPr 7.1 (*Adh3*) are treated as single mutations. The divergences given first are those at the nucleotide level, those in brackets are at the amino acid level.

	Adh1	*Adh2*	*Adh3*	maize *Adh1*	maize *Adh2*
Adh1	0	27.7 (22.4)	22.3 (19.8)	15.8 (8.6)	21.8 (17.7)
Adh2		0	11.0 (10.0)	20.9 (18.7)	13.9 (14.5)
Adh3			0	18.7 (14.5)	13.9 (10.3)
maize *Adh1*				0	18.0 (12.6)
maize *Adh2*					0

and position [38]. No CAAT or AGGA sequences [38] were found 5′ to the TATA box, in common with the absence of such features in the maize *Adh2* gene [11]. Although an S_1 experiment was not performed for the pPr 7.1 gene, an inspection of the available 5′ sequence reveals a likely transcription start; the homologous sequence, TCGTCACCAA, 75 bp upstream of the ATG. A 6 bp motif, CTCACC, is conserved in both genes at a similar position, 10–11 bp 5′ to each presumptive start site.

In the 5′ non-translated regions of the two mRNAs, a 15 bp sequence is conserved in all but one nucleotide, 33–34 bp upstream from the ATG. The maize *Adh1* and *Adh2* mRNAs share a completely different homology at a similar position [11]; although the barley pM9 1.2 leader region has a truncated version of this latter sequence just 2 bp 5′ to the ATG.

Gene identifications

Evidence has been given which suggests that the pM9 15.1 clone is the barley *Adh1-M9* allele. The genetical identities of the similar pM9 1.2. and pPr 7.1 genes are more difficult to establish directly. However, an *Adh* gene has been cloned from wheat which is extremely homologous to the pPr 7.1 sequence. (Mitchell *et al.*, in preparation). A probe derived from the 3′ untranslated region of the wheat clone hybridises to a single restriction fragment in barley total genomic DNA but fails to hybridise to any barley-specific band in wheat/barley addition lines in which the barley chromosomes are added one at a time to the wheat complement. Wheat addition lines containing each of the barley chromosomes were tested except for chromosome 5, which is not available. Since *Adh1* and *Adh2* are known to be on chromosome 4, this negative result indicates that pPr 7.1 is *Adh3* and suggests, furthermore, that this locus resides on chromosome 5. The pM9 1.2 clone would therefore be *Adh2*.

Predicted protein sequences

ADH protein sequences were deduced from the open reading frames identified in the pM9 1.2 and pPr 7.1 genes. pPr 7.1 would encode a polypeptide of 379 amino acids (identical to the length of the maize ADH1 and ADH2 proteins) with a calculated molecular weight of 41 012. The polypeptide encoded by pM9 1.2, shorter by 6 residues, would have a molecular weight of 40512. A comparison of the predicted protein sequences is shown in Fig. 5. The pM9 1.2 polypeptides has four negative charges more than that specified by pPr 7.1, conferring on it a more rapid anodal electrophoretic mobility which would also be consistent with the pM9 1.2 genomic clone being the *Adh2* locus [29].

Discussion

Gene structures and evolution

Alleles of each of the three loci specifying ADH in barley have been isolated by molecular cloning and analysed. On the criteria of DNA and amino acid homology with the maize *Adh* genes, two of the genes (*Adh2* and *Adh3*) have been designated as maize *Adh2*-like and the third, from both sequence comparisons and its developmentally regulated expression, has been identified as an *Adh1* mutant allele. The complete gene structures available for the *Adh2* and *Adh3* genes and the truncated structure for the *Adh1-M9* allele are either identical to or, with intron loss, derivative of that characterising both maize *Adh1* and *Adh2* [10, 11]. An *Adh1*-like gene cloned from pea has a similar structure [35] and the *Arabidopsis Adh* gene displays only intron loss with respect to the maize pattern [8]. This conservation of intron position amongst examples of both monocot and dicot *Adh* genes argues strongly for there having been a single *Adh* gene lineage in higher plants and for *Adh* systems having evolved by repeated gene duplication. The widespread occurrence of at least two *Adh* loci in all monocots and most dicots examined [24] has been interpreted as the result of the initial gene duplication predating this major separation, although the intermediate status of the single *Arabidopsis Adh* gene [8] may challenge this assumption.

What is clear is that within the Gramineae a fur-

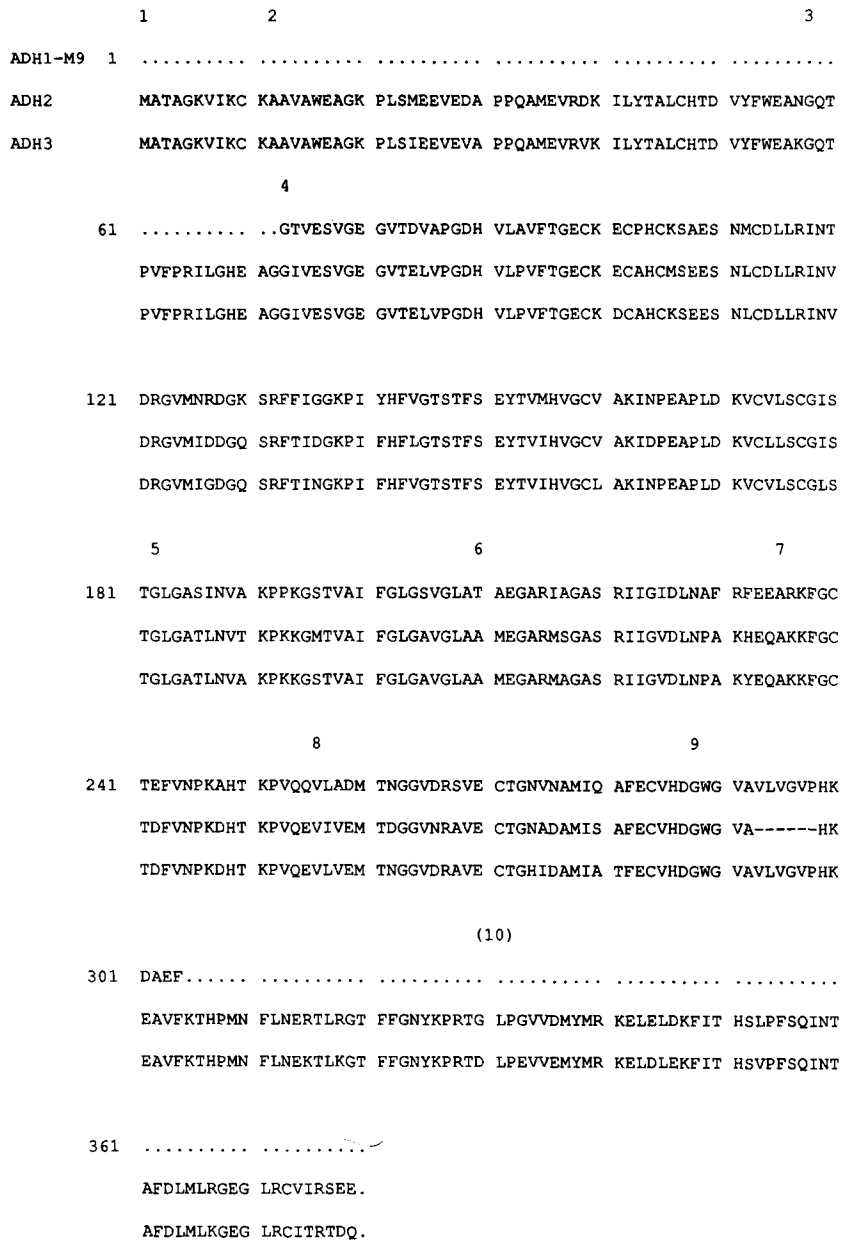


Fig. 5. Predicted amino acid sequences encoded by barley *Adh* genes. Exons are numbered above their locations in the polypeptide sequence.

ther, much more recent, gene duplication has occurred. Despite the lack of data for the *Adh1* locus, the exon fusion shared by the *Adh2* and *Adh3* is likely to be diagnostic. This exon fusion and the acquisition of certain characteristic peptide motifs not found in the maize sequences (for instance, ELV; residues 84–86 in Fig. 5) clearly predated the gene duplication in the barley lineage. The existence of a highly homologous *Adh* gene with an identical exon fusion in hexaploid wheat (Mitchell *et al.*, in preparation) also places the intron loss before in divergence of barley and at least one of the diploid progenitors of wheat.

It is interesting that these inferred phyletic affini-

ties between the barley genes appear not to be reflected in their linkage relationships, thus *Adh1* and *Adh2* are closely linked on chromosome 4 [32] but, from the sequence homologies amongst the three genes cloned, they appear not to be related by the recent gene duplication. Instead, *Adh2* and the unlinked *Adh3* locus are clearly the descendants of a progenitor *Adh2*-like gene with chromosomal transposition of one of the gene copies having occurred either concomitantly with or subsequent to the duplication event. There is considerable heterogeneity in linkage relationships between *Adh* loci amongst the Gramineae. The maize *Adh1* and *Adh2* genes are unlinked, being located on chromosomes 1 and 4 respectively [20]. In wheat, isozyme studies have demonstrated the existence of two linked *Adh* loci, resembling *Adh1* and *Adh2* in expression, on homoeologous group 4 [30, 31] although the *Adh* gene cloned was previously unknown, mapping to group 1 (L. Mitchell *et al.*, in preparation). In Pearl Millet, two extremely tightly linked *Adh* loci occur [1, 2]. On balance, it is probable that a system of two unlinked *Adh* loci existed in the grasses prior to the divergence of the Maydeae and Hordeae. Pearl Millet might then represent a degenerate condition of the three gene system, as suggested by Hanson & Brown [25].

Protein structures

Like the maize enzymes [10, 11], the predicted polypeptides encoded by the barley loci generally show conservation of those residues implicated, by analogy with the horse liver ADH three-dimensional structure [5], in structural or catalytic roles. For instance, the cysteine and histidine amino acids important in binding both zinc ligands are conserved; these particular residues are invariant amongst all plant ADH sequences reported. Likewise, those residues identified for coenzyme binding, catalytic function and subunit interaction are also conserved. Although the ADH1 data are incomplete, the ready heterodimerisation of the maize and barley ADH1 polypeptides to give an active enzyme argues for the conservation of those functional residues not determined in the present work. A similar active association of heterologous ADH1 subunits of *Arabidopsis* and maize has been reported [15] and antibodies raised against maize ADH cross-react with a variety of higher plant ADH proteins [4].

The clear exception to this functional conservation is the deletion of residues 294–299 suffered by the putative ADH2 polypeptide encoded by the pM9 1.2 sequence; this region of the protein lies within the second mononucleotide binding domain [5] and is also implicated in subunit interaction. The sequence after this deletion is very well conserved between the ADH2 and ADH3 polypeptides. The possibility of the deletion being a cloning artefact was excluded by there being two independent isolates of the gene. Although there is no direct evidence of the pM9 1.2 gene specifying the functional ADH2 polypeptide in the plant, S_1 mapping has demonstrated its transcription under anaerobic conditions. The possibility of it representing a very recently generated pseudogene is mitigated by the *Adh* sequence complexity, as revealed in genomic blots, not readily accomodating a fourth locus.

Two copies of the *Adh2* gene were cloned on genomic fragments which show a high level of restriction site variation in flanking sequences (Fig. 3). From the limited sequencing done on the pPr 3.1 gene, it would appear that this variation is completely absent from both coding and intervening sequences within the transcription unit. The pPr 3.1 clone was isolated from a different acquisition of the Proctor variety to that which formed the progenitor of the *Adh1-M9* mutagenised genotype [28] and from which each of the M9 clones were isolated. Thus, this flanking sequence divergence probably reflects residual genetic polymorphism in seed stocks. A comparable amount of divergence in the sequences flanking the structural genes of various maize *Adh1* alleles has been observed [34, 44].

5′ sequences

Despite their probably relatively recent derivation from a common ancestor, the barley *Adh2* and *Adh3* genes show considerable divergence in sequence 5′ to the ATG. There are only two obvious regions of sequence conservation between them. The first of these is located at the mapped or putative transcrip-

tion start sites. The consensus sequence here, TC(A/G)TCA(G/C)CAA, shows homology to that derived from the maize genes, TCCT(-/G)AC-(AGG/CAA) [10, 16]. Located 11 or 10 nt upstream of the transcription start in each barley gene is the motif CTCACC which resembles the core of the sequence; no equivalent is found in the maize genes.

The second region of homology lies within the non-translated leader sequences. The maize *Adh* genes have a conserved sequence in their leaders G(G/T)TCT(C/T)GGAGTGG..ATCGA, approximately 45 nt upstream of the translation start [11], which has been suggested as a possible regulatory element involved in the known selective translation of *Adh* messages, amongst others, during anaerobiosis [43]. Interestingly, the barley *Adh2* gene shows perfect homology to a truncated version of the maize *Adh2* sequence, GTTCTTGGAGTG, but at a different position, only 2 nt upstream of the ATG (Fig. 4). No counterpart to this sequence is found in the barley *Adh3* gene. However, both barley genes do have a highly conserved motif, AGAAG(A/C)AGA-AACAGC, 33-34 nt 5′ to the start of translation which perhaps has evolved to fulfil a similar role.

The sequence data available so far has not permitted an extensive comparison of 5′ regions between the barley and the maize genes. In the maize *Adh1* gene, an anaerobic regulatory element (ARE) has been located in the 5′ region and functionally defined by a deletion analysis of constructs expressed in a transient expression system [33, 51]. This element is composed of two subregions both of which are essential for anaerobic induction. Sequence comparisons between the maize *Adh1* ARE subregions and the upstream regions of other anaerobically regulated genes, including maize *Adh2*, *Arabidopsis* and pea *Adh* and maize aldolase [12], have revealed the apparent conservation of a hexanucleotide "core" sequence, TGGTTT, between positions -170 and -70 relative to the transcription start [13]. An examination of the corresponding region of the 5′ sequence available for the putative barley *Adh2* gene reveals a related, but not perfectly conserved, motif, TGGTTG at position -70. Functional assays would be needed to demonstrate that this sequence was at the core of an equivalent regulatory element in barley. If so, the significance of this deviation from a consensus sequence established amongst considerably more taxonomically distinct plants is unclear.

Acknowledgements

We are grateful to Nick Harberd for discussions and permission to use the heterodimerisation data and to Luke Mitchell for sharing data prior to publication. Thanks are also due to David Webb and John Spencer in Cambridge and to Janice Norman in Canberra for excellent technical assistance. M. T. acknowledges the receipt of a Wain Fund grant towards travel to Australia where part of this work was done.

References

1. Banuett-Bourillon F: Linkage of the alcohol dehydrogenase structural genes in pearl millet (*Pennisetum typhoides*). Biochem Genet 17: 537-552 (1982).
2. Banuett-Bourillon F, Hague DK: Genetic analysis of alcohol dehydrogenase isozymes in pearl millet (*Pennisetum typhoides*) Biochem Genet 17: 537-552 (1979).
3. Benton WD, Davis RW: Screening λgt recombinant clones by hybridisation to single plaques in situ. Science 196: 180-182 (1977).
4. Berendes K: PhD Thesis, Indiana University, Bloomington (1978).
5. Brändén C-I, Eklund H, Cambillau C, Pryor AJ: Correlation of exons with structural domains in alcohol dehydrogenase. EMBO J 3: 1307-1310 (1984).
6. Breathnach R, Chambon P: Organisation and expression of eucaryotic split genes coding for proteins. Ann Rev Biochem 50: 349-383 (1981).
7. Brown AHD: Genetic basis of alcohol dehydrogenase polymorphism in *Hordeum spontaneum*. J Heredity 70: 127-128 (1980).
8. Chang C, Meyerowitz EM: Molecular cloning and DNA sequence of *Arabidopsis thaliana* alcohol dehydrogenase gene. Proc Natl Acad Sci USA 83: 1408-1412 (1986).
9. Davis RW, Botstein D, Roth JR: Advanced Bacterial Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1978).
10. Dennis ES, Gerlach WL, Pryor AJ, Bennetzen JL, Inglis A, Llewellyn D, Sachs MM, Ferl RJ, Peacock WJ: Molecular analysis of the alcohol dehydrogenase (*Adh1*) gene of maize. Nucleic Acids Res 12: 3983-4000 (1984).
11. Dennis ES, Sachs MM, Gerlach WL, Finnegan EJ, Peacock WJ: Molecular analysis of the alcohol dehydrogenase 2 (*Adh2*) gene of maize. Nucleic Acids Res 13: 727-743 (1985).

12. Dennis ES, Gerlach WL, Walker JC, Lavin M, Peacock WJ: An anaerobically regulated aldolase gene of maize. Mol Biol (in press)
13. Dennis ES, Walker JC, Llewellyn DJ, Ellis JG, Singh K, Tokuhisa JG, Wolstenholme DR, Peacock WJ: The response to anaerobic stress: transcriptional regulation of genes for anaerobically induced proteins. In: Wettstein D, Chua N-H (eds.) Plant Molecular Biology 1987, pp. 407–417. Plenum Press, New York (1987).
14. Dlouhy SR: PhD Thesis, Indiana University, Bloomington (1980).
15. Dolferus R, Jacobs M: Polymorphism of alcohol dehydrogenase in *Arabidopsis thaliana* (L.) Heynh.: genetical and biochemical characterisation. Biochem Genet 22: 817–838 (1984).
16. Ellis JG, Llewellyn DJ, Dennis ES, Peacock WJ: Maize *Adh1* promoter sequences control anaerobic regulation: addition of upstream promoter elements from constitutive genes is necessary for expression in tobacco. EMBO J 6: 11–16 (1987).
17. Ferl RJ, Dlouhy SR, Schwartz D: Analysis of maize alcohol dehydrogenase by native SDS two-dimensional electrophoresis and autoradiography. Mol Gen Genet 169: 7–12 (1979).
18. Freeling M: Simultaneous induction by anaerobiosis or 2,4-D of multiple enzymes specified by two unlinked genes: differential ADH1-ADH2 expression in maize. Mol Gen Genet 127: 215–217 (1973).
19. Freeling M, Bennett DC: Maize *Adh1*. Ann Rev Genet 19: 297–323 (1985).
20. Freeling M, Birchler JA: Mutants and variants of the alcohol dehydrogenase-1 gene in maize. In: Setlow JK, Hollaender A (eds.) Genetic Engineering Principles, Methods III, pp. 223–264. Plenum Press, New York (1981).
21. Freeling M, Schwartz D: Genetic relationships between the multiple alcohol dehydrogenases of maize. Biochem Genet 8: 27–36 (1973).
22. Frischauf A-M, Lehrach H, Poustka A, Murray N: Lambda replacement vectors carrying polylinker sequences. J Mol Biol 170: 827–842 (1983).
23. Gerlach WL, Pryor AJ, Dennis ES, Ferl RJ, Sachs MM, Peacock WJ: cDNA cloning and induction of the alcohol dehydrogenase gene (*Adh1*) of maize. Proc Natl Acad Sci USA 79: 2981–2985 (1982).
24. Gottlieb LP: Conservation and duplication of isozymes in plants. Science 216: 373–380 (1982).
25. Hanson AD, Brown AHD: Three alcohol dehydrogenase genes in wild and cultivated barley: characterisation of the products of variant alleles. Biochem Genet 22: 495–515 (1984).
26. Hanson AD, Jacobsen JV, Zwar JA: Regulated expression of three alcohol dehydrogenase genes in barley aleurone layers. Plant Physiol 75: 573–581 (1984).
27. Harberd NP: A genetical investigation of the alcohol dehydrogenase system in barley. PhD Thesis, University of Cambridge (1981).
28. Harberd NP, Edwards KJR: A mutational analysis of the alcohol dehydrogenase system in barley. Heredity 48: 187–195 (1982).
29. Harberd NP, Edwards KJR: Further studies on the alcohol dehydrogenases in barley: evidence for a third alcohol dehydrogenase locus and data on the effect of an alcohol dehydrogenase-1 null mutation in homozygous and in heterozygous condition. Genet Res Camb 41: 109–116 (1983).
30. Hart GE: Evidence for triplicate genes for alcohol dehydrogenase in hexaploid wheat. Proc Natl Acad USA 66: 1136–1141 (1970).
31. Hart GE: Evidence for a second triplicate set of alcohol dehydrogenase structural genes in hexaploid wheat. Genetics 94: s41 (1980).
32. Hart GE, Islam AKMR, Shepherd KW: Use of isozymes as chromosome markers in the isolation and characterisation of wheat-barley chromosome addition lines. Genet Res, Camb 36: 311–325 (1980).
33. Howard EA, Walker JC, Dennis ES, Peacock WJ: Regulated expression of an alcohol dehydrogenase 1 chimeric gene introduced into maize protoplasts. Planta 170: 535–540 (1987).
34. Johns MA, Strommer JN, Freeling M: Exceptionally high levels of restriction site polymorphism in DNA near the maize *Adh1* gene. Genetics 105: 733–743 (1983).
35. Llewellyn DJ, Finnegan EJ, Ellis JG, Dennis ES, Peacock WJ: Structure and expression of an alcohol dehydrogenase 1 gene from *Pisum sativum*.. J Mol Biol 195: 115–123 (1987).
36. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press Laboratory, Cold Spriong Harbor, NY (1982).
37. Marshall DR, Broué P, Oram RW: Genetic control of the alcohol dehydrogenase isozymes in narrow-leafed lupins. J Hered 65: 198–203 (1974).
38. Messing J, Geraghty D, Heidecker G, Hu N-T, Kridl J, Rubenstein J: Plant gene structure. In: Kosuge T, Meredith CP, Hollaender A (eds.) Genetic Engineering of Plants, pp. 211–227. Plenum Press, New York (1983).
39. Peacock WJ, Dennis ES, Gerlach WL, Llewellyn D, Lörz H, Pryor AJ, Sachs MM, Schwartz D, Sutton WD: Gene transfer in maize: controlling elements and the alcohol dehydrogenase genes. In: Downey K, Voellmy RW, Ahmad F, Schultz J (eds.) Advances in Gene Technology; Molecular Genetics of Plants and Animals, pp. 311–325. Academic Press, New York (1983).
40. Peacock WJ, Wolstenholme D, Walker J, Singh K, Llewellyn D, Ellis J, Dennis ES: Developmental and environmental regulation of the maize alcohol dehydrogenasee 1 (*Adh1*) gene: promoter-enhancer interactions. In: McIntosh L, Key J (eds.) Plant Gene Systems. AR Liss, New York, (in press).
41. Rigby PWJ, Dieckmann M, Rhodes C, Berg P: Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J Mol Biol 113: 237–251 (1977).
42. Sachs MM, Freeling M: Selective synthesis of alcohol dehydrogenase during anaerobic treatment of maize. Mol Gen Genet 161: 111–115 (1978).
43. Sachs MM, Freeling M, Okimoto R: The anaerobic proteins of maize. Cell 20: 761–767 (1980).
44. Sachs MM, Dennis ES, Gerlach WL, Peacock WJ: Two alleles of maize alcohol dehydrogenase 1 have 3′ structural and poly

(A) addition polymorphisms. Genetics 113: 449–467 (1986).
45. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
46. Schwartz D: The genetic control of alcohol dehydrogenase in maize: gene duplication and repression. Proc Natl Acad Sci USA 56: 1431–1436 (1966).
47. Schwartz D: Genetic control of alcohol dehydrogenase in maize – a competition model for regulation of gene action. Genetics 67: 411–425 (1971).
48. Schwartz D: Comparison of relative activities of maize *Adh1* alleles in heterozygotes- analysis at the protein (CRM) level. Genetics 74: 615–617 (1973).
49. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis J Mol Biol 98: 503–517 (1975).
50. Wahl GM, Stern M, Stark GR: Efficient transfer of large DNA fragments from agarose gels to diazobenzyloxymethyl-paper and rapid hybridisation by using dextran sulphate. Proc Natl Acad Sci USA 76: 3683–3687 (1979).
51. Walker JC, Howard EA, Dennis ES, Peacock WJ: DNA sequences required for anaerobic expression of the maize alcohol dehydrogenase-1 gene. Proc Natl Acad Sci USA 84: 6624–6628 (1987).