Theor Appl Genet (1999) 98: 815–824

E. Huq · S. Harrington · M. A. Hossain
F. Wen · S. R. McCouch · T. K. Hodges

Molecular characterization of *pdc2* and mapping of three *pdc* genes from rice

Received: 19 May1998 / Accepted: 29 September 1998

Abstract The anaerobic fermentation pathway is thought to play an important role under flooding conditions. The pyruvate decarboxylase 2 (*pdc2*) gene that encodes the first enzyme of this pathway has been cloned and characterized from rice. This gene has an open reading frame that putatively encodes a 603 amino-acid-residue protein with a molecular mass of 64 kDa. *pdc2* has five introns dispersed throughout the coding region, which is also true for rice *pdc1*. Although the length of these introns in rice *pdc2* are different from those in rice *pdc1*, they are located in exactly the same positions based on the deduced amino-acid sequences. The temporal and spatial expression patterns of *pdc1* and *pdc2* show that *pdc2* is induced to a higher level during the early period (1.5–12 h) of anoxia than *pdc1*, which is induced more after longer time periods (24–72 h) of anoxia in both shoots and roots. The map positions of the three *pdc* genes have also been determined. Rice *pdc1* is located on chromosome 5 between BCD454A and RZ67, *pdc2* is located on chromosome 3 between RZ329 and RZ313, and *pdc3* is mapped on chromosome 7 distal to RG351.

Key words Pyruvate decarboxylase (*pdc*) · Molecular mapping · Orthologous loci · Rice · Flooding tolerance

Communicated by G. E. Hart

E. Huq[1] · M. A. Hossain[2] · F. Wen[3] · T. K. Hodges (✉)
Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA
Fax: + 765-494-5896
E-mail: hodges@btny.purdue.edu

S. Harrington · S. R. McCouch
Department of Plant Breeding and Biometry, Cornell University, Ithaca, N.Y. 14859, USA

Present addresses:
[1] Department of Plant Biology, University of California at Berkeley/USDA/Plant Gene Expression Center, 800 Buchanan St., Albany, CA 94710, USA
[2] Department of Biochemistry, University of Dhaka, Dhaka 1000, Bangladesh
[3] Department of Biotechnology, Shandong Agricultural University, Taian, Shandong Province 271018, P.R. China

Introduction

Pyruvate decarboxylase (PDC, EC 4.1.1.1) nonoxidatively decarboxylates pyruvate to acetaldehyde which is then reduced to ethanol by alcohol dehydrogenase (ADH). PDC is a critical enzyme in the anaerobic-specific fermentation pathway, and the product of this reaction, acetaldehyde, is very toxic to plants. Reduction of acetaldehyde to ethanol by ADH regenerates NAD^+, which is then utilized in the glycolytic pathway to maintain carbon flow through this pathway under anaerobic conditions. This switching of the energy production pathway from aerobic glycolysis to anaerobic fermentation is one of the major metabolic adaptations that plants undertake when they are submerged or confronted with a lack of oxygen. The importance of increased rates of alcoholic fermentation (AF) under anaerobic conditions has been demonstrated by several experimental observations (Waters et al. 1991; Perata and Alpi 1993; Crawford and Braendle 1996; Setter et al. 1997): (1) enzymes for AF often increase, (2) mutants without ADH die more rapidly during anoxia, (3) increased tolerance to anoxia comes from hypoxic pretreatments and presumably the induction of enzymes of AF, (4) high sugar supply increases survival, presumably due to continued operation of AF, and (5) rates of AF are related to the tolerance of several species to waterlogging or flooding. It has also been suggested that the rate of AF is limited by PDC (Waters et al. 1991).

Roberts et al. (1989) studied the role of ADH in the metabolism and survival of anoxic maize root tips

by comparing the ethanol production of isogenic lines differing in ADH activity over an approximately 200-fold range. They concluded that ADH activity in wild-type maize root tips was not a limiting factor for energy production via fermentation and did not determine viability under anoxia. This conclusion was further supported by Johnson et al. (1994) who showed that 70% of the hypoxically acclimated root tips of *adh1* null maize survived up to 24 h of anoxia, whereas only 10% of the unacclimated root tips survived for 6 h of anoxia. They also concluded that the high levels of ADH activity inducible in acclimated wild-type maize root tips are in excess of those required to increase rates of fermentation. Thus, PDC, being the first enzyme in the AF pathway, may play a key regulatory role. Bucher et al. (1994) have found in tobacco that overexpression of a *pdc* gene from *Zymomonas mobilis* resulted in higher levels of acetaldehyde and ethanol formation, suggesting that PDC is likely to be the key regulator of anaerobic metabolism. Unfortunately, the measurements made in this study were only for the initial 24 h of anoxic treatment which did not allow for an evaluation of tolerance under long-term anoxia. Moreover, rice has different mechanisms of submergence tolerance from tobacco, as evidenced by the fact that it has a well-developed aerenchyma system to maintain an oxygen supply (Perata and Alpi 1993).

Genes encoding PDC have been cloned and characterized from maize (Kelley et al. 1991), yeast (Kellerman et al. 1986), and bacteria (Conway et al. 1987). Recently we have reported the isolation and characterization of two *pdc* cDNAs (Hossain et al. 1994 a; Huq et al. 1995) and two genomic clones (Hossain et al. 1994 b; 1996) from rice, while Rivoal et al. (1997) reported another partial cDNA, called *pdc4*, from rice. In this paper, we provide a characterization of *pdc2* and the relative induction of *pdc1* and *pdc2* over time in both shoots and roots under anaerobic conditions. We also present the map locations of these three genes in rice chromosomes and predict the locations of orthologous loci in maize, oat, and *Triticeae* chromosomes.

Materials and methods

Seedling growth conditions

Rice (*Oryza sativa* L. variety IR54) seeds were stirred in 75% Clorox for 50 min, rinsed three times with tap water followed by a single brief rinsing with 95% ethanol. Finally, these seeds were washed with sterile de-ionized water twice and placed in Petri dishes on two layers of filter paper still moist following sterilization in an autoclave, and seedlings were grown at 25°C in light (12 h photoperiod). After 5–7 days, seedlings were transferred to Magenta boxes (Magenta Corporation, USA) containing 0.5 MS salts (Murashige and Skoog 1962) (5–10 ml solution, just sufficient to cover the root of the seedlings), and were grown for an additional 7 days under the same light conditions.

Screening of the genomic library and construction of a restriction map of *pdc2*

An IR54 genomic library constructed in a lambda GEM 11 vector (Hossain et al. 1996) was screened with the following probes generated from the plasmid pBGS-PDC containing maize *pdc* cDNA (Kelley 1989): a *Bgl*II-*Bam*HI fragment of 702 bp representing the 5′ end of the gene, and a *Hin*dIII-*Eco*RI fragment of 366 bp of the 3′ *pdc* coding region. Identification of the hybridizing clones and subsequent plaque purifications were carried out using standard procedures (Sambrook et al. 1989). Phage DNA isolation was done using a liquid culture method (Sambrook et al. 1989). A *pdc2*-specific probe was made from the 3′-untranslated region of *pdc2* cDNA (Huq et al. 1995) and was used to identify the *pdc2* gene. Phage DNAs were digested with different restriction enzymes, electrophoresed on 0.9% agarose gels, and hybridized with two different probes after Southern blotting. The size of the bands was calculated in comparison with migration of a 1-kb ladder (Life Technologies) run on the same gel, and a restriction map of the positive phage clone, 2B, was constructed.

Cloning and sequencing of *pdc2*

pdc2 was subcloned into the plasmid vector pGEM7Zf(+) (Promega) as three fragments from the positive phage clone 2B: a 2.4-kb *Nsi*I-*Bam*HI fragment representing most of the 5′-upstream region, a 1-kb *Bam*HI-*Nsi*I fragment representing the middle region, and a 2.5-kb fragment containing approximately 1.8 kb of the 3′-coding and untranslated regions. The inserted *pdc* fragments were serially deleted by exonuclease III using the Erase-a-Base protocol of Promega. Both strands of the deleted clones were sequenced using a Sequenase version 2 protocol (United States Biochemicals). When required, synthetic oligonucleotides (19 - mers) were used as primers for further sequencing. In some cases compressions were resolved by performing the reactions at 70°C with *Taq* DNA polymerase (United States Biochemicals) and using the deaza-GTP analog. DNA sequences were analysed using the University of Wisconsin Genetics Computer Group (GCG) package (Devereux et al. 1984) and the DNA Strider program (Marck 1988).

Isolation of total RNA

Two-week-old uniform seedlings of the variety IR54 were transferred to 2-l flasks. The flasks were air-tight and well-ventilated. Seedlings were completely submerged in distilled water through which nitrogen gas (99.5%) was continuously bubbled to produce anaerobic conditions. The anaerobic treatment was in the dark at 25°C. Control seedlings were maintained in the dark at 25°C in Magenta boxes. Roots and shoots including coleoptiles were harvested separately and immersed in liquid nitrogen. Extraction of RNA was carried out according to the procedure of Logemann et al. (1987) with modifications. The frozen tissue (2 gm) was ground in liquid nitrogen in an autoclaved morter and pestle, and transferred to autoclaved plastic centrifuge tubes. Extraction buffer (4 M guanidine hydrochloride, 20 mM MES pH 7.0, 20 mM EDTA, 4 ml/g of tissue) was added to each tube. After vigorous shaking and complete mixing the suspension was extracted once with phenol : chloroform : isoamyl alcohol (25 : 24 : 1) and centrifuged for 45 min at room temperature at 10 000 rpm in a Beckman centrifuge using a JA-20 rotor. Then 0.2 vol of 1 M acetic acid [diethyl pyrocarbonate (DEPC)-treated] were added to the supernatant, and RNA was precipitated by adding 0.7 vol of ethanol after 1-h incubation at −80°C following centrifugation at 10 000 rpm for 10 min. The RNA pellet was subsequently dissolved in DEPC-treated water and re-precipitated using 1/3 vol of 8 M LiCl (overnight at 4°C). The RNA sample was subsequently washed with 80% ethanol and

dissolved in DEPC-water after drying, and then quantified spectrophotometrically.

Synthesis of antisense probes

Gene-specific antisense RNA probes were synthesized for use in ribonuclease protection assays using the Maxiscript II System (Ambion). *pdc1* (Hossain et al. 1994 a), *pdc2* (Huq et al. 1995), and *ubi1* (Huq et al. 1997) cDNAs were all subcloned into the pSPORT1 plasmid vector (Life Technologies) as *Eco*RI-*Not*I fragments. *pdc1* was linearized with *Xho*I, *pdc2* with *Sca*I, and *ubi1* with *Nco*I. These linearized plasmids were transcribed with SP6 RNA polymerase to produce gene-specific antisense RNA probes from the 3′-untranslated regions. Transcription reactions were performed in a 20 µl vol containing 1 µg of linearized DNA template, 1 transcription buffer, 0.5 mM ATP, 0.5 mM GTP, 0.5 mM UTP, 10 units of SP6 RNA polymerase, and 3.5 µM of [α-^{32}P]CTP (800 Ci/mmol, 20 mCi/ml), and then incubated for 1 h at 37°C. Following transcription, all reactions were treated with RNase-free DNase I (0.4 units/µl, Ambion) for 15 min at 37°C to remove contaminating template DNA. Full-length probes were purified from 5% denaturing polyacrylamide gels as described (Sambrook et al. 1989).

Ribonuclease protection assays

Ribonuclease protection assays (RPAs) were performed using total RNA isolated from shoots and roots of the 2-week-old IR54 variety treated under anaerobic conditions for 0, 1.5, 3, 6, 12, 24, 48, and 72 h time periods. Assays were carried out using the RPAII System (Ambion) according to the protocol supplied by the company. Five micrograms of total RNA samples for each time point were hybridized with $2–4 \times 10^4$ CPM of high-activity (typically $6–7 \times 10^8$ cpm/µg) antisense probes specific for either *pdc1* or *pdc2*. A rice *ubi1* cDNA isolated from variety IR54 (Huq et al. 1997) was used as an internal control to show the amount of RNA employed. Ribonuclease digestions were performed using 1 : 100 dilutions of the supplied ribonuclease cocktail (final concentration = 0.2 units/ml RNase A; 2.0 units/ml RNase T1). Following RNase treatment, total samples were subjected to electrophoresis through 5% denaturing polyacrylamide gels, wrapped in plastic wrap, and exposed to X-ray film with an intensifying screen for autoradiography. The expected sizes of the full-length and protected fragments using the various gene-specific probes were as follows: 343 and 286, respectively, for *pdc1*; 235 and 178, respectively, for *pdc2*; 300 and 243, respectively, for rice *ubi1* (internal control).

RFLP mapping

A mapping population consisting of 113 backcross (BC) individuals derived from the cross *O. sativa*/*O. longistaminata*//*O. sativa*, maintained at Cornell University, was used to place *pdc1*, *pdc2*, and *pdc3* onto rice chromosome maps (Causse et al. 1994). DNA from the *indica* recurrent parent (BS125) and the interspecific F_1 (BS125/WLO2) was digested with six restriction enzymes (*Eco*RV, *Hin*dIII, *Xba*I, *Sca*I, *Dra*I, *Eco*RI), electrophoresed overnight on 0.9% agarose gels, and blotted onto Hybond N$^+$ (Amersham Corporation) according to the manufacturer's instructions and used in parental polymorphism surveys. For mapping purposes, DNA from 113 backcross progeny was digested and blotted using the same procedures.

The following were used as probes: (1a) *pdc1*, a 2.2-kb cDNA in a pSPORT1 vector (Life Technologies) that contained an open reading frame for *pdc1* which is highly homologous to the other *pdc*'s and is not specific for *pdc1* (hereafter referred to as clone 1A), (1b) *pdc1*, a 115-bp *Mlu*I-*Nco*I fragment of the 5′ untranslated region specific for *pdc1* (hereafter referred to as clone 1B), (2a) *pdc2*, a 700-bp 3′ fragment of *pdc2* cDNA in a pSPORT1 vector (Life Technologies) (hereafter referred to as clone 2A), (2b) *pdc2*, a 200-bp *Sca*I-*Not*I fragment from the 3′-untranslated region specific for *pdc2* (hereafter referred to as clone 2B) and (3) *pdc3*, a 2.8-kb 5′ fragment of *pdc3* in a pUC19 vector (Life Technologies) (hereafter referred to as clone 3). The fragments were labeled by the random priming method using an Amersham Megaprime Labeling System and used as probes in Southern analyses. Filters were hybridized overnight at 65°C and washed three times for 20 min each at 65°C at successive stringencies of 2×, 1×, and 0.5×SSC (with 0.1% SDS). Labeled filters were exposed to X-ray film with an intensifying screen for 5 days at −80°C. For mapping, segregation in the backcross progeny was scored by the presence or absence of the polymorphic band from *O. longistaminanta*. Linkage analysis was performed using Mapmaker version 2.0 (Lander et al. 1987) on a Macintosh Performa 475. Genetic distance is expressed in Kosambi cM (Kosambi 1944) and the map was constructed using a LOD 2.0 significance threshold.

Results

Isolation and sequence analysis of *pdc2*

Two probes from the maize *pdc* cDNA (Kelley 1989) were used to screen an IR54 rice genomic library, and 18 positive clones that hybridized to both the 5′- and 3′-regions of the maize *pdc* cDNA were identified. Eight of these 18 clones were found to be *pdc3*, and two of them were found to be *pdc1* (Hossain et al. 1996). Two clones, 2B and 14C, were similar in restriction pattern and hybridized to a *pdc2*-specific probe (data not shown). The 2B phage clone was selected for further subcloning and sequencing. In order to subclone *pdc2*, we constructed a partial restriction map of the phage clone using different restriction endonucleases. This restriction map is shown in Fig. 1, and the region that hybridized to the 5′- and 3′-region probes is shown at a larger scale. The 5.9-kb fragment that is shown in larger scale was subcloned as three fragments and was sequenced from both ends. The vector containing *pdc2* was called pRgpdc2. The nucleotide and predicted

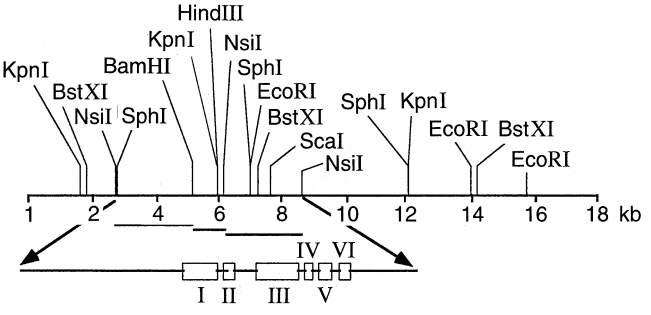

Fig. 1 Restriction map of lambda clone 2B containing the *pdc2* gene. The 5.9-kb region which contained the gene and about 2.38 kb of the 5′ upstream region is shown in larger scale. The *empty boxes* are the exons and the *lines* in between the boxes are the introns. The three fragments that were subcloned and sequenced are shown by *thin lines*. The restriction enzymes used are shown in the figure

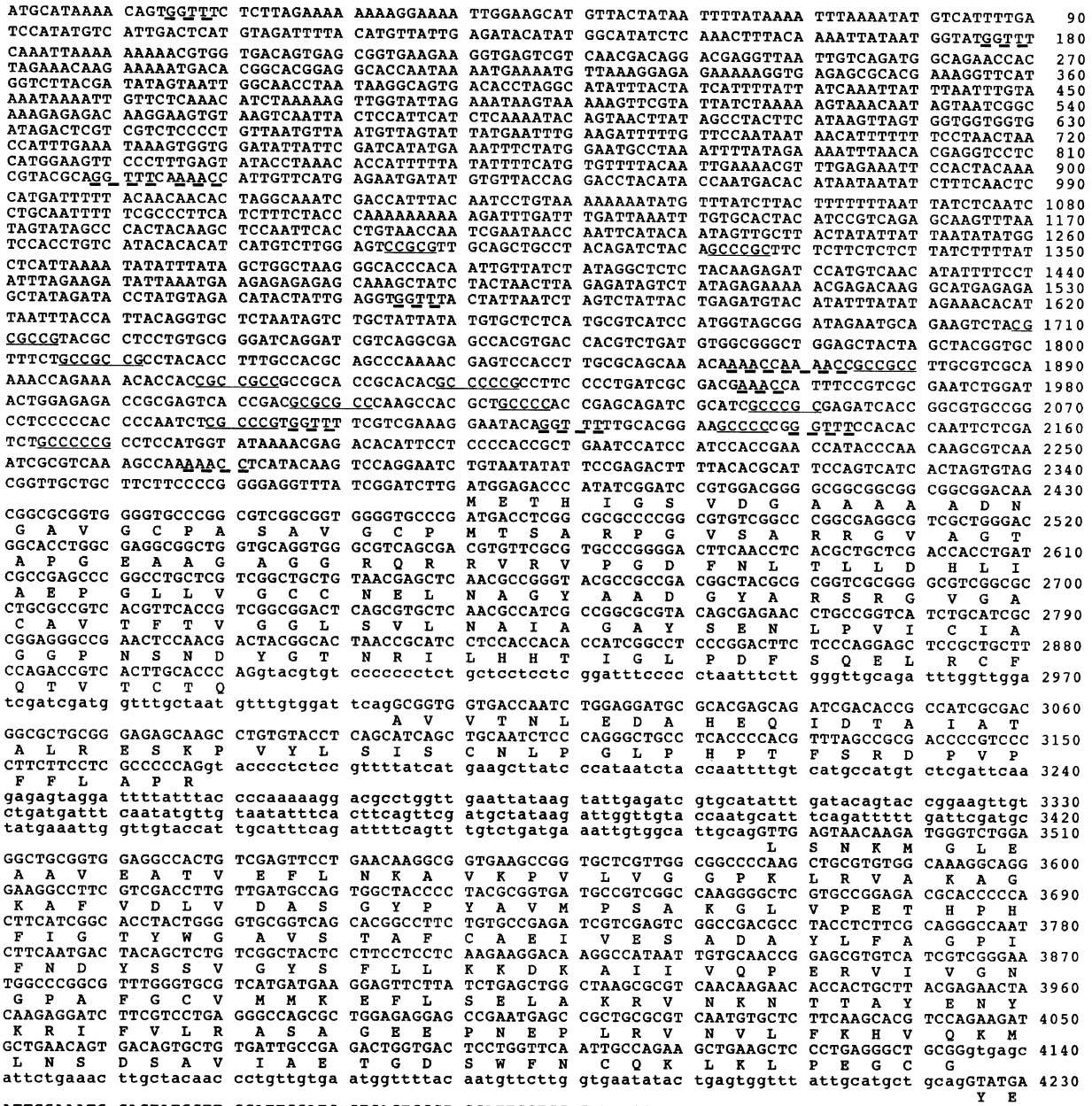

Fig. 2 Nucleotide and deduced amino-acid sequences of the rice *pdc2* gene. The presumptive TATA box is from nucleotides 2180 to 2186. The start and end of the cDNA, pRcpdc2 (Huq et al. 1995), are at nucleotides 2242 and 5070, respectively. The GC- and GT-motif-like sequences are shown by *solid and dashed underlines*, respectively. A G-box-like sequence is present at nucleotides 1752–1759. The GenBank accession number for the sequence reported in this paper is U38199

amino-acid sequences of rice *pdc2*, including an approximately 2.38-kb region upstream of the translational start site, is shown in Fig. 2. The sequence of the coding region exactly matched the sequence of the previously isolated *pdc2* cDNA (Huq et al. 1995) including the 5′- and 3′-untranslated regions. It has an open-reading-frame of 1812 nucleotides that encodes a 603 amino-acid-residue polypeptide with molecular mass of 64 kDa. The sequence also revealed that there were five introns present in the coding region (Fig. 2).

Ribonuclease protection assays

We performed ribonuclease protection assays (RPAs) to determine the spatial and temporal expression patterns of *pdc2* and *pdc1*. RPAs using a *pdc2*-specific probe showed that an approximately 170-bp fragment is protected, which is close to the expected size (178 bp) for the *pdc2*-specific messages (Fig. 3 a). The results also showed that this gene is highly induced within 1.5 h of anoxia in both shoots and roots of IR54. The induction is higher in shoots than in roots. The message levels peaked at around 3 h of anoxia and then gradually decreased in both shoots and roots. The rice *ubi1* gene has been found not to be induced under anoxia. Consequently, the *ubi1*-specific probe was used as an internal control to show the amount of RNA used in each sample. The *ubi1* message levels showed that a little higher amount of RNA was used in the 3-h shoot sample, and that the 24-h root sample was probably degraded.

When we used a *pdc1*-specific probe, an approximately 280-bp fragment was protected, which is close to the expected size (286 bp) for the *pdc1*-specific transcripts (Fig. 3 b). The results showed that *pdc1* is also highly induced under anoxia, and that the induction is greater in shoots than in roots. The message levels peaked around 6 h of anoxia and gradually decreased with the time of anoxia in both shoots and roots. These results also showed that *pdc2* was more intensively induced during the early period of time (1.5–12 h) than *pdc1* which was induced more strongly during the later period of time (24–72 h) under anoxia in both shoots and roots.

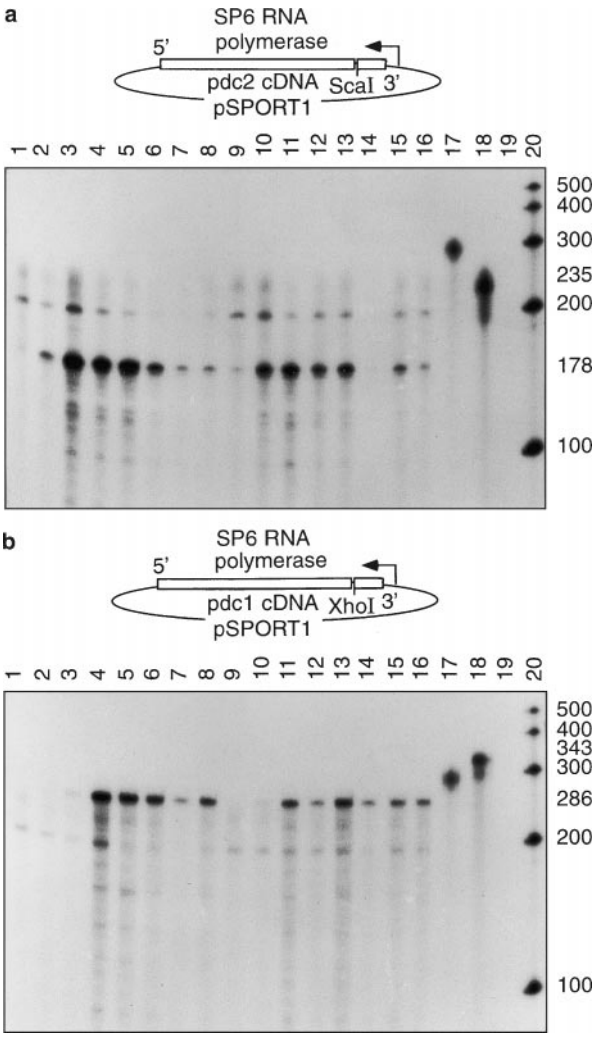

Fig. 3a, b RNase protection assays (RPAs) showing the temporal and spatial expression patterns of rice *pdc2* **a** and *pdc1* **b**. RPAs were performed on 5 μg of total RNA isolated from IR54 seedlings treated under anaerobic conditions for various time periods and were hybridized with gene-specific antisense probes. *Lanes 1–8* are RNA samples from shoots of variety IR54 treated under anaerobic conditions for 0, 1.5, 3, 6, 12, 24, 48, and 72 h. *Lanes 9–16* are RNA samples from roots of variety IR54 treated under anaerobic conditions for 0, 1.5, 3, 6, 12, 24, 48, and 72 h. *Lanes 17–18* are full-length *ubi1*- and *pdc2*-specific probes in (**a**) and *ubi1*- and *pdc1*-specific probes in (**b**), respectively, mixed with 5 μg of yeast RNA digested without RNases. *Lane 19* is a full-length *pdc2*-specific probe in (**a**) and a *pdc1*-specific probe in (**b**) mixed with 5 μg of yeast RNA digested with RNases. *Lane 20* is an RNA Century marker (Ambion) transcribed with T7 RNA polymerase in both (**a**) and (**b**). The sizes of the RNA markers in nucleotides are shown on the right side. The expected sizes of the full-length and protected fragments using the various gene-specific probes were as follows: 343 and 286 bases, respectively, for *pdc1*; 235 and 178 bases, respectively, for *pdc2*; 300 and 243 bases, respectively, for rice *ubi1* (internal control)

Evolutionary relationship of the *pdc*s

The intron lengths and positions were compared among rice *pdc1* (Hossain et al. 1996) and *pdc2*, and maize *pdc1* (Kelley et al. 1991) (Fig. 4). Rice *pdc*s have five introns, whereas maize *pdc1* has only three introns. The length of the first intron in both the rice *pdc1* and *pdc2* are more conserved than the other four introns (Fig. 4 a). The lengths of the three maize introns are not consistent with the corresponding introns of rice *pdc1* and *pdc2* (Fig. 4 a). However, the positions of the corresponding introns are highly conserved in the deduced amino-acid sequences among rice *pdc1* and *pdc2*, and maize *pdc1* (Fig. 4 b).

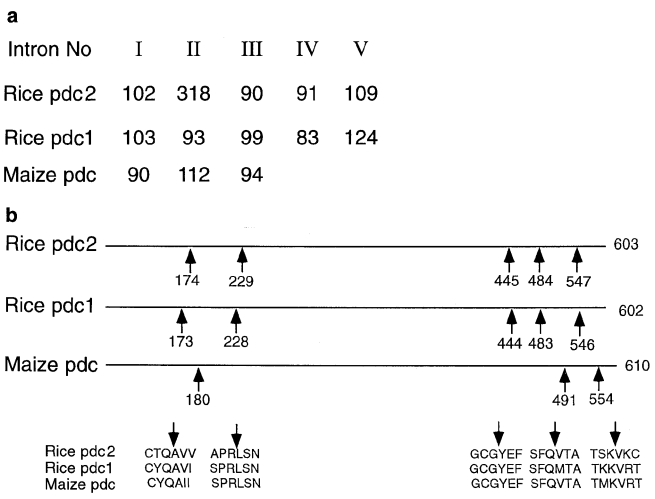

Fig. 4a, b Intron lengths (in nucleotides) **a** and intron positions in the amino-acid sequences **b** of rice *pdc2*, *pdc1*, and maize *pdc1* genes. The amino-acid sequences on either side of the introns are shown in **b**

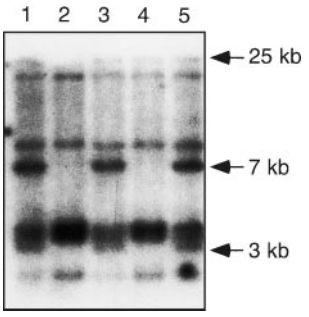

Fig. 5 Autoradiogram of *Sca*I-digested BC_1 individuals from the rice mapping population probed with a 2.2-kb *pdc1* probe in the pSPORT vector. The 7-kb and 3-kb bands (seen in *lanes 1, 3 and 5*) represent the *pdc1* locus and the 25-kb band (*lanes 1, 3, 4 and 5*) represents the *pdc3* locus

Mapping of the *pdc*s

The rice *pdc1* probe (1A) detected a multiple-copy gene family with polymorphic fragments observed in *Eco*RV-, *Hin*dIII-, *Dra*I-, *Xba*I- and *Sca*-I digested parental DNA. The probe was then hybridized to *Sca*I-digested DNA from the mapping population. Three scorable polymorphic fragments were seen that were estimated to be 25 kb, 7 kb, and 3 kb, respectively (Fig. 5). To determine which of these fragments hybridized specifically to *pdc1*, the *pdc1*-specific probe, 1B, was hybridized to DNA digested with the five enzymes mentioned above. In the *Sca*I-digested DNA, only the 7-kb polymorphic fragment gave a hybridization signal (data not shown). This band was mapped to rice chromosome 5 between BCD454A and RZ67 (Fig. 6 a). On mapping filters probed with clone 1A, the 7-kb fragment co-segregated with the 3-kb fragment in *Sca*I-digested DNA. Both mapped to the same location on rice chromosome 5.

When *pdc2* clone 2A was used as a probe on the same parental survey filters, a previously undetected 7.8-kb polymorphic fragment was observed using *Dra*I-digested DNA. To confirm that this fragment was specific to *pdc2*, clone 2B was used as a probe on the same parental survey filters. An identical hybridization pattern was evident. The 7.8-kb fragment was mapped to rice chromosome 3 between RZ329 and RZ313 using *Dra*I-digested DNA from the BC population (Fig. 6 b).

When *pdc3* clone 3 was used to probe *Sca*I-digested DNA, a 25-kb polymorphic fragment was observed. This band was identical in molecular weight to that seen when the *pdc1* 2.2-kb cDNA (clone 1 A) was used as a probe. This band mapped to the bottom of rice chromosome 7 distal to RG351 (Fig. 6 c), suggesting that the *pdc* probes 1A and 3 detected the same 25-kb *Sca*I genomic fragment. Because the *pdc1*-specific probe 1B did not detect this 25-kb band on *Sca*I-digested DNA, it may be concluded that this band is *pdc3*-specific.

Discussion

We have previously reported the isolation and characterization of two *pdc*s and two *pdc* cDNAs from rice (Hossain et al. 1994 a, b, 1996; Huq et al. 1995). The present investigation details the isolation and characterization of another gene, called *pdc2*, from rice. The *pdc2* gene has an open-reading-frame of 1812 nucleotides that presumably encodes a 603 amino-acid-residue polypeptide. The molecular weight of the deduced polypeptide is 64 kDa. A 64-kDa polypeptide was previously purified from rice along with a 62-kDa polypeptide, and it was concluded that the PDC holoenzyme consisted of a tetramer of two different molecular-weight polypeptides in rice (Rivoal et al. 1990, 1997). This provides evidence that the 64-kDa deduced polypeptide encoded by *pdc2* might be a PDC polypeptide. Moreover, both the deduced amino-acid and nucleotide sequences of this gene are highly homologous to those of the rice and maize *pdc1*. Rice PDC2 is 88% similar and 78% identical to rice PDC1, and 88% similar and 79% identical to the maize PDC1 enzyme. *pdc2* also has five introns as found in rice *pdc1*. The exon-intron splice junctions conform to the consensus sequences found in plants (Hanley and Schuler 1988). While, the translation initiation site of this gene does not conform to the consensus sequences for plant translational initiation sites (Joshi 1987) or the Kozak consensus sequences (Kozak 1981), this is not unexpected because several different variations have been found in these sequences in plants (Joshi 1987).

The open-reading-frame of *pdc2* exactly matched with a previously isolated cDNA including the 5′- and 3′-untranslated regions (Huq et al. 1995). We have sequenced approximately 2.38 kb upstream of the translational start site (Fig. 2). A TATA-box-like sequence has been found from nucleotides 2180 to 2186 which is 65 nucleotides upstream of the start site of the

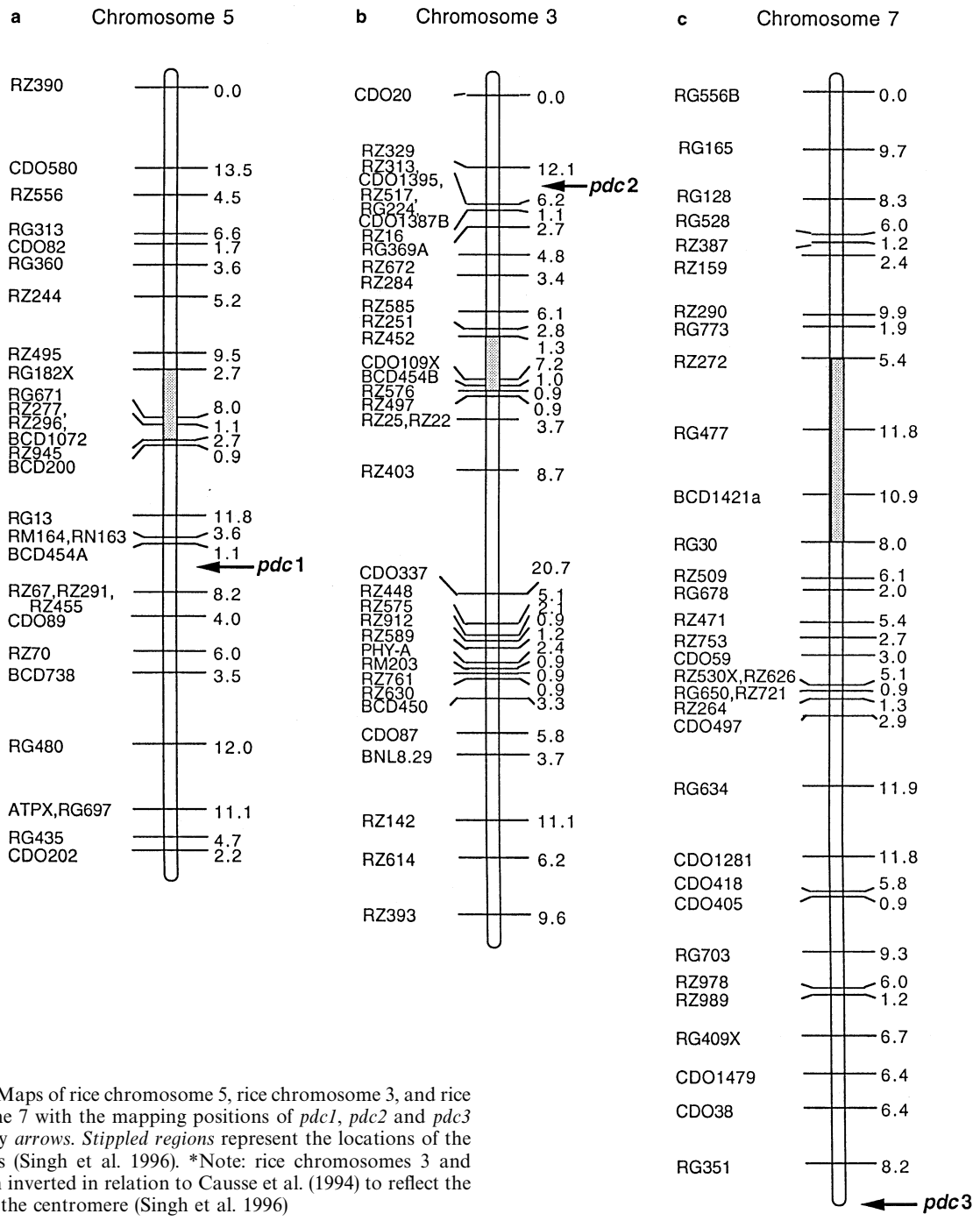

Fig. 6 a–c Maps of rice chromosome 5, rice chromosome 3, and rice chromosome 7 with the mapping positions of *pdc1*, *pdc2* and *pdc3* indicated by *arrows*. *Stippled regions* represent the locations of the centromeres (Singh et al. 1996). *Note: rice chromosomes 3 and 7 have been inverted in relation to Causse et al. (1994) to reflect the position of the centromere (Singh et al. 1996)

pRcpdc2 cDNA. This suggests that the cDNA was near full-length and the sequence from 2180 to 2186 might be the actual TATA box. Moreover, when this 2.38-kb region was fused to *uid*A the resulting construct showed GUS activity after being shot into rice suspension cells (data not shown). The 2.38-kb region has multiple copies of the GT- and GC-motif-like sequences and one G-box-like sequence (Fig. 2). The GT and GC motifs whose core sequences are 5′-GGTTT-3′ and 5′-GC(G/C)CC-3′, respectively, are found in many anaerobically inducible genes (Olive et al. 1990, 1991), including rice *pdc1* (Hossain et al. 1996), and the G-box-like sequence has been found to be involved in altering expression in response to many different environmental stresses (Ferl and Laughner 1989; Dolferus et al. 1994; de Vetten and Ferl 1994). These observations require further experimental evidence to confirm the exact role of these boxes in the *pdc2* promoter.

The spatial and temporal expression patterns of *pdc2* and *pdc1* were investigated by RNase protection assays

(Fig. 3 a, b). The expression patterns showed that both genes were highly induced in both shoots and roots. Moreover, *pdc2* was induced more intensively in shoots and roots during the early phase of anoxia (1.5–12 h) than *pdc1*, which was induced slightly more during the later period of anoxia (24–72 h). This was in contrast to our previous results where we showed that the *pdc1* gene was more inducible than *pdc2* using antisense RNA probes from 5′- or 3′-untranslated regions of *pdc1* or *pdc2*, respectively, on Northern blots (Hossain et al. 1996). Recently, Rivoal et al. (1997) reported isolation of another partial rice cDNA, called *pdc4*, which is 96 and 95% identical to our pRcpdc1 (Hossain et al. 1994 a) in nucleic-acid and amino-acid sequences in the coding region, respectively. Since *pdc1* and *pdc*4 are highly homologous, the hybridizing signals detected by the 5′-untranslated region probe in our previous report (Hossain et al. 1996) might represent the combined signal of *pdc1* and *pdc4*, which could not be distinguished on Northern blots. This observation was supported by the work of Rivoal et al. (1997) as these authors also could not distinguish hybridization between these two genes. Since RPAs are very sensitive and more specific than Northern blots, this cross-hybridization between *pdc1* and *4* was eliminated and the results in Fig. 3 b represent only the induction of *pdc1*. Additional evidence was found in support of this conclusion when we compared the number of GC- and GT-boxes present in the promoter regions of *pdc1* and *pdc2*. The *pdc1* gene has five copies of each of these boxes in the 1.3-kb promoter region (Hossain et al. 1996), while *pdc2* has ten copies of each of these boxes within the 1-kb promoter region (Fig. 2). These boxes have been shown to be involved in the anaerobic induction of reporter gene expression (Olive et al. 1990, 1991) and, hence, the presence of these boxes in higher copies in the *pdc2* promoter than in the *pdc1* promoter also supports the higher induction of *pdc2*. Therefore, we believe that *pdc2* is more inducible than *pdc1* in both shoots and roots under anaerobic conditions. The fainter and smaller bands observed in both Figs. 3 a and b might arise from hybridization between radiolabelled probes and partially degraded RNAs, as well as from degraded cross-hybridization products among other genes.

The evolutionary relationship of the *pdc*s was studied by comparing their intron positions and intron lengths (Fig. 4 a and b). Although intron lengths were found to be variable among rice *pdc1* and *pdc2*, and maize *pdc1*, the intron positions were found to be highly conserved among these three genes. This might indicate an evolutionary conservation of the functional domains, or even the whole protein in rice and maize.

The map locations of three *pdc*s have been determined in the rice genome. The mapping of these genes confirmed the presence of three independent *pdc* loci in rice supporting our previous results that *pdc* is encoded by three or more genes in rice (Hossain et al. 1996). The use of gene-specific probes was essential for resolving the locations of each *pdc*. *pdc1* has an internal *Sca*I site at the 3′-end (Hossain et al. 1996). Thus, co-segregation of the *pdc1*-specific 7-kb band and the 3-kb band that hybridized only to probe 1A on the *Sca*I digested DNA suggested that this 3-kb band might represent the 3′-end of *pdc1*. Mapping of this 3-kb band on the same position as the *pdc1*-specific band also supports this conclusion.

The orthologous counterpart to rice *pdc1* in maize would be predicted to lie on either of two internally duplicated regions of maize chromosome 6 or 8, based on the alignment of comparative rice-maize maps (Paul et al. 1997). As the maize *pdc1* locus has previously been mapped to a heterologous region in the middle of maize chromosome 8 between umc173 and umc12 (Peschke and Sachs 1993), these two *pdc1* genes appear to represent homologues in the rice and the maize genomes. Branching out to other members of the Gramineae, an orthologous locus would be predicted to be located near the middle of oat linkage group A (VanDeynze et al. 1995 a; Paul et al. 1997) and near the centromere on *Triticeae* chromosome 1 (VanDeynze et al. 1995 b; Paul et al. 1997; Figs. 1A, B, and C).

Rice *pdc2* mapped to rice chromosome 3 and the predicted orthologous locus in maize would be expected to lie on one of the duplicated regions of maize chromosome 1 or 9. However, a locus known as maize *pdc2* has been mapped to maize chromosome 8 (within 40 cM of maize *pdc1*) (Peschke and Sachs 1993) in a region that shows no evidence of conserved linkage with the markers near the *pdc2* locus in rice. Therefore, we have no data to suggest that the maize *pdc2* locus is orthologous to this newly mapped *pdc2* locus in rice. Based on the position of rice *pdc2* on chromosome *3* in this study, we would still predict that if there is an orthologous counterpart to this gene in maize, it would lie on either chromosome 1 or 9, and on linkage group F in oat and chromosome 4 in the *Triticeae*.

Rice *pdc3* mapped to the distal portion of rice chromosome 7. Based on the placement of the closest linked marker mapped in both maize and rice (CDO38), the predicted orthologue in maize would be expected to lie on either maize chromosome 1 or 7. The report by Peschke and Sachs (1993) indicates that maize *pdc3* lies in the predicted region of maize chromosome 1; therefore, we conclude that these *pdc3* genes are probably orthologous.

We have compared the map positions of three *pdc*s to that of the *sub1* locus (Xu and MacKill 1996) and other QTLs for submergence tolerance (Nandi et al. 1997) in rice. Only *pdc3* is located on chromosome 7, in which a QTL has also been mapped (Nandi et al. 1997). However, *pdc3* is located, towards the end of chromosome 7 whereas the QTL is located in the middle of the same chromosome suggesting that they may not be the same locus. Submergence tolerance has been found to be a complex trait, and significant additive and

non-additive gene action, as well as the involvement of a few to several genes, has been reported (Mohanty and Khush 1985; Haque et al. 1989). Therefore, the apparent lack of correspondence in the map locations of *pdc*s to that of the *sub1* locus and other QTLs may not be sufficient to reach any conclusion about the role of *pdc*s in submergence tolerance.

In summary, our study on the *pdc* family in rice is of interest for two reasons: it provides a model for studying the evolution of gene families in related genera, in addition to its important biological role in anaerobic-specific fermentation. These genes also provide useful targets for engineering submergence tolerance in rice and other cereals.

Acknowledgments This research was supported by The Rockefeller Foundation under the Rice Biotechnology program and as Rockefeller fellowships to M.A.H., F.W., and E.H. The authors wish to thank Dr. Jeffrey R. Vincent for assisting in computer analyses of the sequences. This is journal paper number 15868 of the Purdue University Agricultural Experimentation Station.

Reference

Bucher M, Brandle R, Kuhlemeier C (1994) Ethanolic fermentation in transgenic tobacco expressing *Zymomonas mobilis* pyruvate decarboxylase. EMBO J 13 : 2755–2763

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE, Second G, McCouch SR, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138 : 1251–1274

Conway T, Osman YA, Konnan JI, Hoffman EM, Ingram LO (1987) Promoter and nucleotide sequences of the *Zymomonas mobilis* pyruvate decarboxylase. J Bacterial 169 : 949–954

Crawford RMM, Braendle R (1996) Oxygen deprivation stress in a changing environment. J Exp Bot 47 : 145–159

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

Dolferus R, Jacobs M, Peacock WJ, Dennis ES (1994) Differential interactions of promoter elements in stress responses of the *Arabidopsis adh* gene. Plant Physiol 105 : 1075–1087

Ferl RJ, Laughner BH (1989) In vivo detection of regulatory factor binding sites of *Arabidopsis thaliana adh*. Plant Mol Biol 12 : 357–366

Hanley BA, Schuler MA (1988) Plant intron sequences: evidence for distinct groups of introns. Nucleic Acids Res 16 : 7159–7176

Haque QA, Hillerislambers D, Tepora NM, Delacruz QD (1989) Inheritance of submergence tolerance in rice. Euphytica 41 : 247–251

Hossain MA, Huq E, Hodges TK (1994 a) Sequence of a cDNA from *Oryza sativa* (L.) encoding the pyruvate decarboxylase 1 gene. Plant Physiol 106 : 799–800

Hossain MA, McGee JD, Grover A, Dennis ES, Peacock WJ, Hodges TK (1994 b) Nucleotide sequence of a rice genomic pyruvate decarboxylase gene that lacks introns: a pseudo gene? Plant Physiol 106 : 1697–1698

Hossain MA, Huq E, Grover A, Dennis ES, Peacock WJ, Hodges TK (1996) Characterization of pyruvate decarboxylase genes from rice. Plant Mol Biol 31 : 761–770

Huq E, Hossain MA, Hodges TK (1995) Cloning and sequencing of a cDNA encoding the pyruvate decarboxylase 2 gene (Accession No. U27350) from rice. Plant Physiol 109 : 722

Huq, E, Hirayama, L, Hossain, MA, Hodges, TK (1997) Characterization of a cDNA encoding the polyubiquitin gene in rice (Accession No. U37687). Plant Physiol 113 : 305

Johnson JR, Cobb BG, Drew MC (1994) Hypoxic induction of anoxia tolerance in roots of *adh1* null *Zea mays* L. Plant Physiol 105 : 61–67

Joshi CP (1987) An inspection of the domain between the putative TATA box and the translation start site in 79 plant genes. Nucleic Acids Res 15 : 6643–6653

Kellerman E, Seeboth PG, Hollenberg CP (1986) Analysis of the primary structure and promoter function of a pyruvate decarboxylase gene (PDC1) from *Saccharomyces cerevisiae*. Nucleic Acids Res 14 : 8963–8977

Kelley PM (1989) Maize pyruvate decarboxylase mRNA is induced anaerobically. Plant Mol Biol 13 : 213–222

Kelley PM, Godfrey K, Lal SK, Alleman M (1991) Characterization of the maize pyruvate decarboxylase gene. Plant Mol Biol 17 : 1259–1261

Kosambi DD (1944) The estimation of map distance from recombination values. Ann Eugen 12 : 172–175

Kozak M (1981) Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res 9 : 5233–5252

Lander S, Green P, Abrahamson J, Barlow A, Day MJ (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1 : 174–181

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Anal Biochem 163 : 16–20

Marck C (1988) 'DNA strider' : a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res 16 : 1829–1836

Mohanty HK, Khush GS (1985) Diallel analysis of submergence tolerance in rice, *Oryza sativa* L. Theor Appl Genet 70 : 467–473

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plant 15 : 473–497

Nandi S, Subudhi PK, Senadhira D, Manigbas NL, Sen-Mandi S, Huang N (1997) Mapping QTLs for submergence tolerance in rice by AFLP analysis and selective genotyping. Mol Gen Genet 255 : 1–8

Olive MR, Walker JC, Singh K, Dennis ES, Peacock WJ (1990) Functional properties of the anaerobic responsive element of the maize *adh1* gene. Plant Mol Biol 15 : 593–604

Olive MR, Peacock WJ, Dennis ES (1991) The anaerobic response element contains two GC-rich sequences essential for binding a nuclear protein and hypoxic activation of the maize *adh1* promoter. Nucleic Acids Res 19 : 7053–7060

Paul E, Blinstrub M, McCouch SR (curators) (1997) Rice genes. Database accessible via World Wide Web, http://probe.nalusda.gov:8300

Perata P, Alpi A (1993) Plant responses to anaerobiosis. Plant Sci 93 : 1–17

Peschke VM, Sachs MM (1993) Multiple pyruvate decarboxylase genes in maize are induced by hypoxia. Mol Gen Genet 240 : 206–212

Rivoal J, Ricard B, Pradet A (1990) Purification and partial characterization of pyruvate decarboxylase from *Oryza sativa* L. Eur J Biochem 194 : 791–797

Rivoal J, Thind S, Pradet A, Rivard B (1997) Differential induction of pyruvate decarboxylase subunits and transcripts in anoxic rice seedlings. Plant Physiol 114 : 1021–1029

Roberts JKM, Chang K, Webster C, Callis J, Walbot V (1989) Dependence of ethanolic fermentation, cytoplasmic pH regulation, and viability on the activity of alcohol dehydrogenase in hypoxic maize root tips. Plant Physiol 89 : 1275–1278

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Setter TL, Ellis M, Laureles EV, Ella ES, Senadhira D, Mishra SB, Sarkarung S, Datta S (1997) Physiology and genetics of submergence tolerance in rice. Ann Bot 79 : 67–77

Singh K, Ishii T, Parco A, Huang N, Brar DS, Khush GS(1996) Centromere mapping and orientation of the molecular linkage map of rice (*Oryza sativa* L.) Proc Natl Acad Sci USA 93 : 6163–6168

VanDeynze AE, Nelson JC, O'Donoughue LS, Ahn SN, Siripoonwiwat W, Harrington SE, Yglesias ES, Braga DP, McCouch SR, Sorrells ME (1995 a) Comparative mapping in grasses. Oat relationships. Mol Gen Genet 249 : 349–356

VanDeynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995 b) Comparative mapping in grasses. Wheat relationships. Mol Gen Genet 248 : 744–754

Vetten NC de, Ferl RJ (1994) Transcriptional regulation of environmentally inducible genes in plants by an evolutionarily conserved family of G-box binding factors. Int J Biochem 26 : 1055–1068

Waters I, Morrell S, Greenway H, Colmer TD (1991) Effects of anoxia on wheat seedlings. J Exp Bot 42 : 1437–1447

Xu K, Mackill DJ (1996) A major locus for submergence tolerance mapped on rice chromosome 9. Mol Breed 2 : 219–224