NADP-Malate Dehydrogenase Gene Evolution in Andropogoneae (Poaceae): Gene Duplication Followed by Sub-functionalization

P. RONDEAU1, C. ROUCH1 and G. BESNARD2,*

1Université de la Réunion, LBGM, 15 Avenue R. Cassin, 97715 St-Denis Messag Cedex 9, La Réunion, France and 2Department of Ecology and Evolution, Biology Building, University of Lausanne, 1015 Lausanne, Switzerland

Received: 22 November 2004 Returned for revision: 18 May 2005 Accepted: 15 September 2005 Published electronically: 21 October 2005

INTRODUCTION

C4 photosynthesis provides a selective advantage for plants growing in tropical areas (Sage, 2004). The C4 pathway serves as a pump, which provides an elevation of the CO2 concentration at the site of Rubisco, leading to limited photorespiration and high water use efficiency in C4 species (von Caemmerer and Furbank, 2003). This pathway has appeared several times during plant evolution from the Oligocene and Miocene, at a period when atmospheric CO2 levels were gradually declining (Edwards et al., 2001; Sage, 2004). C4 genes are of great interest since an aim of biotechnology is to introduce the C4 traits to cultivated C3 plants, possibly improving their performance under warm conditions (Matsuoka et al., 2001). Phylogenetic and biochemical studies of C4 genes have also provided insights into the molecular evolution of the C4 pathway (e.g. Sheen, 1999; Gehrig et al., 2001; Besnard et al., 2003; Monson, 2003; Westhoff and Gowik, 2004). In Poaceae, multiple appearances of C4 photosynthesis during the PACCAD (Panicoideae, Arundinoideae, Chloridoideae, Centothecoideae, Aristidoideae, Danthonioideae) clade diversification have probably occurred (Kellogg, 2000). Using a phylogenetic approach on the grass phosphoenolpyruvate carboxylase multigene family, it has been proposed that such multiple appearances should be associated with multiple inductions of a pre-existing system (Besnard et al., 2003). Study of other multigene families including a C4 isoform (NADP-malate dehydrogenase, pyruvate orthophosphate dikinase) is required to understand thoroughly how the C4 pathway evolved during grass diversification.

In plants, plastid NADP-dependent malate dehydrogenases (NADP-MDHs; EC 1.1.1.82) are essential enzymes involved in photosynthesis (Scheibe, 1987; Sheen, 1999). These enzymes are implicated in two main functions. First, in all plants, NADP-MDH helps control export of the reducing equivalents (malate) from the chloroplast to the cytosol (Scheibe, 1987; Ocheretina et al., 2000). Secondly, NADP-MDH is also involved in photosynthetic carbon assimilation of both C4 NADP-malic enzyme type and C4 phosphoenolpyruvate carboxykinase (PCK) enzyme type plants because it converts oxaloacetate to malate, which is then transferred from the mesophyll cells to the bundle sheath cells (von Caemmerer and Furbank, 2003). NADP-MDH is a redox-regulated enzyme (Miginiac-Maslow et al., 2000). Redox regulation appeared relatively late in evolution (probably in green algae; Ocheretina et al., 2000). NADP-MDH is a highly conserved enzyme in plants, and is apparently encoded by a small multigene family...
(Luchetta et al., 1991). Phylogenetic trees based on this gene family should be appropriate for increasing our understanding of the evolution of NADP-MDH isoforms and particularly to study the appearance(s) of the C_4 NADP-MDH isoform during plant evolution.

In maize and sorghum, two grasses belonging to tribe Andropogoneae of subfamily Panicoideae (GPWG, 2001), NADP-MDH genes have been characterized (Metzler et al., 1989; Luchetta et al., 1991). Only one NADP-MDH nuclear gene was identified in maize (Metzler et al., 1989) and in the complete genome sequence of rice (Kikuchi et al., 2003). Conversely, in sorghum, the existence of two tandemly repeated genes encoding NADP-MDH was demonstrated (Luchetta et al., 1991). The regulation of the two genes was differential, with one being light induced (C_4 gene; NMDH-I) and the other (NMDH-II) being constitutively expressed at a low level. The function of the latter was not clearly elucidated but it could be involved in the export of reducing equivalents. Based on these data, Luchetta et al. (1991) postulated that the grass C_4 NADP-MDH gene could have been derived from a non-C_4 NADP-MDH. In contrast, McGonigle and Nelson (1995) proposed that only changes in the regulation of a pre-existing gene were involved in the C_4 NADP-MDH appearance in the genus Flaveria (Asteraceae), without gene duplication during the evolution from C_3 to C_4 plants. Only a few DNA sequences of NADP-MDH genes are currently available, and consequently the evolution of the NADP-MDH multigene family is still not well understood (Ocheterina et al., 2000).

In this study, PCR-based cloning was used to isolate cDNAs encoding NADP-MDHs from 15 species belonging to Panicoideae. From green leaves, cDNAs related to the two sorghum isoforms were isolated. A phylogenetic approach was used to provide insights into the evolution of the NADP-MDH multigene family in Panicoideae. Episodic selection along some branches of the phylogenetic tree was tested. In addition, an analysis of transcription specificities of NADP-MDH genes was conducted for three genera of Andropogoneae.

**MATERIALS AND METHODS**

**Molecular analysis**

In order to isolate cDNA sequences encoding NADP-MDHs in grasses, 15 species belonging to Panicoideae were analysed (Table 1). Twelve are C_4 NADP-malic enzyme type plants (representatives of Andropogoneae plus *Paspalum panicum* and *Paspalum geminatum*), *Urochloa maxima* (Jacq.) R. Webster and *Melinis repens* are C_4 PCK enzyme type plants, and *Oplismenus compositus* is a C_3 plant. For each accession, green leaves were sampled at noon. Total RNA was extracted and then reverse transcription reactions were performed as previously described (Besnard et al., 2002).

From the available NADP-MDH sequences of sorghum [EMBL accessions X53453 (NMDH-I) and S55884 (NMDH-II); Luchetta et al., 1991] and maize (EMBL accession X16084; Metzler et al., 1989), specific primer pairs were designed to amplify cDNA segments of NADP-MDH using reverse transcription–polymerase chain reaction (RT–PCR). Primers were defined in conserved regions of available grass NADP-MDH genes. Two forward primers were defined: 5' GAGTGCTTCGGGGTCTTCTGC 3' (A) and 5' AATGGCCCTGTCACAGYTTATCC 5' (B). The same reverse primer was used in all reactions: 5' TACATTTCTCTGTTAGCATGTATC 3' (R). Long cDNA segments

**Table 1.** Species origin, accession number and reference of NADP-MDH sequences used for analyses

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession number</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poaceae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-family Panicoideae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tribe Andropogoneae</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Dichanthium aristatum</em> (Poir.) C. Hubb.</td>
<td>AJ512368</td>
<td>This study</td>
</tr>
<tr>
<td><em>Heteropogon contortus</em> (L.) P. Beauv.</td>
<td>AJ621582</td>
<td>This study</td>
</tr>
<tr>
<td><em>Hyparrhenia rufa</em> (Nees) Stapf</td>
<td>AJ621580</td>
<td>This study</td>
</tr>
<tr>
<td><em>Ischaemum coleostachys</em> (Steu.) Hack.</td>
<td>AJ512369</td>
<td>This study</td>
</tr>
<tr>
<td><em>Pogonatherum panicum</em> (Lam.) Hack.</td>
<td>AJ512373</td>
<td>This study</td>
</tr>
<tr>
<td><em>Saccharum officinarum</em> L.</td>
<td>AJ344432, AJ565857, AJ33960</td>
<td>This study</td>
</tr>
<tr>
<td><em>Saccharum spontaneum</em> L.</td>
<td>AJ416567, AJ565856</td>
<td>This study</td>
</tr>
<tr>
<td><em>Sorghum arundinaceum</em> (Desv.) Stapf</td>
<td>AJ512374</td>
<td>This study</td>
</tr>
<tr>
<td><em>Sorghum bicolor</em> (L.) Moench</td>
<td>XS3453, S55884, M31965</td>
<td>This study</td>
</tr>
<tr>
<td><em>Themeda quadrivalvis</em> (L.) Kuntze</td>
<td>AJ512376</td>
<td>This study</td>
</tr>
<tr>
<td><em>Vetiveria zizanioides</em> (L.) Nash</td>
<td>AJ344433, AJ512375</td>
<td>This study</td>
</tr>
<tr>
<td><em>Zea mays</em> L.</td>
<td>X16084</td>
<td>Metzler et al. (1989)</td>
</tr>
<tr>
<td>Tribe Paniceae</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Melinis repens</em> (Willd.) Zizka</td>
<td>AJ621581</td>
<td>This study</td>
</tr>
<tr>
<td><em>Oplismenus compositus</em> (L.) P. Beauv.</td>
<td>AJ512370</td>
<td>This study</td>
</tr>
<tr>
<td><em>Paspalidium geminatum</em> (Forsk.) Stapf</td>
<td>AJ33959</td>
<td>This study</td>
</tr>
<tr>
<td><em>Paspalum panicum</em> L.</td>
<td>AJ512372</td>
<td>This study</td>
</tr>
<tr>
<td><em>Urochloa maxima</em> (Jacq.) R. Webster</td>
<td>AJ512371</td>
<td>This study</td>
</tr>
<tr>
<td>Sub-family Erhartiodeae</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Oryza sativa</em> L.</td>
<td>AK105935</td>
<td>Kikuchi et al. (2003)</td>
</tr>
<tr>
<td>Asteraceae</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Flaveria trinervia</em> (Spreng.) C. Mohr</td>
<td>U22533</td>
<td>McGonigle and Nelson (1995)</td>
</tr>
</tbody>
</table>

(Luchetta et al., 1991). Phylogenetic trees based on this gene family should be appropriate for increasing our understanding of the evolution of NADP-MDH isoforms and particularly to study the appearance(s) of the C_4 NADP-MDH isoform during plant evolution.

In maize and sorghum, two grasses belonging to tribe Andropogoneae of subfamily Panicoideae (GPWG, 2001), NADP-MDH genes have been characterized (Metzler et al., 1989; Luchetta et al., 1991). Only one NADP-MDH nuclear gene was identified in maize (Metzler et al., 1989) and in the complete genome sequence of rice (Kikuchi et al., 2003). Conversely, in sorghum, the existence of two tandemly repeated genes encoding NADP-MDH was demonstrated (Luchetta et al., 1991). The regulation of the two genes was differential, with one being light induced (C_4 gene; NMDH-I) and the other (NMDH-II) being constitutively expressed at a low level. The function of the latter was not clearly elucidated but it could be involved in the export of reducing equivalents. Based on these data, Luchetta et al. (1991) postulated that the grass C_4 NADP-MDH gene could have been derived from a non-C_4 NADP-MDH. In contrast, McGonigle and Nelson (1995) proposed that only changes in the regulation of a pre-existing gene were involved in the C_4 NADP-MDH appearance in the genus *Flaveria* (Asteraceae), without gene duplication during the evolution from C_3 to C_4 plants. Only a few DNA sequences of NADP-MDH genes are currently available, and consequently the evolution of the NADP-MDH multigene family is still not well understood (Ocheterina et al., 2000).

In this study, PCR-based cloning was used to isolate cDNAs encoding NADP-MDHs from 15 species belonging to Panicoideae. From green leaves, cDNAs related to the two sorghum isoforms were isolated. A phylogenetic approach was used to provide insights into the evolution of the NADP-MDH multigene family in Panicoideae. Episodic selection along some branches of the phylogenetic tree was tested. In addition, an analysis of transcription specificities of NADP-MDH genes was conducted for three genera of Andropogoneae.

**MATERIALS AND METHODS**

**Molecular analysis**

In order to isolate cDNA sequences encoding NADP-MDHs in grasses, 15 species belonging to Panicoideae were analysed (Table 1). Twelve are C_4 NADP-malic enzyme type plants (representatives of Andropogoneae plus *Paspalum panicum* and *Paspalum geminatum*), *Urochloa maxima* and *Melinis repens* are C_4 PCK enzyme type plants, and *Oplismenus compositus* is a C_3 plant. For each accession, green leaves were sampled at noon. Total RNA was extracted and then reverse transcription reactions were performed as previously described (Besnard et al., 2002).

From the available NADP-MDH sequences of sorghum [EMBL accessions X53453 (NMDH-I) and S55884 (NMDH-II); Luchetta et al., 1991] and maize (EMBL accession X16084; Metzler et al., 1989), specific primer pairs were designed to amplify cDNA segments of NADP-MDH using reverse transcription–polymerase chain reaction (RT–PCR). Primers were defined in conserved regions of available grass NADP-MDH genes. Two forward primers were defined: 5' GAGTGCTTCGGGGTCTTCTGC 3' (A) and 5' AATGGCCCTGTCACAGYTTATCC 5' (B). The same reverse primer was used in all reactions: 5' TACATTTCTCTGTTAGCATGTATC 3' (R). Long cDNA segments
(of approx. 1100 bp with primers AR and 1300 bp with primers BR) were PCR generated. The PCR conditions described in Besnard et al. (2002) were used with an annealing temperature of 50 °C for each primer pair. RT–PCR fragments were cloned before sequencing. The pGEM®-T vector (Promega, Madison, WI, USA) and Escherichia coli strain DH5α were used according to the supplier’s recommendations. For each species of Andropogoneae, distinct sequences were looked for among 15 clones using their restriction profile. The insert in each clone was amplified by PCR using SP6 and T7 primers (Promega). PCR products were restricted with HpaII, separated on a 3 % agarose gel and stained with ethidium bromide. Each clone displaying a different profile was then sequenced. Double-stranded DNA sequencing reactions were performed by the ESGS Society (Euro Sequence Gene Service, Evry, France). Sequences have been deposited in the EMBL databank.

**Phylogenetic analysis**

NADP-MDH cDNA sequences from *Sorghum*, *Zea* and *Oryza* from GenBank were added to the data (Table 1). For *Saccharum* and *Sorghum*, both cultivated forms (*Saccharum officinarum* and *Sorghum bicolor*) and close wild relatives (*S. spontaneum* (= *S. verticilliflorum*)) were studied (Dillon et al., 2001; Hodkinson et al., 2002). Only coding sequences (cds) were used for analysis. In addition, the phylogenetic trees were rooted with an outgroup sequence. Since no complete NADP-MDH sequence from another monocot was available in GenBank, a cDNA sequence from *Flaveria* (U22533; Asteraceae) was used. A sample of 25 sequences (Table 1) displaying the major part of the cds was considered. Sequence alignments were performed using Clustal W (Thompson et al., 1994). Phylogenetic analyses were made using a large internal cDNA segment of 1050 bp in which indels are very infrequent (only one 3 bp deletion was detected in *S. spontaneum* sequences). In the sorghum sequence X53453, the 1050 bp segment is located between nucleotides 220 and 1269. Phylogenetic analyses were conducted using MEGA version 2 (Kumar et al., 2004). Mean frequencies of each nucleotide were similar (T = 25 %, A = 29 %, G = 26 %, C = 20 %), but transitions were twice as frequent as transversions. Consequently, to consider this unequal probability of the substitution types, the Kimura 2-parameter (Kimura, 1980) was used to compute distances between each pair of sequences. Phylogenetic trees were then constructed using the neighbour joining algorithm (NJ; Saitou and Nei, 1987). Bootstrap values were computed using 1000 replicates to evaluate support for the groupings. In addition, a maximum parsimony (MP) analysis was conducted. All characters were equally weighted. Gaps were treated as missing data. A heuristic search was used to find the most-parsimonious trees. The close-neighbour-interchange algorithm was used with a search level of 3 (Kumar et al., 2004). Searches included 100 replications of random addition sequences. All the best trees were retained. A strict consensus tree was generated from the equally most-parsimonious trees. Bootstrap values were computed using 1000 replicates with MEGA (Kumar et al., 2004) to evaluate support for the branches.

**Compared evolution of NADP-MDH genes in Andropogoneae**

Within Andropogoneae, two NADP-MDH genes were identified (NMDH-I and NMDH-II; see below). Using all available sequences, amino acid sites providing a distinction between these two genes were investigated. A segment of 350 amino acids was used. In the *Sorghum* NMDH-I protein (429 amino acids; X53453), this segment is situated between amino acids 70 and 419. In addition, in order to determine possible selective divergence between the two NADP-MDH genes, the rates of non-synonymous and synonymous substitutions between pairs of sequences from *S. officinarum*, *S. bicolor* and *Vetiveria zizanioides* were compared independently for NMDH-I and NMDH-II. The method of Nei and Gojobori (1986) was applied.

Within Andropogoneae, the occurrence of episodic positive selection following gene duplication was also tested on the 1050 bp segment. Eleven sequence accesions from *Saccharum*, *Sorghum*, *Vetiveria*, *Zea* and *Oryza* were considered (M31965, S55884, X16084, X53453, AJ344432, AJ344433, AJ416567, AJ512374, AJ512375, AJ565857 and AK105935). The AJ565856 accession was not used in this analysis since it probably encodes a non-functional enzyme (see below). Furthermore, in the genus *Saccharum*, only one NMDH-I sequence (AJ565857) was used because accessions AJ565857 and AJ833960 only differ by one non-synonymous substitution. The structure of the phylogenetic tree was defined based on an MP analysis as previously described. The $d_S/d_D$ ratios ($\omega$) were estimated for all branches of the tree by using a maximum likelihood (ML) analysis as described by Yang (1998). We used the CODEML program implemented in the PAML package (Yang, 1997) using the ‘free ratio’ model assuming an independent $d_S/d_D$ ratio for each branch of the phylogenetic tree (Yang, 1998). The F3X4 codon model was used; it takes into account differences in transversion $vs.$ transition and codon usages between analysed sequences.

**NADP-MDH gene transcription analysis in Andropogoneae**

To quantify differential NADP-MDH RNA accumulation between different plant materials (i.e. green leaf, etiolated leaf, root and stem) in *S. officinarum*, we used a Northern dot-blot membrane (see Besnard et al., 2003). For each tissue sample of cultivar R570, 10 mg of total RNA was directly located onto an N+-Hybond nylon membrane (Amersham, Bucks, UK). An NADP-MDH segment of 357 bp (accession AJ320266) was hybridized as a probe using the procedure described in Besnard et al. (2003). In *Saccharum*, this probe displays great similarity (at least 94 %) with both genes NMDH-I and NMDH-II. Consequently, this probe cannot be considered as specific to one NADP-MDH gene. The membrane was also independently hybridized with a ribosomal 18S probe. Ribosomal 18S gene expression was considered to be constitutive and
RESULTS

Data sequences and phylogenetic relationships between NADP-MDH

Nineteen cDNA sequences encoding NADP-MDHs were generated in the species analysed (Table 1). One sequence from *S. spontaneum* (AJ565856) probably corresponds to a pseudo-gene due to the presence of a termination codon in the middle of the sequence. The phylogenetic trees based on the sequence analysis of the 1050 bp segment are presented in Fig. 1. The different methods used to reconstruct phylogenies led to similar results. *Oryza*, as expected with this level of sampling, is sister to Panicoideae. In a non-monophyletic tribe Paniceae, *Paspalum* is separated from *Paspalidium, Melinis, Urochloa* and *Oplismenus*, although with only weak bootstrap support. *Paspalum* is sister to Andropogoneae in the MP tree, whereas it is sister to the rest of the Panicoideae in the NJ tree. Lastly, phylogenetic reconstruction shows that the clade including all NADP-MDHs from Andropogoneae is well supported (>98%; Fig. 1). Nevertheless, in this clade, two groups of sequences can be recognized (named NMDH-I and NMDH-II; Fig. 1B). In each of these two sub-clades, there is at least one sequence from *Saccharum, Sorghum* and *Vetiveria*. The sequence of *Zea* is not clearly assignable to either of these two sub-clades.

Evolution of NADP-MDH genes in Andropogoneae

The comparison of mutation rates between the same species for each NADP-MDH gene revealed that the ratio $\omega (d_{s}/d_{c})$ is 1.3–2.0 times higher in NMDH-II than in NMDH-I (Table 2). This suggests that the evolution of NMDH-II was slightly faster than that of NMDH-I. Using an ML analysis to estimate the $d_{s}/d_{c}$ ratios along phylogenetic branches, a value >1.0 was obtained at the base of the NMDH-II clade ($\omega = 1.60$; Fig. 2). This probably indicates an episodic positive selection on the NMDH-II gene. Within the 350 amino acid sequence, three sites allow a distinction between NMDH-I and NMDH-II. First, at position 329 (in the sorghum sequence X53453), NMDH-I proteins have an alanine whereas all NMDH-II proteins have a valine. The other sequences from Paniceae, *Zea* and *Oryza* also have a valine. (This modification is a conservative mutation that should not induce a major change in the enzyme characteristics.) Secondly, at position 292 (in X53453), a lysine is found in proteins belonging to NMDH-I in Paniceae, *Zea* and *Oryza*, whereas a glutamine or a glutamate is found in NMDH-II proteins. Thirdly, at position 414, an asparagine is found in all proteins belonging to NMDH-I in Paniceae, *Zea* and *Oryza*, whereas an aspartate is found in NMDH-II proteins.

Comparative NADP-MDH gene transcription in Andropogoneae

Using a Northern dot-blot approach, higher NADP-MDH gene transcription was revealed in green leaves of sugarcane cultivar R570 (Fig 3A) than in roots, stem and etiolated leaves. This confirms that the transcription of NADP-MDH genes is light induced in *Saccharum*. It was estimated that the transcript accumulation was about seven times higher in green leaves than in etiolated leaves. Using the semi-quantitative PCR method, it was revealed that both NADP-MDH genes are transcribed in green leaves of *Saccharum, Sorghum* and *Vetiveria* (Fig. 3B), and it was shown that the amount of cDNA was usually higher for NMDH-I than for NMDH-II. However, similar amounts of cDNA were observed for both genes in cultivar R570 (54 and 46%; Fig. 3B).

DISCUSSION

The phylogenetic trees based on NADP-MDH sequences (Fig. 1) allow a distinction to be made between a paraphyletic Paniceae and a monophyletic Andropogoneae in subfamily Panicoideae. The separation of *Paspalum* from other Paniceae analysed (*Paspalidium, Melinis, Urochloa* and
Fig. 1. Phylogenetic relationships between cDNAs encoding NADP-malate dehydrogenase in grasses. The analysis is based on a 1050 bp segment (423 variable sites of which 228 are potentially parsimony informative). (A) An MP tree based on a maximum parsimony analysis (904 steps, CI = 0.64, RI = 0.57). The consensus tree is given (from the three most-parsimonious trees). Species belonging to either Andropogoneae or Paniceae are indicated. (B) An NJ tree based on Kimura 2-parameter distance and the neighbour-joining algorithm. The two gene lineages NMDH-I and NMDH-II are specified. Bootstrap values >50% are indicated on each tree.
**Table 2.** Non-synonymous \((d_S)\) and synonymous \((d_S)\) mutation rates computed by comparison of pairs of sequences from *Saccharum officinarum*, *Sorghum bicolor* and *Vetiveria zizanioides*

<table>
<thead>
<tr>
<th>NMDH gene</th>
<th>Pair of sequences</th>
<th>(d_S)</th>
<th>(d_S)</th>
<th>(\omega^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMDH-I</td>
<td><em>Saccharum</em> (AJ565857)</td>
<td>0.014</td>
<td>0.098</td>
<td>1.43</td>
</tr>
<tr>
<td>NMDH-II</td>
<td><em>Saccharum</em> (AJ565857)</td>
<td>0.020</td>
<td>0.105</td>
<td>1.90</td>
</tr>
<tr>
<td>NMDH-I</td>
<td><em>Sorghum</em> (X53453)</td>
<td>0.010</td>
<td>0.095</td>
<td>1.05</td>
</tr>
<tr>
<td>NMDH-II</td>
<td><em>Sorghum</em> (X53453)</td>
<td>0.017</td>
<td>0.081</td>
<td>2.10</td>
</tr>
</tbody>
</table>

The parameters were computed independently for each NADP-MDH gene (NMDH-I and NMDH-II) using the method of Nei and Gojobori (1986). These computations were based on the 1050 bp segment.

\({}^*\omega\) is the ratio \(d_S/d_S\).

**Fig. 2.** An MP tree based on a subset of 11 NADP-MDH cDNAs representative of the gene lineages NMDH-I and NMDH-II in the tribe Andropogoneae (genera *Saccharum*, *Sorghum* and *Vetiveria*). Only one most-parsimonious tree was obtained (289 steps, CI = 0.85, RI = 0.71). Bootstrap values >50% are indicated. The tree topology was used to estimate \(d_S/d_S\) (\(\omega\)) values along each branch based on a maximum likelihood analysis and a free ratio model. The numbers in parentheses are the maximum likelihood estimates of the numbers of synonymous and non-synonymous substitutions, respectively, along the basal branch of NMDH-II.

*Oplismenus*; Fig. 1B) was expected since paraphyly in this tribe has been demonstrated using other gene sequences (e.g. Giussani et al., 2001; Besnard et al., 2002; Duvall et al., 2003). In Andropogoneae, the existence of two distinct genes encoding NADP-MDH (named NMDH-I and NMDH-II) is supported by our study, and this confirms the observation in sorghum (Luchetta et al., 1991). These two genes are closely related, but the sequence from *Zea* falls in an unresolved position relative to the clades representing NMDH-I and NMDH-II (Fig. 1B). This suggests that the gene duplication event leading to the appearance of these two gene lineages is specific to some Andropogoneae (at least in species deriving from the common ancestor of *Saccharum*, *Sorghum* and *Vetiveria*). The duplicate gene probably exists in some of the other genera analysed (*Dichanthium*, *Heteropogon*, *Hyparrhenia*, *Ischaemum* and *Pogonatherum*), but this was not demonstrated in our analysis (probably due to a lower transcription level of NMDH-II than of NMDH-I).

Theoretical models predict that maintenance of duplicate genes could be associated with selective pressures via neo-functionalization, in which one copy acquires a new function, or by sub-functionalization, in which the original function is partitioned across both copies (Monson, 2003; Walsh, 2003). Gene duplication and subsequent modifications are considered to be essential for the multiple C4 pathway appearance. Many of the duplicated genes destined for C4 roles should have been shaped by selection to produce novel C3 roles, prior to being recruited for C4 photosynthesis (Monson, 2003). However, the C4 NADP-MDH isoform selection could be a relatively simple event compared with the selection of other C4 genes (e.g. C4 PEPC genes) since it may not always have involved gene duplication (McGonigle and Nelson, 1995; Monson, 2003). As proposed by McGonigle and Nelson (1995), NADP-MDH enzymes involved in C4 photosynthesis have been selected for their high expression level in the mesophyll cells (as is probably the case in *Zea* in which only one NADP-MDH gene has been identified). Nevertheless, selective pressures favouring new mutations during the evolution from C3 to C4 NADP-MDH are likely to have occurred. Gene duplication should thus be an opportunity to partition the original functions of NADP-MDHs across different copies. In *S. bicolor*, NMDH-I was reported to have a high transcription level in...
green leaves and is likely to be involved in C₄ photosynthesis, whereas NMDH-II, which displays a lower transcription level, could be involved in the reducing equivalent export (Luchetta et al., 1991). Our results (Fig. 3) confirm that the transcript accumulation in green leaves is usually stronger for NMDH-I than for NMDH-II, sustaining the hypothesis that these two genes were maintained in some Andropogoneae due to sub-functionalization. Furthermore, we expect that each isoform displays different kinetic properties since positive selection has probably occurred during the period of time following the duplication event (on the basal branch of the NMDH-II clade; Fig. 2). In this way, two amino acid sites (Lys292 and Asn414 in sorghum NMDH-I) were detected which could induce significant differences in kinetic properties between NMDH-I and NMDH-II. Different electric properties of the amino acids at these variable sites should result in conformational changes between NMDH-I and NMDH-II. To assess the importance of such sites, additional studies are needed. For example, experiments could be done using site-directed mutagenesis (Schepens et al., 2000) or by construction of reciprocal enzyme chimeras (Bläsing et al., 2000), offering new perspectives for biochemical research on the NADP-MDH enzymes (Miginiac-Maslow et al., 2000).

ACKNOWLEDGEMENTS

We would like to thank Trevor Hodkinson, Gerald Edwards and Mike Fay for their constructive suggestions. We are also grateful for helpful discussions with J. P. Jacquot, N. Corradi, C. Parisod and B. Offmann.

LITERATURE CITED


Duvall MR, Saar DE, Grayburn WS, Holbrook GP. 2003. Complex transitions between C₃ and C₄ photosynthesis during the evolution of Paniceae: a phylogenetic case study emphasizing the position of


