DNA Research **19**, 231–244, (2012) doi:10.1093/dnares/dss007
Advance Access publication on 24 February 2012

Alternative Splicing Regulates Targeting of Malate Dehydrogenase in *Yarrowia lipolytica*

Philomène Kabran[1,2], Tristan Rossignol[1], Claude Gaillardin[3], Jean-Marc Nicaud[4], and Cécile Neuvéglise[1,*]

INRA, UMR 1319 Micalis, F-78352 Jouy-en-Josas, France[1]; Division of Biological Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA[2]; AgroParisTech, Micalis, F-78352 Jouy-en-Josas, France[3] and CNRS, Micalis, F-78352 Jouy-en-Josas, France[4]

*To whom correspondence should be addressed. INRA, UMR 1319 Micalis, Biologie intégrative du métabolisme lipidique microbien, Bât. CBAI, F-78850 Thiverval-Grignon, France. Tel. +33 1-30-81-54-78. Fax. +33 1-30-81-54-57. Email: cecile.neuveglise@grignon.inra.fr

Edited by Takashi Ito
(Received 10 November 2011; accepted 30 January 2012)

Abstract

Alternative pre-mRNA splicing is a major mechanism contributing to the proteome complexity of most eukaryotes, especially mammals. In less complex organisms, such as yeasts, the numbers of genes that contain introns are low and cases of alternative splicing (AS) with functional implications are rare. We report the first case of AS with functional consequences in the yeast *Yarrowia lipolytica*. The splicing pattern was found to govern the cellular localization of malate dehydrogenase, an enzyme of the central carbon metabolism. This ubiquitous enzyme is involved in the tricarboxylic acid cycle in mitochondria and in the glyoxylate cycle, which takes place in peroxisomes and the cytosol. In *Saccharomyces cerevisiae*, three genes encode three compartment-specific enzymes. In contrast, only two genes exist in *Y. lipolytica*. One gene (*YlMDH1*, YALI0D16753g) encodes a predicted mitochondrial protein, whereas the second gene (*YlMDH2*, YALI0E14190g) generates the cytosolic and peroxisomal forms through the alternative use of two 3′-splice sites in the second intron. Both splicing variants were detected in cDNA libraries obtained from cells grown under different conditions. Mutants expressing the individual YlMdh2p isoforms tagged with fluorescent proteins confirmed that they localized to either the cytosolic or the peroxisomal compartment.

Key words: yeast; TCA cycle; glyoxylate cycle; MDH2; intron

1. Introduction

The hemiascomycetous yeast *Yarrowia lipolytica* has relatively few introns compared with other Opisthokonts such as vertebrates or even compared with basidiomycota and filamentous ascomycota. However, with introns in ~15% of its genes,[1] *Y. lipolytica* has many more introns than any other hemiascomycete yeast, e.g. four times more than *Saccharomyces cerevisiae*.[2] Despite this paucity of introns, the *S. cerevisiae* transcriptome is proving to be more complex than previously appreciated.[3,4] Alternative transcription start sites have been reported and, for some intron-containing genes, transcript variants result from either alternative splicing (AS) or intron retention. Overall, alternative transcripts leading to different proteins remain uncommon in *S. cerevisiae* and only one case of AS (*SRC1*[5]) and one of intron retention (*PTC7*[6]) have been described, along with several cases of alternative start sites (e.g. *YCAT*[7] and *SUC2*[8]). The large majority of alternative transcripts in yeasts are presumed to be untranslated, as most of them contain premature termination codons and therefore activate the nonsense-mediated mRNA

decay pathway that degrades defective mRNAs prior to translation (for review, see Behm-Ansmant *et al.*[9] and Stalder and Muhlemann[10]).

AS has been investigated at the genome scale in *Y. lipolytica* and all known modes of AS have been observed, i.e. exon skipping and alternative 5′- or 3′-splice sites, as well as intron retention and alternative start sites.[1] Several of these genes are involved in central carbon metabolism. The malate dehydrogenase (MDH) gene (*YlMDH2*, YALI0E14190g) was particularly attractive, as AS does not generate a premature termination codon, but leads to putative functional variants of the protein.

The MDH isoenzymes catalyse the conversion of malate into oxaloacetate with a concomitant reduction of NAD^+ (see Minarik *et al.*[11] for a review). This reaction, which is reversible, represents an important step in the tricarboxylic acid (TCA) cycle, a central metabolic pathway occurring in mitochondria and critical for cellular respiration and ATP production.[12,13] The reaction also takes place in the glyoxylate cycle, which is a variant of the TCA cycle that shares three of its five enzymes with the TCA cycle, including MDH. In plants, nematodes and yeasts, the glyoxylate cycle is partly localized in peroxisomes, and thus contributes to the degradation of free fatty acids by providing coenzyme A and/or reoxidizing NADH for the β-oxidation cycle. MDH is also involved in gluconeogenesis, which takes place in the cytosol and plays a significant role in the malate/aspartate shuttle across the mitochondrial membrane.

In *S. cerevisiae*, three MDH genes (*MDH1*, *MDH2* and *MDH3*) have been identified and each encodes an enzyme targeted for a different subcellular compartment. Mdh1p is localized in mitochondria,[14,15] Mdh2p in the cytosol[16,17] and Mdh3p in peroxisomes.[18,19] Large sequence variations have been observed in the MDH regions that encode the putative domains involved in compartmental targeting. For example, *MDH1* encodes a 17-amino acid N-terminal extension that is absent from the other isozymes and is removed upon mitochondrial import.[15] Mdh3p has a unique C-terminal tripeptide sequence, Ser-Lys-Leu, characteristic of peroxisomal targeting sequences (PTSs).[20] This targeting sequence, called PTS1, is conserved in *S. cerevisiae*, but other conservative variants, such as Ala-Lys-Ile, exist among yeasts.[21–23]

Homologues of all three *S. cerevisiae* MDH genes are present in many other hemiascomycetous species.[24] In contrast, the oleagineous yeast *Y. lipolytica* contains only two MDH genes: YALI0D16753g (*YlMDH1*) and YALI0E14190g (*YlMDH2*). Whereas the first gene encodes a protein predicted to have a mitochondrial location, the second gene encodes a putatively cytosolic form. The presence of only two MDH genes in *Y. lipolytica* compared with the three genes possessed by other hemiascomycetes, and the fact that *YlMDH2* is potentially subject to AS,[1] suggest a possible role for AS in the regulation of MDH compartmentalization in this yeast. We thus focused on these *Y. lipolytica* MDH genes to decipher the roles and regulation of the different isoenzymes. In this study, we show that *YlMDH2* encodes the cytosolic and peroxisomal forms of MDH. This dual localization is due to the presence of an alternative 3′-splice site in the second intron of the gene, located at the 3′-end of the coding sequence. The use of this alternative splice site creates an mRNA that encodes a carboxyl-terminal peroxisomal targeting sequence (PTS1) and allows specific peroxisomal localization, which was revealed by colocalization studies.

2. Materials and methods

2.1. Strains and culture conditions

The yeast and bacterial strains used in this study are listed in Table 1. The bacterial strains, Mach1 T1® (Invitrogen, Cergy Pontoise, France) and DH5α (Gibco BRL, Rockville, MD, USA) used for the amplification of recombinant plasmids, were grown at 37°C in Luria-Bertani medium supplemented with 100 μg/l ampicillin or 40 μg/l kanamycin, if required.

Growth media and conditions used for *Y. lipolytica* have been previously described.[25] The yeasts were grown on rich medium YPD and minimum medium YNB[26] or YNBE (YNB supplemented with 0.1% yeast extract), supplemented with 0.2% (w/v) casa-amino acids or 0.05 g/ml uracil, if required. Carbon sources were dextrose (D), oleic acid (OA), tributyrin (TB), triolein (TO), alkane (hexadecane), acetate (ammonium or sodium), glycerol or ethanol at a final concentration of 2%. The hydrophobic substrates were emulsified by sonication in a mixture containing 20% OA, TB or TO and 0.625% (v/v) Tween 40. For solid media, 2% agar was added. Yeast strains were grown at 28°C. Growth on 96-well plates in 100 μl was performed under constant agitation with glucose, malate, succinate or citrate at 0.5% as the carbon source. Growth was monitored by measuring the optical density at 600 nm every 10 min using a microtiter plate reader (Biotek, Colmar, France). For each strain and condition, four biological replicates were performed. Calculations of average OD, blank reduction, the lag phase (μ) and the maximum OD were performed using the grofit package in R.[27] Differences between growth curves were evaluated by analysis of variance with R statistical software.[28]

2.2. Molecular biology techniques

Standard molecular biology techniques were used throughout this study.[29] Restriction enzymes were

Table 1. *Yarrowia lipolytica* and bacterial strains used in this study

Strain	Genotype	Reference	Comments
Yeasts			
PO1d	Mat A, *leu2−270, ura3−302, xpr2−322*	25	
JMY1699	PO1d, *URA3-YALI0E14190g Δintron 1335-1422*	This study	Cytosolic form of *YlMDH2*
JMY1711	PO1d, *URA3-YALI0E14190g Δintron 1335-1426*	This study	Peroxisomal form of *YlMDH2*
JMY1685	PO1d, *URA3-YALI0E14190g*	This study	Wild-type form of *YlMDH2*
JMY2416	JMY1699, *LEU2*	This study	Cytosolic form of *YlMDH2*
JMY2426	JMY1711, *LEU2*	This study	Peroxisomal form of *YlMDH2*
JMY2428	JMY1685, *LEU2*	This study	Wild-type form of *YlMDH2*
JMY2451	PO1d, *URA3-pPOX-eYFP-YALI0E14190g Δintron 1335-1422*	This study	Cytosolic form of *YlMDH2*
JMY2457	PO1d, *URA3-pPOX-eYFP-YALI0E14190g Δintron 1335-1426*	This study	Peroxisomal form of *YlMDH2*
JMY2459	PO1d, *URA3-pPOX-eYFP-YALI0E14190g*	This study	Wild-type form of *YlMDH2*
JMY2499	JMY2451, *LEU2-pTEF-REDSTAR2*	This study	Cytosolic form of *YlMDH2*
JMY2501	JMY2457, *LEU2-pTEF-REDSTAR2*	This study	Peroxisomal form of *YlMDH2*
JMY2500	JMY2459, *LEU2-pTEF-REDSTAR2*	This study	Wild-type form of *YlMDH2*
JMY2493	JMY2451, *LEU2-pTEF-REDSTAR2-SKL*	This study	Cytosolic form of *YlMDH2*
JMY2496	JMY2457, *LEU2-pTEF-REDSTAR2-SKL*	This study	Peroxisomal form of *YlMDH2*
JMY2495	JMY2459, *LEU2-pTEF-REDSTAR2-SKL*	This study	Wild-type form of *YlMDH2*
Bacteria	**Plasmid/genotype**		
JME802	JMP62 *LEU2ex*, pPOX2 expression vector with the excisable *LEU2ex* marker	66	
JME803	JMP62 *URA3ex*, pPOX2 expression vector with the excisable *URA3ex* marker	56	
JME1018	JMP62 *URA3* pPOX eYFP N-Ter	This study	
JME1392	JMP62 *LEU2ex* pTEF Redstar2-SKL	This study	
JME1394	JMP62 *LEU2ex* pTEF Redstar2	This study	

obtained from Ozyme (Saint-Quentin-en-Yvelines, France). Genomic DNA from yeast was prepared as previously reported.[30] Polymerase chain reaction (PCR) amplifications were performed using an Eppendorf 2720 thermocycler, Pyrobest DNA polymerase (Lonza, Levallois-Perret, France) and the primers are listed in Table 2. PCR fragments were purified using the QIAquick PCR Purification Kit (Qiagen, Courtaboeuf, France). DNA fragments were recovered from agarose gels using the QIAquick Gel Extraction Kit (Qiagen).

2.3. Genetic modifications

Three constructions were made to determine the phenotypes of the strains containing deletions of either the short or long introns of *YlMDH2*. The primers used for PCR amplification are listed in Table 2. The forward primer contained a *Cla*I restriction site and stop codons in the three phases. The reverse primers contained an *Avr*II site and sequences corresponding to the wild-type gene and to the two variants. The corresponding *Cla*I-*Avr*II PCR fragments were cloned into the JMP62 *URA3ex* vector[31] previously digested with *Cla*I-*Avr*II. The resulting plasmids were purified, digested with *Bam*HI (a *Bam*HI site is present in the middle of the gene), and used to transform the PO1d strain by the lithium acetate method.[32] After single crossing-over, the mutated version had integrated at the *YlMDH2* locus under its own promoter and the genomic version was invalidated by the introduction of stop codons at the 5′-end of the wild-type gene (Supplementary Fig. S1). The resulting strains JMY1699, JMY1707 and JMY1711 (Table 1) expressed the predicted cytosolic, peroxisomal and wild-type versions of the *YlMDH2* gene products and were confirmed by southern blot (data not shown). Prototrophic strains were obtained by transforming the different mutants with the *LEU2* marker, which gave rise to the strains JMY2416, JMY2426 and JMY2428, respectively (Table 1).

We monitored localization of the MDH proteins in different subcellular compartments by expressing recombinant genes that encoded the peroxisomal or cytosolic mutant or the wild-type protein fused to the eYFP fluorescent protein tag at their N-termini. The three variant constructs were PCR amplified as *Bcl*I-*Avr*II fragments using specific primers (Table 2) and then cloned into the JMP62 *URA3ex*-YFP-N vector previously digested by *Bam*HI and *Avr*II. This vector was derived from the JMP62 *URA3ex*

Table 2. Oligonucleotides used in this study

Primer	Sequence	Purpose
E14190F1	GCCTACATCTACCTTGAC	cDNA sequencing
MDH1-Start	CAGATCGAT**TAA**C**TGA**C**TAG**GTTAAAGCTGTCGTTGCCGGAGC	MDH disruption cassette
MDH1-Sa	AGGCCTAGGTTTAGATCCTAGCTAGAATGGTTAGTGATCGTGTAGTTCAAATGG	MDH disruption cassette
MDH1-04	AGGCCTAGGTTTAGATCCTAGTTGGCAGGAGGAGGGTTAACAATG.	MDH disruption cassette
MDH1-09	AGGCCTAGGTTTAGATCTTGGCAGGAGGAGGGTTAACAATG.	MDH disruption cassette
MDHS1	AACAGGCAACAATGGCATGC	Probe for Southern blot
MDHS2	AACTGGATTCGCTTGACGAG	Probe for Southern blot
MDHgfp1	CGCTGATCA**A**gtgagttatcatggtgggag	Fusion cassette
MDHgfp2	CGCTGATCA**ATG**GTTAAAGCTGTCGTTGCC	Fusion cassette
RedSKL1	CACGGATCCCACAATGAGTGCTTCTTCTGAAGATG	Localization cassette
RedSKL2	TGGTGCCTAGGCTGC***TTAAAGCTTGGA***CAAGAACAAGTGGTGTCTACC	Localization cassette
RedNoSKL2	TGGTGCCTAGGCTGC**TTA**CAAGAACAAGTGGTGTCTACC	Localization cassette

Restriction sites are underlined: *Cla*I ATCGAT, *Avr*II CCTAGG, *Bcl*I TGATCA, *Bam*HI GGATCC. Stop codons and methionine codons (or the first nucleotide of the codon) are in bold. Nucleotides complementary to the S-K-L codons are in bold italic. Intron nucleotide sequence is in lower case.

vector[31] and allows N-terminal, eYFP-tagged proteins to be produced under control of the *POX2* promoter. The three different expression cassettes were excised from their respective vectors by *Not*I digestion, gel purified, and subsequently used to transform PO1d to create the JMY2451, JMY2457 and JMY2459 strains, respectively (Table 1). These strains were subsequently transformed with constructions containing the sequence of the fluorescent protein RedStar2. The RedStar2 gene was PCR amplified from the plasmid pYM43[33] using primers containing *Bam*HI and *Avr*II restriction sites (Table 2) and cloned into the JMP62 *LEU2ex* pTEF vector previously digested by *Bam*HI and *Avr*II. The JMP62 *LEU2ex* pTEF plasmid corresponds to JMP62 *LEU2ex* in which the pPOX2 promoter has been exchanged with the pTEF promoter.[34] Primers were designed to add or not add a PTS1 consensus sequence (Ser-Lys-Leu; SKL) onto the C-terminus of RedStar2 and allow expression of cytosolic or peroxisome-targeted RedStar2 protein. Each RedStar2 expression cassette was released by *Not*I digestion, gel purified, and subsequently used to transform the JMY2451, JMY2457 and JMY2459 strains. The resulting strains harboured one of the three different eYFP-tagged MDH protein expression constructs controlled by the *POX2* promoter and an addition construct for expressing either the RedStar2 or the RedStar2-SKL protein under control of pTEF. All the strains generated are listed in Table 1.

2.4. Microscopy

Images were acquired using a Zeiss Axio Imager.M2 microscope (Zeiss, Le Pecq, France) using a 100× objective and Zeiss filters 45 and 46 for fluorescent microscopy. Image acquisition was performed with the Axiovision 4.8 software (Zeiss). All images were post-processed similarly (background reduction) and merged using ImageJ software (http://rsbweb.nih.gov/ij/). Strains were grown 12 h on YNBE 2% OA, YNBE 2% D or YNBE with both 1% D and 2% OA.

2.5. RNA-seq analysis of *YlMDH2* transcripts

RNA-seq data produced in the lab were screened for the different splicing variants of *YlMDH2*. Total RNAs from cells cultivated in six different conditions (glucose, OA, TO, TB, alkane, and glycerol) were prepared using the Qiagen RNeasy kit (Qiagen). mRNAs were purified by the selection of poly(A)$^+$ transcripts, which were then sequenced by the Illumina Solexa sequencing technology with either a Genome Analyzer IIX or a HiSeq sequencing system. Thirteen to 29 million single-end reads were generated per sample with a read length of 36 nt, 50 nt or 100 nt (accession numbers E-MTAB-939 and E-MTAB-940). The reads were aligned to both forms of *YlMDH2* spliced transcripts (long or short intron) using SOAPaligner version 2.20.[35] Only reads aligned to the exon−exon junction (34 nt, 48 nt or 98 nt from each exon, depending on the read length) of intron 2 were counted and the ratio between the two isoforms was determined. We used Fisher's exact tests to determine whether the expression levels of *YlMDH2* differed in the different growth conditions and to determine the significance of the differences observed among the AS ratios.

2.6. Bioinformatics

Homologues of *S. cerevisiae* MDH were identified in fully sequenced *Y. lipolytica* genomes from Génolevures (http://www.genolevures.org/) using BLASTp.[36] Homologues were also found in *Pichia pastoris* (Pichia_GS115.pep_0509; https://bioinformatics.psb.ugent.be/gdb/pichia/) and in *Schizosaccharomyces pombe* (pompep_17022010; http://www.sanger.ac.uk/Projects/S_pombe/). Amino acids sequences were then aligned using MUSCLE[37] or Multalin[38] and columns of gap-containing residues were removed manually. Phylogenetic trees were constructed with the Neighbour-Joining algorithm using ClustalX.[39]

Predictions for mitochondrial protein targeting were performed using MITOPROT II version 1.101 (http://ihg.gsf.de/ihg/mitoprot.html).[40] PTS1 targeting sequence predictions were performed using the PTS1 predictor using default cut-off (http://mendel.imp.ac.at/mendeljsp/sat/pts1/PTS1predictor.jsp).[41]

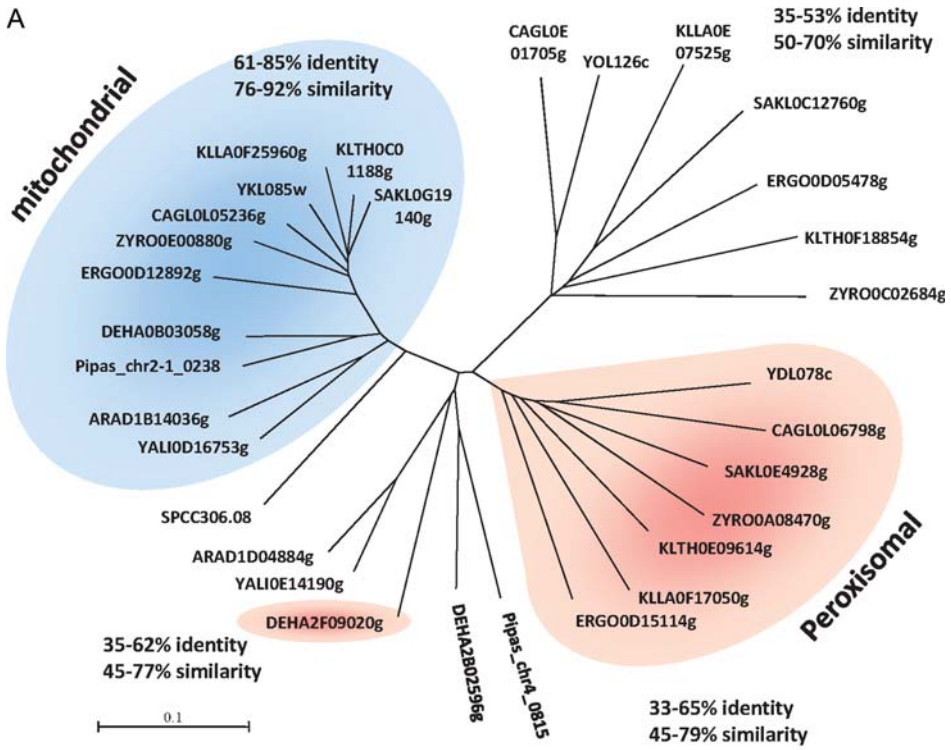
3. Results

3.1. Malate dehydrogenase protein families in hemiascomycetous yeasts

In *S. cerevisiae*, the three genes that encode MDH are: YKL085W (*MDH1*), which encodes the mitochondrial enzyme;[14,15] YOL126C (*MDH2*), which encodes the cytosolic enzyme;[16,17] and YDL378C (*MDH3*), which encodes the peroxisomal enzyme.[18,19] The three isozymes show amino acid identities ranging from 43 to 50%.[19] A Blast search for *S. cerevisiae* MDH homologues in the *Y. lipolytica* genome[42] identified only two genes, YALI0D16753g (*YlMDH1*) and YALI0E14190g (*YlMDH2*). Both genes exhibit the highest levels of amino acid conservation with the *S. cerevisiae MDH1* gene, with 63% identity and 79% similarity for YALI0D16753g and 47% identity and 62% similarity for *YlMDH2*. The computed proteome families of the Génolevures database[24] revealed that both *Y. lipolytica* MDH genes belong to the GL3R0092 family, which includes all of the MDH proteins encoded by the nine fully sequenced yeast genomes in the Génolevures database. All of these genomes contain three MDH genes. The only exceptions are *Y. lipolytica* (two genes) and *Lachancea (Kluyveromyces) thermotolerans*, which possesses three additional genes. We found these additional genes (KLTH0D00440g, KLTH0G19536g and KLTH0G19558g) to be closely related to each other (64–66% identity and 80–83% similarity) but very divergent from all other hemiascomycetous MDH genes. Due to their mutual similarity and their localization (one is subtelomeric and two are repeated in tandem), they likely represent species-specific gene duplications. Given their divergence from the other hemiascomycete proteins, we did not consider them in our phylogenetic analysis. As *Y. lipolytica* is relatively isolated among the hemiascomycetes from a phylogenetic point of view, we added two other yeasts, not included in the previous yeast set, more closely related to *Y. lipolytica*: *Arxula adeninivorans* (recently sequenced and annotated by the Génolevures consortium, to be published) and *P. pastoris*.[43] Like *Y. lipolytica*, these two species contain only two MDH genes. We used these data to construct a phylogenetic tree of MDHs in the hemiascomycetous yeasts with the unique MDH gene from *S. pombe* as the outgroup (Fig. 1A). Simultaneously, we performed *in silico* sequence analysis to predict mitochondrial and peroxisomal localization for the different genes using MITOPROT and PTS1 prediction online software (Fig. 1B). The phylogenetic tree highlights clear distinctions between the three types of MDH. The most conserved branch corresponds to the mitochondrial MDH, with high levels of amino acid conservation among the different proteins: 61–85% identity and 76–92% similarity (Fig. 1A). The peroxisomal branch is less conserved, with minimums of 34% identity and 44% similarity for the most distantly related proteins. The cytosolic MDHs segregated into two branches, with similarity and identity values close to those obtained for the peroxisomal branch (35–62% identity and 45–77% similarity). According to the localization predictions, only one gene, DEHA2F09020g from *Debaryomyces hansenii*, which has a PTS1 peroxisome targeting sequence (C-terminal sequence SKL and PTS1 predictor score of −8.044), is mislocalized in the cytosolic branch (Fig. 1). *Eremothecium* (*Ashbya*) *gossypii* has three MDH genes, but the localization prediction failed to identify any PTS1 sequence in these three genes. The ERGO0D15114g protein belonging to the peroxisomal branch harbours an unexpectedly long C-terminal extension of ∼150 amino acids. In a recently corrected version of the *Ashbya* Genome Database (http://agd.vital-it.ch/index.html[44]; version 11 of 26 May 2011), the gene AGOS_ADR252W encodes a protein of 339 amino acids (*vs* 486 aa in ERGO0D15114p) and has a putative PTS1 with a significant score (C-terminal sequence ARL and PTS1 predictor score 3.865).

The product of the *Y. lipolytica YlMDH1* gene is predicted to be targeted to the mitochondria (MITOPROT probability 0.9891), whereas the product of the *YlMDH2* gene is not (MITOPROT probability 0.3528). Neither of these genes harbours any PTSs. This is in line with their positions in the phylogenetic tree, i.e. *YlMDH1* is positioned within the mitochondrial branch, and *YlMDH2* is within the cytosolic branch of the tree. Similarly, *A. adeninivorans* and *P. pastoris* both have a mitochondrial form (ARAD1B14036g and Pipas_chr2-1_0238, respectively) and a cytosolic

A



B

Sequence	Mitoprot score	PTS1 Predictor Score	PTS1 Predictor C-ter sequence	Theorical localization
ARADD04884g	0.2470	-109.78	HTLKYLWWEEHR	
CAGL0E01705g	0.0053	-59.906	EKNIAKGLDFSA	
DEHA2B02596g	0.1094	-39.711	GNIDKGVAFIKK	
ERGO0D05478g	0.1637	-53.306	KNIETGETLGSN	
KLLA0E07525g*	**0.9578**	-45.822	LEGNIETGLKLV	Cytosolic
KLTH0F18854g	0.1081	-37.298	VDRGLEWAHAHR	
Pipas_chr4_0815	0.2762	-52.68	NIAKGTAFIAGN	
SAKL0C12760g	0.2209	-53.361	QGNIETGLNFVN	
YALI0E14190g-cyto	0.3528	-46.363	GEEFIVNPPPAN	
YOL126c	0.0388	-34.796	KGLEFVASRASS	
ZYRO0C02684g	0.0457	-25.858	KLEGGIRNGLSL	
ARAD1B14036g	**0.9922**	-44.77	IKKGQDFVAQNP	
CAGL0L05236g	**0.9927**	-39.675	IEKGINFVKSSN	
DEHA2B03058g	**0.9658**	-38.538	IQKGVDFVKQNP	
ERGO0D12892g	**0.9930**	-48.301	KQNIEKGINFAK	
KLLA0F25960g	**0.9850**	-46.178	NIEKGQKFVTGN	
KLTH0C01188g	**0.9592**	-25.565	KNIEKGTAFVKS	
Pipas_chr2-1_0238	**0.9482**	-37.73	IAKGQEFVKQNP	
SAKL0G19140g	**0.9867**	-44.23	NIEKGENFVNQN	Mitochondrial
SPCC306.0	**0.9829**	-44.69	KKSITKGEQFVA	
YALI0D16753g	**0.9891**	-38.829	IEKGVNFVKQNP	
YKL085w	**0.9879**	-48.496	NIEKGVNFVASK	
ZYRO0E00880g	**0.9931**	-47.948	KNIEKGVQFVNN	
AGOS_ADR252W**	0.2112	**3.865**	GVDFVHAPK**ARL**	
CAGL0L06798g	0.1482	**8.632**	KGKDFVSGQ**TKL**	
DEHA2F09020g	0.1316	**-8.044**	DQGTTFVTG**SKL**	
KLLA0F17050g	0,1922	**-0.713**	KKGLAFSKQ**TKL**	
KLTH0E09614g	0.2328	**6.376**	VQKGKRFVQ**SKL**	
SAKL0E04928g	0.2828	**-9.838**	KGKQFVIGG**SKL**	
YALI0E14190g-pero	0.3528	-1.439	EEFIVNPPP**AKI**	
YDL078c	0.1779	**1.269**	KGKSFILDS**SKL**	
ZYRO0A08470g	0.2043	**-5.434**	DKGKKFALG**AKL**	Peroxisomal

Figure 1. (A) Phylogenetic tree of the MDHs from 11 fully sequenced hemiascomycetous yeast species using the unique MDH gene from *S. pombe* (SPCC306.08) as an outgroup. The yeasts are *Candida glabrata* (CAGL), *D. hansenii* (DEHA), *Kluyveromyces lactis* (KLLA), *Lachancea thermotolerans* (KLTH), *Lachancea kluyveri* (SAKL), *Zygosaccharomyces rouxii* (ZYRO), *Eremothecium gossypii* (ERGO), *P. pastoris* (Pipas), *A. adeninivorans* (ARAD), *Y. lipolytica* (YALI) and *S. cerevisiae*. The blue zone indicates genes with a typical mitochondrial targeting sequence and the pink zone indicates genes with a potential PTS. The percentages of amino acid identity and similarity among the proteins of each group (peroxisomal, mitochondrial and cytosolic) are indicated next to each group. (B) *In silico* predictions of protein targeting. The MITOPROT score, PTS1 predictor score and C-terminal sequence are indicated for each MDH.

form (ARAD1D04884g and Pipas_chr4_0815, respectively) according to phylogeny and signal sequence prediction. Consequently, for these three closely related species, sequence analysis of their complete genomes did not identify any peroxisomal forms of MDH.

3.2. *The YlMDH2 gene is alternatively spliced*

The peroxisomal form of MDH has been reported to be mandatory for fatty acid degradation in *S. cerevisiae*.[45] *Yarrowia lipolytica* is an oleagineous yeast, with an affinity for fatty acids as a carbon source. Thus, the absence of a peroxisomal form of MDH was unlikely. A deeper analysis of the sequence of *YlMDH2* revealed a putative alternative 3′-splice site for the second intron, defined by comparison with the splicing sequence pattern of the genosplicing database (http://genome.jouy.inra.fr/genosplicing2). The gene model of *YlMDH2* is represented in Fig. 2A, with the positions of the putative intron boundaries (5′- and 3′-splice sites) and the branch point (BP) at the exon 2−intron 2−exon 3 junctions in Fig. 2B. The putative AS event leads to a very minor change in the length of the mRNA sequence (4 nt) but introduces a 3′-frameshift and thus generates different stop codons. If the intron is spliced at the upstream 3′-splice site (short intron), the amino acid sequence of the C-terminus encoded by the *YlMDH2* mRNA is P-A-N. If splicing occurs at the downstream 3′-splice site (long intron), the C-terminal amino acid sequence ends with A-K-I (Fig. 2B), which is a characteristic PTS1 sequence identified in several yeasts.[21,23]

We generated three cDNA libraries from *Y. lipolytica* cells cultivated in media containing either glucose (exponential and stationary phases) or oleate (exponential phase) as the carbon source[1] and obtained 30 cDNA clones corresponding to *YlMDH2* transcripts. All of these clones derived from cells harvested from glucose medium: 10 clones from cells harvested in the exponential phase (expo library) and 20 clones from cells harvested in the stationary phase (stat library). We sequenced the 3′-ends of 11 of these *YlMDH2* cDNA clones and found that seven corresponded to mRNAs spliced at the upstream 3′-splice site (short intron 2 splice) and four corresponded to mRNAs spliced at the downstream 3′-splice site (long intron 2). As expected from our model, the two groups of cDNAs differed by only four nucleotides (Fig. 2B) and thus validated AS for *YlMDH2*. The different numbers of cDNAs for the two spliced forms (five cDNAs with a short intron and two cDNAs with a long intron in the expo library, and only four cDNAs with a long intron, none with a short intron, in the expo library) was a first indication that the AS was regulated by the environment of the cells, i.e. the growth conditions. However, due to the low number of cDNAs sequenced, the statistical significance of this difference could not be confirmed. This prompted us to investigate larger data sets using RNA-seq experiments.

In a first experiment (Illumina Solexa with single reads of 36 bp; Table 3) involving six different growth conditions, we retrieved a total of 1829 RNA-seq reads containing the exon−exon junction of the second intron of *YlMDH2*: 1181 reads corresponded

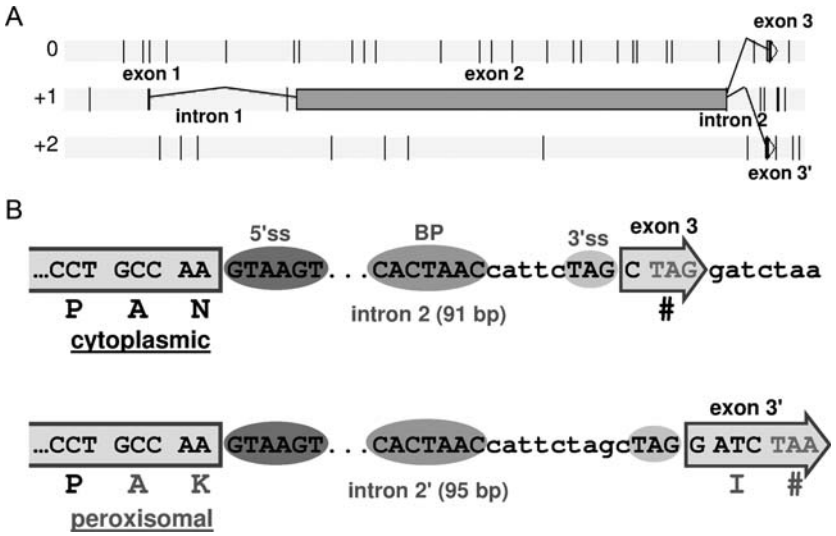


Figure 2. Gene model for *YlMDH2*. (A) Schematic representation of alternative transcripts from the multi-intronic MDH gene *YlMDH2*. Exons are represented by grey rectangles and introns are symbolized by thin black articulated lines. Vertical bars on each of the three phases (0, +1 and +2) represent in-frame stop codons. (B) Sequence representation of the 3′ regions centred on the second intron. Coloured circles indicate the 5′-splice site, the BP and the two 3′-splice sites used to generate the two mRNA variants. Exon parts are represented by grey rectangles and the two putative C-terminal protein sequences are indicated.

Table 3. Number of reads specific for the short or long intron 2 of *YlMDH2* obtained by RNA-seq analysis under various growth conditions

Carbon source	Illumina Solexa sequencing technology and read length	Total reads	Reads mapping short intron 2	Reads mapping long intron 2	Ratio short/long
Triolein	GAIIX single reads 36 nt	19 589 043	94 (57%)	71 (43%)	1.33
Tributyrin	GAIIX single reads 36 nt	15 996 762	298 (83%)	59 (17%)	5.05[a]
Glycerol	GAIIX single reads 36 nt	22 935 285	308 (62%)	186 (38%)	1.66
Alkane	GAIIX single reads 36 nt	23 903 839	99 (47%)	110 (53%)	0.9[b]
Glucose	GAIIX single reads 36 nt	13 022 908	121 (68%)	56 (32%)	2.16
Glucose	Hiseq single reads 50 nt	12 965 094	202 (67%)	86 (33%)	2.21
Glucose	Hiseq single reads 50 nt	10 592 928	203 (63%)	116 (37%)	1.70
Oleic acid	GAIIX single reads 36 nt	19 270 870	261 (61%)	166 (39%)	1.57
Oleic acid	Hiseq paired-end 100 nt	25 974 957	118 (60%)	78 (40%)	1.51
Oleic acid	Hiseq paired-end 100 nt	13 316 230	260 (67%)	130 (33%)	2.00

[a]The ratio on tributyrin is statistically different from that on all other media ($P < 0.0001$).[b]The ratio on alkane is statistically different from that on all other media, except on triolein ($P = 0.0765$).

to the short intron variant, and 648 reads corresponded to the long intron variant. The short intron/long intron usage ratio varied from 0.9 to 5.05 according to the carbon source used for growth (Table 3). These results strongly suggest that environmental conditions influence the regulation of *YlMDH2* AS. For example, for *Y. lipolytica* grown on TB medium, the number of transcripts containing the upstream 3′-splice site is five times the number of transcripts containing the downstream 3′-splice site. This ratio is significantly different from those found in all other conditions ($P < 0.0001$, for all comparisons to TB). In contrast, the levels of both variants were equivalent in cells grown on alkane. Additional RNA-seq experiments with two replicates on glucose (HiSeq with single reads of 50 bp; Table 3) and OA (HiSeq with paired-end reads of 2×100 bp) media confirmed that both ratios are conserved. However, these promising results on splicing regulation are presented as preliminary results since they came from a single time point, with a single concentration of carbon source and using RNA-seq as a single method of quantification. They obviously should be confirmed by additional experiments and alternative methods to understand the kinetic and the carbon source concentration-dependent regulations. This will be further investigated and will constitute the aim of a separate study.

3.3. *Absence of either the peroxisomal or the cytosolic form of the MDH does not affect growth rate, irrespective of the carbon source*

In order to evaluate the biological roles of the two MDH isoforms encoded by *YlMDH2*, we constructed strains that expressed cDNAs encoding either the cytosolic (cyto), peroxisomal (pero) or wild-type (wt) form of YlMdh2p at the *YlMDH2* locus in the *Y. lipolytica* PO1d strain (JMY2416, JMY2426 and JMY2428 strains, respectively). After complementation to restore prototrophy, strains were screened for growth on various substrates known to affect the growth of *S. cerevisiae* MDH mutants, i.e. glucose, acetate, ethanol and OA.[17–19,46,47] Surprisingly, we did not observe any growth rate differences among the three strains in either liquid or solid culture using these substrates as the sole carbon sources (see Fig. 3 for growth on solid substrates and Fig. 4A for glucose on liquid medium). We also screened these strains for growth on other lipid substrates, i.e. TB and TO, but failed to identify any specific growth phenotypes (Fig. 3). We also performed growth rate comparisons in a 96-well plate system using various carbon sources. Under these experimental conditions, we observed significant increases in the growth rate and maximum cell growth for the cyto strain (JMY2416) for all carbon sources tested, but observed no differences in the lag phase. For example, the *P*-values calculated for four replicates grown on glucose are $1.544e^{-2}$ and $1.184e^{-4}$ for the growth rate and maximum cell growth, respectively (Fig. 4B). We speculated that this result may be due to low oxygenation of the culture in a microtiter plate. In order to simulate low oxygen levels in a flask culture, we complemented the growth medium with antimycin A, a drug that impairs the respiratory chain, at a concentration that allows *Y. lipolytica* to grow but with a reduced rate (0.1 μg/ml). However, the growth rates of all strains were similarly reduced (data not shown), which indicated that the increased growth rate of the cyto strain is not directly linked to respiration efficiency. This phenomenon could be linked to oxygen access for other pathways or could reflect a different morphological state that

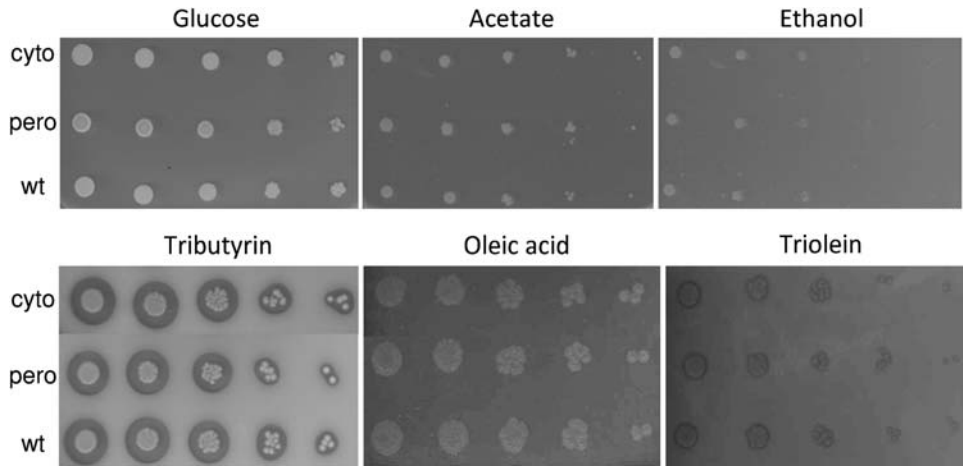


Figure 3. Expression of alternatively spliced forms of *YlMDH2* and growth on different substrates. Serial dilutions (serial dilution factor of five) of cultures of the wild-type (wt—JMY2428) strain, the cytosolic variant (cyto—JMY2416) and the peroxisomal variant (pero—JMY2426) were inoculated on YNB medium supplemented with different carbon sources. No growth differences between the mutants were detected; both of them were able to grow on all the media.

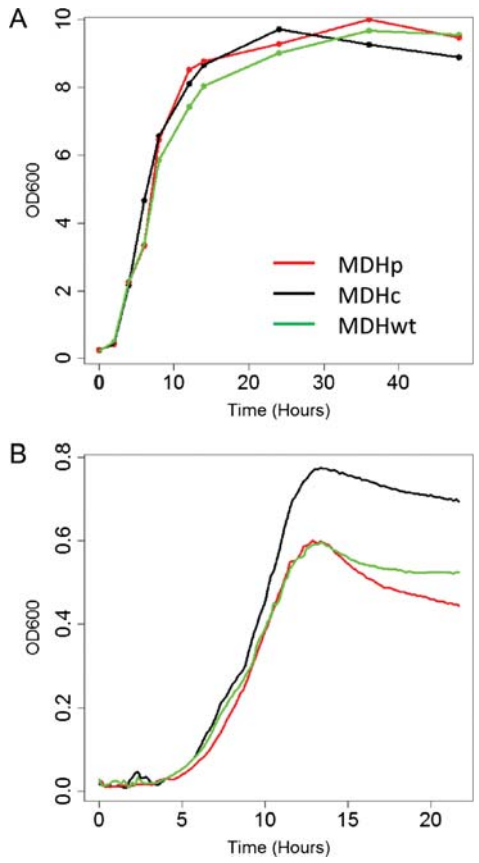


Figure 4. Comparative growth of *YlMDH2* mutants. (A) Growth curves in flasks with agitation in YNBE medium with glucose 2%. (B) Growth curves on 96-well plates in YNBE with glucose 0.5%. Coloured curves represent the different splicing mutants of *YlMDH2*: the cytosolic form is in black (MDHc), the peroxisomal form is in red (MDHp) and the wild-type is in green (MDHwt). OD, optical density measured at 600 nm.

affects optical density and may not rely on growth rate directly.

3.4. *Malate dehydrogenase localization*

In order to locate the two MDH isoforms produced by AS, we constructed vectors to express the eYFP fluorescent protein fused to the N-termini of the cytosolic and peroxisomal MDH isoforms and to the N-terminus of MDH expressed from the wt *YlMDH2* gene. *Yarrowia lipolytica* strains containing these constructs were subsequently transformed with a construction encoding the RedStar2 fluorescent protein without or with a PTS1 sequence (Ser-Lys-Leu; SKL) and used for colocalization studies. The native RedStar2 protein was cytosolic, whereas the RedStar2-SKL protein was targeted to the peroxisomes, which were visible as fluorescent dots in the cells (Fig. 5). In the strain expressing the predicted cytosolic form of eYFP-MDH, green fluorescence colocalized with red fluorescence produced by cytosolic RedStar2 (Fig. 5; JMY2499), but did not colocalize with the peroxisomal form of RedStar2 (Fig. 5; JMY2493). Conversely, the predicted peroxisomal form of eYFP-MDH colocalized with the peroxisomal form of RedStar2 (colocalization of fluorescent dots; Fig. 5; JMY2496) but not with cytosolic RedStar2 (Fig. 5; JMY2501). These colocalization experiments confirmed that the upstream splice variant mRNA encodes a cytosolic form of MDH, whereas the downstream splice variant encodes MDH with a C-terminal PTS1 and is peroxisomal. For all these constructions, we observed similar localizations and colocalizations with either glucose or OA as the carbon source. For the eYFP-tagged wild-type copy of YlMdh2p, which

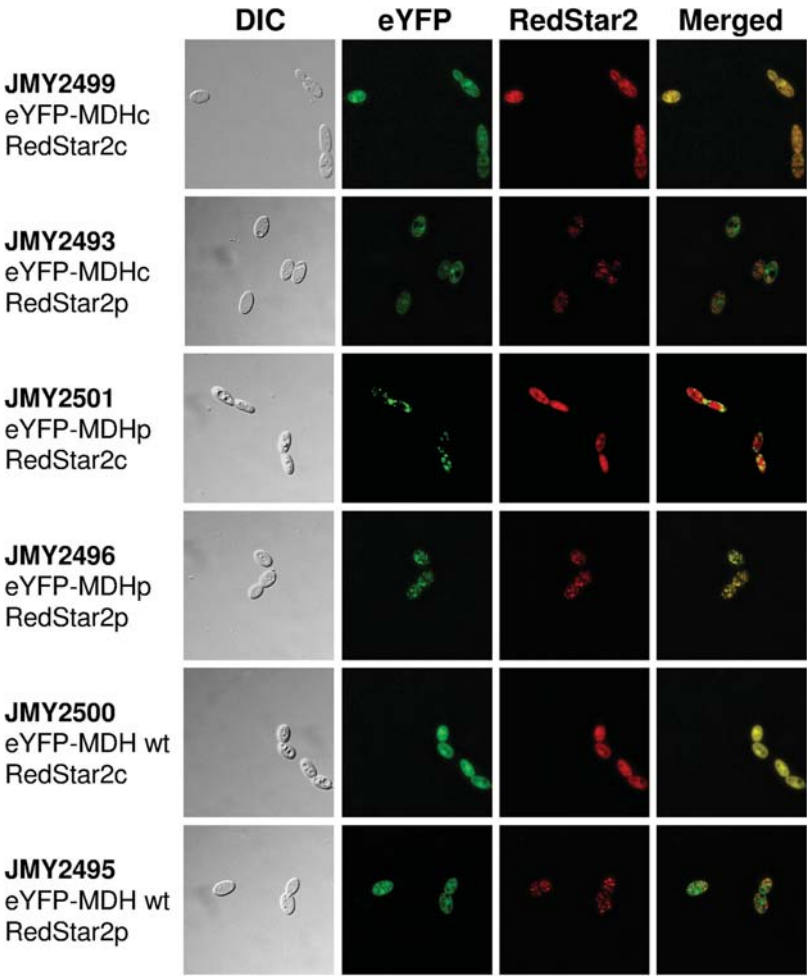


Figure 5. Colocalization of the two YlMdh2p isoforms with cytosolic or peroxisomal forms of the RedStar2 protein. eYFP-tagged peroxisomal YlMdh2p (eYFP-MDHp), cytosolic YlMdh2p (eYFP-MDHc) or wild-type YlMdh2p (eYFP-MDHwt) were co-expressed with either the peroxisomal (redstar2p) or cytosolic (redstar2c) forms of the RedStar2 protein. For each strain, both proteins (MDH and RedStar2) were visualized simultaneously by fluorescence microscopy. Cells were imaged after 12 h of growth in YNBE 2% OA using differential interference contrast (DIC) for eYFP fluorescence (eYFP, green) and for RedStar2 fluorescence (redstar2, red). eYFP and RedStar2 images were merged (right panel). Yellow colour indicates overlapping fluorescence and evidences colocalization.

encoded both MDH isoforms, we observed clear cytosolic fluorescence (Fig. 5; JMY2500) but no clear labelling of peroxisomes (Fig. 5; JMY2495). However, overlapping fluorescence might have prevented distinguishing the peroxisomal and cytosolic localizations, especially since the expression level of the cytosolic isoform is higher than that of the peroxisomal form, as suggested by the cyto/pero RNA ratio (Table 3).

3.5. AS of MDH is unique to Y. lipolytica

A comparative study of the MDH gene family revealed that *P. pastoris* and *A. adeninivorans* also contain only two genes coding for MDH, and in each case, a peroxisomal form is absent. Whereas the *P. pastoris* MDH genes do not contain any introns, model predictions for the gene that encodes the cytosolic form of MDH in *A. adeninivorans* revealed the presence of two introns. The second intron is located at exactly the same position as the second intron of *YlMDH2* (Supplementary Fig. S2). Unexpectedly, the terminal exon has undergone a large expansion, probably by fusion with a downstream CDS, which is conserved in synteny in *Y. lipolytica*. Despite the presence of the same positioned intron, no alternative 3′-splice site was detected in *A. adeninivorans*.

4. Discussion

Yarrowia lipolytica genes contain more introns than those of *S. cerevisiae* (15 and 4%, respectively) or other hemiascomycetes, and *Y. lipolytica* exhibits all known forms of AS.[1,2] This suggests possible roles for splicing in the post-transcriptional regulation of gene expression and/or in generating additional proteome complexity in *Y. lipolytica,* and that splicing might have a much higher impact in *Y. lipolytica* than

in *S. cerevisiae*. As an example of protein targeting regulation, we found clear evidence that the MDH gene *YlMDH2* is alternatively spliced and leads to the production of two proteins destined for distinct cellular compartments: cytosol and peroxisomes. The *HAC1* gene in *Y. lipolytica*[48] is also subject to functional AS; however, this gene possesses a non-spliceosomal intron and the splicing mechanism is mediated under stress conditions by Ire1p, a transmembrane kinase/endonuclease.[49] Thus, the *YlMDH2* gene constitutes the first case of functional AS involving a spliceosomal intron in *Y. lipolytica*.

Overall, few cases of AS have been described in hemiascomycete yeasts, which agrees with their low intron content. The only examples that contribute to the diversity of both the transcriptome and proteome were found in *S. cerevisiae*. In two reported cases, the protein isoforms encoded by the alternatively spliced transcripts are targeted to different subcellular compartments: *PTC7* proteins localize to the nuclear envelope and the mitochondria[6] and *YCAT* proteins to the mitochondria and peroxisomes.[7] In the case of *YlMDH2*, AS also leads to dual localization. We clearly demonstrated by cDNA sequencing that two transcript variants exist, which result from the use of two alternative 3′-splice sites separated by 4 nt. The predicted protein isoforms differ by only two C-terminal amino acids (AN or AKI), the latter of which constitutes a typical PTS1. Visualization of the eYFP-labelled cyto and pero isoforms showed that both proteins localized to their respective predicted compartments.

In *S. cerevisiae*, the distinct phenotypes of *mdh2*- and *mdh3*-deleted strains have been extensively studied. The *mdh2* mutant, which is equivalent to the *Y. lipolytica* pero strain (deleted for the cytosolic form), exhibits no growth defect on medium with glucose as a carbon source but is unable to grow on acetate or ethanol in minimum medium.[17–19,46,47] We did not observe a similar phenotype for the *Y. lipolytica* pero strain. In *S. cerevisiae*, a functional glyoxylate cycle is mandatory for growth on C2 carbon sources and the cytosolic oxaloacetate-malate conversion step by Mdh2p is likely to be required for the glyoxylate cycle.[16] Our finding that a *Y. lipolytica* strain deleted for cytosolic MDH is able to grow on ethanol and acetate suggests that the peroxisomal form can perform the malate–oxaloacetate conversion of the glyoxylate cycle. If we assume that the *Y. lipolytica* glyoxylate cycle follows the *S. cerevisiae* model, this result has two implications. First, the cytosolic and peroxisomal MDH proteins in *Y. lipolytica* must share the same enzymatic activities, which is entirely likely as the two isoforms differ by only two C-terminal amino acids. Moreover, in *S. cerevisiae*, cytosolic versions of Mdh3p (lacking the PTS1) or Mdh1p (lacking the mitochondrial targeting sequence) are able to complement an *mdh2* deletion for growth on ethanol, indicating that all isoforms share similar enzymatic activities despite their amino acid sequence differences.[18,50] Second, in cases of growth on C2 compounds, the peroxisomal form may be retained in the cytosol independent of the splicing event and the intermediates of the glyoxylate cycle, i.e. malate and oxaloacetate, may be able to cross the peroxisomal membrane in both directions. Examples of mislocalization of peroxisomal proteins have been reported in *S. cerevisiae:* Mls1p is sequestered in the cytosol when the yeast is cultivated on ethanol[51] and the Mdh3p–GFP fusion protein remained in the cytosol in an *mdh2Δ* mutant.[52] However, an alternative model to that of *S. cerevisiae* is that the entire glyoxylate cycle takes place in the peroxisome. Indeed, in *Y. lipolytica*, isocitrate lyase, one of the two enzymes specific to the glyoxylate cycle, is located in peroxisomes, whereas it is located into the cytosol in *S. cerevisiae*.[53] Thus, peroxisomal MDH may be involved in the malate–oxaloacetate conversion. However, the deletion of the peroxisomal form, as exemplified by our cyto strain, had no effect on growth on C2 compounds, suggesting that the *S. cerevisiae* model also reflects the situation in *Y. lipolytica*, at least in the absence of peroxisomal MDH.

van Roermund *et al.* reported that growth of the *S. cerevisiae mdh3* mutant (peroxisomal form disrupted) on OA was impaired. They showed that β-oxidation is also impaired in this mutant, which suggested an indirect role for Mdh3p in this process through reoxidation of the NADH generated during the β-oxidation of fatty acids.[45] This relationship has also been demonstrated in *Arabidopsis thaliana*.[54] In *Y. lipolytica*, strains expressing only the cytosolic MDH (cyto strain) grow as well as the wild-type strain on various lipid substrates and we did not observe any other distinct phenotypes. Thus, the peroxisomal MDH is not essential for NADH reoxidation in the peroxisome. This suggests that reoxidation is carried out by an alternative pathway in peroxisomes and might take the form of shuttles based on either lactate dehydrogenase (LDH) or glycerol-3-phosphate dehydrogenase (G3PDH), which are present in mammalian peroxisomes that do not contain MDH.[55] This alternative activity remains to be discovered in *Y. lipolytica*.

Several less convincing alternative hypotheses can also be addressed. First, cytosolic MDH may be targeted to the peroxisomes by a PTS distinct from PTS1 or may be imported through a peroxisomal protein complex. Second, the *Y. lipolytica* peroxisomal membrane may be permeable to NADH metabolites, which would allow NADH recycling through a cytosolic reoxidation. However, this is clearly not the case in *S. cerevisiae*.[45] Third, another β-oxidation pathway

allowing growth on fatty acids in the absence of peroxisomal MDH may exist in *Y. lipolytica*. In line with this concept, vestigial mitochondrial β-oxidation may exist in *Y. lipolytica*.[56,57]

From the sequencing of our cDNA libraries and the RNA-seq data, it appears that *YlMDH2* splicing is probably regulated or affected by the carbon source and/or the growth conditions. Several examples of splicing regulation in yeast, principally in *S. cerevisiae*, are known to exist. The unspliced form of SUS1 mRNA (intron retention) accumulates after a temperature shift.[58,59] Splicing of the *MER* genes in response to sporulation conditions and/or meiosis has also been well documented.[3,60–62] AS autoregulation by the protein product of the gene itself, through negative or positive feedback, has been reported for *RPL30*,[63] *YRA1*[64] and *SUS1*.[58] In *Y. lipolytica*, regulation of *YlMDH2* splicing could thus control the level of MDH in the peroxisome or the cytosol. Dual localization of a protein can be regulated at the level of transcription, usually involving alternative start sites, and at the post-transcriptional level. The differential targeting of YCAT is controlled at the transcriptional level by means of alternative transcription initiation sites, whose frequency of use depends on the carbon source available.[7] In the case of *PTC7* (intron retention), splicing is not regulated in response to carbon source availability, but the ratio of the two protein isoforms produced by the gene depends on whether the available carbon sources are fermentable or not, indicating that PTC7 isoform production is regulated post-transcriptionally.[6]

Thus, splicing regulation, as well as dual protein localization, appears to be more prevalent in yeasts than previously believed.[65] Whether this is reminiscent of ancestral attributes or corresponds to evolutionary changes that led to selective advantages remains to be investigated. Our data illustrate distinct differences in the metabolic behaviour of *Y. lipolytica* and *S. cerevisiae* in terms of their glyoxylate and TCA pathways and their MDH enzymatic activities. This underscores a possible role in the adaptation for growth on lipid substrates and for maintaining efficient β-oxidation across evolution. The AS mechanism that regulates MDH localization in *Y. lipolytica* remains so far an exclusive innovation of this yeast. *YlMDH2* AS regulation and its impact on protein localization and metabolism are under investigation and will provide new insights into the role of splicing on the proteome complexity in *Y. lipolytica*.

Acknowledgements: We thank François Brunel for RNA-seq library construction, Nicolas Morin for his help in RNA-seq data analysis and Stéphanie Michely for her assistance with R software utilization. We also thank all our colleagues from the Génolevures consortium for access to the genome sequence of *A. adeninivorans*, which is not yet publicly available.

Supplementary Data: Supplementary Data are available at www.dnaresearch.oxfordjournals.org.

Funding

This work was funded by the GDR CNRS 2354 Génolevures-3, the ANR Genarise (ANR-2005-BLAN-0331) and the ANR YeastIntrons (ANR-2010-BLAN-1620). P.K. was supported by a grant from the president of Republic of Côte d'Ivoire and by an INRA-MICA department fellowship.

References

1. Mekouar, M., Blanc-Lenfle, I., Ozanne, C., et al. 2010, Detection and analysis of alternative splicing in *Yarrowia lipolytica* reveal structural constraints facilitating nonsense-mediated decay of intron-retaining transcripts, *Genome Biol.*, **11**, R65.
2. Neuveglise, C., Marck, C. and Gaillardin, C. 2011, The intronome of budding yeasts, *C R Biol.*, **334**, 662–70.
3. Juneau, K., Palm, C., Miranda, M. and Davis, R.W. 2007, High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing, *Proc. Natl Acad. Sci. USA*, **104**, 1522–7.
4. Nagalakshmi, U., Wang, Z., Waern, K., et al. 2008, The transcriptional landscape of the yeast genome defined by RNA sequencing, *Science*, **320**, 1344–9.
5. Grund, S.E., Fischer, T., Cabal, G.G., Antunez, O., Pérez-Ortin, J.E. and Hurt, E. 2008, The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression, *J. Cell Biol.*, **182**, 897–910.
6. Juneau, K., Nislow, C. and Davis, R.W. 2009, Alternative splicing of PTC7 in *Saccharomyces cerevisiae* determines protein localization, *Genetics*, **183**, 185–94.
7. Elgersma, Y., van Roermund, C.W., Wanders, R.J. and Tabak, H.F. 1995, Peroxisomal and mitochondrial carnitine acetyltransferases of Saccharomyces cerevisiae are encoded by a single gene, *EMBO J.*, **14**, 3472–9.
8. Carlson, M. and Botstein, D. 1982, Two differentially regulated mRNAs with different 5′ ends encode secreted with intracellular forms of yeast invertase, *Cell*, **28**, 145–54.
9. Behm-Ansmant, I., Kashima, I., Rehwinkel, J., Sauliere, J., Wittkopp, N. and Izaurralde, E. 2007, mRNA quality control: an ancient machinery recognizes and degrades mRNAs with nonsense codons, *FEBS Lett.*, **581**, 2845–53.
10. Stalder, L. and Muhlemann, O. 2008, The meaning of nonsense, *Trends Cell Biol.*, **18**, 315–21.
11. Minarik, P., Tomaskova, N., Kollarova, M. and Antalik, M. 2002, Malate dehydrogenases—structure and function, *Gen. Physiol. Biophys.*, **21**, 257–65.

12. McCammon, M.T., Epstein, C.B., Przybyla-Zawislak, B., McAlister-Henn, L. and Butow, R.A. 2003, Global transcription analysis of Krebs tricarboxylic acid cycle mutants reveals an alternating pattern of gene expression and effects on hypoxic and oxidative genes, *Mol. Biol. Cell*, **14**, 958–72.
13. Przybyla-Zawislak, B., Gadde, D.M., Ducharme, K. and McCammon, M.T. 1999, Genetic and biochemical interactions involving tricarboxylic acid cycle (TCA) function using a collection of mutants defective in all TCA cycle genes, *Genetics*, **152**, 153–66.
14. McAlister-Henn, L. and Thompson, L.M. 1987, Isolation and expression of the gene encoding yeast mitochondrial malate dehydrogenase, *J. Bacteriol.*, **169**, 5157–66.
15. Thompson, L.M., Sutherland, P., Steffan, J.S. and McAlister-Henn, L. 1988, Gene sequence and primary structure of mitochondrial malate dehydrogenase from *Saccharomyces cerevisiae*, *Biochemistry*, **27**, 8393–400.
16. Kunze, M., Pracharoenwattana, I., Smith, S.M. and Hartig, A. 2006, A central role for the peroxisomal membrane in glyoxylate cycle function, *Biochim. Biophys. Acta (BBA)—Mol. Cell Res.*, **1763**, 1441–52.
17. Minard, K.I. and McAlister-Henn, L. 1991, Isolation, nucleotide sequence analysis, and disruption of the MDH2 gene from *Saccharomyces cerevisiae*: evidence for three isozymes of yeast malate dehydrogenase, *Mol. Cell. Biol.*, **11**, 370–80.
18. McAlister-Henn, L., Steffan, J.S., Minard, K.I. and Anderson, S.L. 1995, Expression and function of a mislocalized form of peroxisomal malate dehydrogenase (MDH3) in yeast, *J. Biol. Chem.*, **270**, 21220–5.
19. Steffan, J.S. and McAlister-Henn, L. 1992, Isolation and characterization of the yeast gene encoding the MDH3 isozyme of malate dehydrogenase, *J. Biol. Chem.*, **267**, 24708–15.
20. Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J. and Subramani, S. 1989, A conserved tripeptide sorts proteins to peroxisomes, *J. Cell Biol.*, **108**, 1657–64.
21. Aitchison, J.D., Murray, W.W. and Rachubinski, R.A. 1991, The carboxyl-terminal tripeptide Ala-Lys-Ile is essential for targeting *Candida tropicalis* trifunctional enzyme to yeast peroxisomes, *J. Biol. Chem.*, **266**, 23197–203.
22. Aitchison, J.D., Nuttley, W.M., Szilard, R.K., Brade, A.M., Glover, J.R. and Rachubinski, R.A. 1992, Peroxisome biogenesis in yeast, *Mol. Microbiol.*, **6**, 3455–60.
23. Aitchison, J.D., Szilard, R.K., Nuttley, W.M. and Rachubinski, R.A. 1992, Antibodies directed against a yeast carboxyl-terminal peroxisomal targeting signal specifically recognize peroxisomal proteins from various yeasts, *Yeast*, **8**, 721–34.
24. Sherman, D.J., Martin, T., Nikolski, M., Cayla, C., Souciet, J.-L. and Durrens, P. 2009, Génolevures: protein families and synteny among complete hemiascomycetous yeast proteomes and genomes, *Nucleic Acids Res.*, **37**, D550–4.
25. Barth, G. and Gaillardin, C. 1996, The dimorphic fungus *Yarrowia lipolytica*. In: Wolf, K. (ed.), *Genetics, Biochemistry and Molecular Biology of Non-Conventional Yeasts*. Springer Verlag: Heidelberg, pp. 313–88.
26. Mlickova, K., Roux, E., Athenstaedt, K., et al. 2004, Lipid accumulation, lipid body formation, and acyl coenzyme a oxidases of the yeast *Yarrowia lipolytica*, *Appl. Environ. Microbiol.*, **70**, 3918–24.
27. Kahm, M., Hasenbrink, G., Lichtenberg-Fraté, H., Ludwig, J. and Kschischo, M. 2010, grofit: fitting biological growth curves with R, *J. Stat. Softw.*, **33**, 1–21.
28. Ihaka, R. and Gentleman, R. 1996, R: a language for data analysis and graphics, *J. Comput. Graph. Stat.*, **5**, 299–314.
29. Sambrook, J., Maniatis, T. and Fritsch, E.F. 1989, *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor.
30. Querol, A., Barrio, E., Huerta, T. and Ramon, D. 1992, Molecular monitoring of wine fermentations conducted by active dry yeast strains, *Appl. Environ. Microbiol.*, **58**, 2948–53.
31. Nicaud, J.-M., Madzak, C., van den Broek, P., et al. 2002, Protein expression and secretion in the yeast Yarrowia lipolytica, *FEMS Yeast Res.*, **2**, 371–9.
32. Le Dall, M.T., Nicaud, J.M. and Gaillardin, C. 1994, Multiple-copy integration in the yeast *Yarrowia lipolytica*, *Curr. Genet.*, **26**, 38–44.
33. Janke, C., Magiera, M.M., Rathfelder, N., et al. 2004, A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, **21**, 947–62.
34. Müller, S., Sandal, T., Kamp-Hansen, P. and Dalbøge, H. 1998, Comparison of expression systems in the yeasts *Saccharomyces cerevisiae, Hansenula polymorpha, Klyveromyces lactis, Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Cloning of two novel promoters from *Yarrowia lipolytica*, *Yeast*, **14**, 1267–83.
35. Li, R., Li, Y., Kristiansen, K. and Wang, J. 2008, SOAP: short oligonucleotide alignment program, *Bioinformatics*, **24**, 713–4.
36. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990, Basic local alignment search tool, *J Mol Biol.*, **215**, 403–10.
37. Edgar, R.C. 2004, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.*, **32**, 1792–7.
38. Corpet, F. 1988, Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res.*, **16**, 10881–90.
39. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997, The CLUSTAL_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools, *Nucleic Acids Res.*, **25**, 4876–82.
40. Claros, M.G. and Vincens, P. 1996, Computational method to predict mitochondrially imported proteins and their targeting sequences, *Eur. J. Biochem.*, **241**, 779–86.
41. Neuberger, G., Maurer-Stroh, S., Eisenhaber, B., Hartig, A. and Eisenhaber, F. 2003, Prediction of peroxisomal targeting signal 1 containing proteins from amino acid sequence, *J. Mol. Biol.*, **328**, 581–92.
42. Dujon, B., Sherman, D., Fischer, G., et al. 2004, Genome evolution in yeasts, *Nature*, **430**, 35–44.

43. De Schutter, K., Lin, Y.-C., Tiels, P., et al. 2009, Genome sequence of the recombinant protein production host *Pichia pastoris*, *Nat. Biotech.*, **27**, 561–6.
44. Gattiker, A., Rischatsch, R., Demougin, P., et al. 2007, Ashbya Genome Database 3.0: a cross-species genome and transcriptome browser for yeast biologists, *BMC Genomics*, **8**, 9.
45. van Roermund, C.W., Elgersma, Y., Singh, N., Wanders, R.J. and Tabak, H.F. 1995, The membrane of peroxisomes in *Saccharomyces cerevisiae* is impermeable to NAD(H) and acetyl-CoA under *in vivo* conditions, *EMBO J.*, **14**, 3480–6.
46. Lee, Y.J., Jang, J.W., Kim, K.J. and Maeng, P.J. 2011, TCA cycle-independent acetate metabolism via the glyoxylate cycle in *Saccharomyces cerevisiae*, *Yeast*, **28**, 153–66.
47. McCammon, M.T. 1996, Mutants of *Saccharomyces cerevisiae* with defects in acetate metabolism: isolation and characterization of Acn- mutants, *Genetics*, **144**, 57–69.
48. Oh, M.H., Cheon, S.A., Kang, H.A. and Kim, J.Y. 2010, Functional characterization of the unconventional splicing of *Yarrowia lipolytica* HAC1 mRNA induced by unfolded protein response, *Yeast*, **27**, 443–52.
49. Babour, A., Kabani, M., Boisrame, A. and Beckerich, J.M. 2008, Characterization of Ire1 in the yeast Yarrowia lipolytica reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, *Curr. Genet.*, **53**, 337–46.
50. Small, W.C. and McAlister-Henn, L. 1997, Metabolic effects of altering redundant targeting signals for yeast mitochondrial malate dehydrogenase, *Arch. Biochem. Biophys.*, **344**, 53–60.
51. Kunze, M., Kragler, F., Binder, M., Hartig, A. and Gurvitz, A. 2002, Targeting of malate synthase 1 to the peroxisomes of *Saccharomyces cerevisiae* cells depends on growth on oleic acid medium, *Eur. J. Biochem.*, **269**, 915–22.
52. Wolinski, H., Petrovic, U., Mattiazzi, M., et al. 2009, Imaging-based live cell yeast screen identifies novel factors involved in peroxisome assembly, *J. Proteome Res*, **8**, 20–7.
53. Barth, G. and Scheuber, T. 1993, Cloning of the isocitrate lyase gene (ICL1) from Yarrowia lipolytica and characterization of the deduced protein, *Mol Gen Genet*, **241**, 422–30.
54. Pracharoenwattana, I., Cornah, J.E. and Smith, S.M. 2007, Arabidopsis peroxisomal malate dehydrogenase functions in beta-oxidation but not in the glyoxylate cycle, *Plant J.*, **50**, 381–90.
55. Visser, W.F., van Roermund, C.W., Ijlst, L., Waterham, H.R. and Wanders, R.J. 2007, Metabolite transport across the peroxisomal membrane, *Biochem J.*, **401**, 365–75.
56. Haddouche, R., Poirier, Y., Delessert, S., et al. 2011, Engineering polyhydroxyalkanoate content and monomer composition in the oleaginous yeast Yarrowia lipolytica by modifying the ß-oxidation multifunctional protein, *Appl. Microbiol. Biotechnol.*, **91**, 1327–40.
57. Vorapreeda, T., Thammarongtham, C., Cheevadhanarak, S. and Laoteng, K. 2011, Alternative routes of acetyl-CoA synthesis identified by comparative genomic analysis: Involvement in lipid production of oleaginous yeast and fungi, *Microbiology*, in press, doi:10.1099/mic.0.051946-0.
58. Cuenca-Bono, B., Garcia-Molinero, V., Pascual-Garcia, P., et al. 2011, SUS1 introns are required for efficient mRNA nuclear export in yeast, *Nucleic Acids Res.*, doi:10.10.93/nar/gkr496.
59. Hossain, M.A., Rodriguez, C.M. and Johnson, T.L. 2011, Key features of the two-intron *Saccharomyces cerevisiae* gene SUS1 contribute to its alternative splicing, *Nucleic Acids Res.*, **39**, 8612–27.
60. Davis, C.A., Grate, L., Spingola, M. and Ares, M. Jr. 2000, Test of intron predictions reveals novel splice sites, alternatively spliced mRNAs and new introns in meiotically regulated genes of yeast, *Nucleic Acids Res.*, **28**, 1700–6.
61. Engebrecht, J., Voelkel-Meiman, K. and Roeder, G.S. 1991, Meiosis-specific RNA splicing in yeast, *Cell*, **66**, 1257–68.
62. Nakagawa, T. and Ogawa, H. 1999, The *Saccharomyces cerevisiae* MER3 gene, encoding a novel helicase-like protein, is required for crossover control in meiosis, *EMBO J.*, **18**, 5714–23.
63. Li, B., Vilardell, J. and Warner, J.R. 1996, An RNA structure involved in feedback regulation of splicing and of translation is critical for biological fitness, *Proc. Natl Acad. Sci. USA*, **93**, 1596–600.
64. Preker, P.J. and Guthrie, C. 2006, Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA, *RNA*, **12**, 994–1006.
65. Ben-Menachem, R.B., Tal, M. and Pines, O. 2011, A third of the yeast mitochondrial proteome is dual localized: a question of evolution, *Proteomics*, **11**, 4468–76.
66. Haddouche, R., Delessert, S., Sabirova, J., Neuvéglise, C., Poirier, Y. and Nicaud, J.-M. 2010, Roles of multiple acyl-CoA oxidases in the routing of carbon flow towards β-oxidation and polyhydroxyalkanoate biosynthesis in *Yarrowia lipolytica*, *FEMS Yeast Res.*, **10**, 917–27.