Clustering of the *YNA1* gene encoding a $\text{Zn(II)}_2\text{Cys}_6$ transcriptional factor in the yeast *Hansenula polymorpha* with the nitrate assimilation genes *YNT1*, *YNI1* and *YNR1*, and its involvement in their transcriptional activation

Julio ÁVILA, Celedonio GONZÁLEZ, Nélida BRITO and José M. SIVERIO¹

Departamento de Bioquímica y Biología Molecular, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Canarias, Spain

The genes encoding the nitrate transporter (*YNT1*), nitrite reductase (*YNI1*) and nitrate reductase (*YNR1*) are clustered in the yeast *Hansenula polymorpha*. In addition, DNA sequencing of the region containing these genes demonstrated that a new open reading frame called *YNA1* (yeast nitrate assimilation) was located between *YNR1* and *YNI1*. The *YNA1* gene encodes a protein of 529 residues belonging to the family of $\text{Zn(II)}_2\text{Cys}_6$ fungal transcriptional factors, and has the highest similarity to the transcriptional factors encoded by *nirA*, and to a smaller extent to *nit-4*, involved in the nitrate induction of the gene

involved in the assimilation of this compound in filamentous fungi. Northern blot analysis showed the presence of the *YNA1* transcript in cells incubated in nitrate, nitrate plus ammonium, ammonium, and nitrogen-free media, with a decrease in its levels in those cells incubated in ammonium. In nitrate the strain $\Delta yna1::URA3$, with a disrupted *YNA1* gene, neither grew nor expressed the genes *YNT1*, *YNI1* and *YNR1*. In the gene cluster *YNT1-YNI1-YNA1-YNR1*, the four genes were transcribed independently in the *YNT1* → *YNR1* direction and the transcription start sites were determined by primer extension.

INTRODUCTION

Nitrate assimilation in those yeasts able to use nitrate follows the same pathway as that described for plants and filamentous fungi. Once nitrate enters the cells, it is reduced to ammonium by the consecutive actions of nitrate reductase and nitrite reductase [1]. In the yeast *Hansenula polymorpha* the genes encoding a nitrate transporter (*YNT1*), nitrite reductase (*YNI1*) and nitrate reductase (*YNR1*) are clustered [2–4] as in filamentous fungi [5–7] and *Chlamydomonas reinhardtii* [8]. In contrast, no linkage between the nitrate assimilation genes has been found in plants [9]. In general, as in *H. polymorpha*, the genes involved in nitrate assimilation are induced by nitrate and nitrite and repressed by sources of reduced nitrogen [9,10]. The mechanisms underlying nitrate induction and nitrogen catabolite repression, as well as the signalling of these events, are not well known. The current knowledge on these topics has been derived mainly from filamentous fungi, in which two transcriptional factors have been closely studied: *nit-4* and *nit-2* in *Neurospora crassa* and their counterparts *nirA* and *areA* in *Aspergillus nidulans*. *nit-4* and *nirA* encode proteins belonging to $\text{Zn(II)}_2\text{Cys}_6$ fungal transcriptional factors, involved specifically in the induction by nitrate of the nitrate assimilatory genes [11,12]. As general features, the fungal transcriptional factors of the $\text{Zn(II)}_2\text{Cys}_6$ type recognize in the promoter of the target genes the sequences CGGX_nCCG [13]. However, NIRA- and NIT4-binding sites are far from the consensus sequence CGGX_nCCG ; NIRA binds four sites with consensus sequence CTCCCGHGG in the intergenic region of the divergently transcribed genes *nirA* and *nirB* encoding nitrate and nitrite reductases. The NIT4-binding

site contains the symmetrical octameric sequence TCCGCGGA [14,15]. With regard to the structure of the $\text{Zn(II)}_2\text{Cys}_6$ proteins, X-ray studies of protein–DNA crystals of members of this family such as Gal4p, Ppr1p and Put3p revealed that these proteins bind to the DNA as homodimers [16–18]. The monomers interact through a coiled-coil element and bind DNA at the CGGX_nCCG sites by means of the $\text{Zn(II)}_2\text{Cys}_6$ domain. The homodimerization of NIRA has also been shown rigorously [19]. In addition *nit-4* and *nirA* in *N. crassa*, and *nit-2* and *areA* in *A. nidulans*, encode proteins containing a single zinc-finger DNA-binding motif characteristic of the GATA family of the *trans*-acting regulatory proteins [20,21] and are involved in the nitrogen catabolite derepression of the genes involved in the catabolism of nitrate, nitrite and purines. In *N. crassa* another gene, *nmr*, seems to be involved in repression by nitrogen catabolites [22].

The genes involved in the transcriptional regulation of nitrate assimilation in yeast, as in algae and plants, are unknown. The isolation and characterization of these genes in yeast and the creation of strains with disrupted genes will allow the cloning by functional complementation of genes with similar functions from heterologous sources. In this article we report the isolation and characterization of *YNA1* from *H. polymorpha*, which is involved in the transcriptional activation of the genes encoding the enzymes of the nitrate assimilation pathway.

MATERIALS AND METHODS

Yeast strains and growth conditions

H. polymorpha strain NCYC 495 was used to construct the genomic DNA library in lambda EMBL3. The NCYC495

Abbreviations used: ORF, open reading frame; *YNT1*, *YNI1* and *YNR1*, genes encoding nitrate transporter, nitrite reductase and nitrate reductase respectively.

¹ To whom correspondence should be addressed (e-mail jsiverio@ull.es).

The *YNA1* nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number AJ223294.

Figure 1 Expression of *YNT1*, *YNR1* and *YNI1* in the wild-type and $\Delta yna1::URA3$ strains

The amount of total RNA loaded in each lane was 10 μ g. The cells from which RNA was isolated were grown in ammonium, washed and incubated at 10 mg wet weight of cells/ml for 2 h in the nitrogen sources indicated, each at 5 mM. *YNT1*, *YNR1* and *YNI1* indicate the probe used; the bottom panel shows the RNA samples stained with ethidium bromide.

derivative strain *leu1-1,ura3* was used as the recipient strain to perform gene disruption; it is referred to in this paper as the wild type. Cells were grown at 37 °C with shaking in liquid medium containing 0.17% yeast nitrogen base without amino acids and ammonium sulphate (Difco), 2% (w/v) glucose and either 5 mM $NaNO_3$ or 5 mM NH_4Cl as nitrogen source, unless stated otherwise. The *leu1-1,ura3* strain was grown in the same medium, supplemented with 0.23 mM leucine and 0.19 mM uracil.

Disruption of the *YNA1* gene

The strategy followed to disrupt the *YNA1* gene is shown in Figure 1. The plasmid pNiA2 harbouring a *SalI*–*BamHI* fragment of approx. 2.9 kb containing part of the *YNA1* coding region was used. An internal 806 bp *EcoRV*–*NdeI* fragment from *YNA1* was replaced by a 1796 bp *BamHI*–*BglII* fragment containing the *H. polymorpha URA3* gene [23]. The resulting plasmid, pYNA1-UR2, when digested with *SalI*–*BamHI* produced a fragment of approx. 3.9 kb containing the *URA3* marker flanked by *YNA1* and *YNI1* regions. This fragment was separated by electrophoresis, isolated from the agar and used to electrotransform the *H. polymorpha leu1-1,ura3* strain [24]. Cells were subsequently plated on synthetic medium without uracil and with ammonium as nitrogen source. Prototroph cells for uracil were further screened by Southern blotting (Figure 1B). Of 16 prototrophs (URA^+) analysed, three contained disrupted *YNA1*. To restrict the possible growth sustained by leucine, the $\Delta yna1::URA3,leu$ strain was transformed to LEU^+ by the integration of a plasmid containing the *H. polymorpha LEU2* gene [25].

Nucleic acid isolation

Yeast DNA was isolated as described [26]; the cells were harvested in the early exponential phase of growth (2 mg wet weight/ml). Total yeast RNA was isolated as in [27] and fractionated by electrophoresis on formaldehyde agarose gel. Poly(A)⁺ RNA was prepared as described [28]. DNA lambda phage and plasmids were isolated as described [29]. λ JA13 containing the gene cluster was isolated from a *H. polymorpha* genomic library by PCR with degenerate oligonucleotide primers as described previously [2].

Southern and Northern blot analyses

Southern and Northern blot analyses were performed as described [29,30]. Nitrocellulose membranes (Schleicher and Schuell, Dassel, Germany) and positively charged nylon membranes (Boehringer Mannheim, Germany) were used in Southern and Northern blot analyses respectively. The probes were labelled with the digoxigenin system from Boehringer Mannheim. The detection method used in Southern blot analysis was the enhanced chemiluminescence system (Amersham, Madrid, Spain). In Northern blot analysis the detection was performed with the CDP-Star system (Boehringer Mannheim, Germany). The 743 bp *NcoI*–*BglII*, 931 bp *SacI*–*HindIII*, 533 bp *NdeI*–*BamHI* and 454 bp *XbaI* DNA fragments from *YNT1*, *YNI1*, *YNA1* and *YNR1* respectively were used as probes in Northern analysis.

DNA sequencing

Bluescript phagemid plasmids from Stratagene (Heidelberg, Germany) were used. Exonuclease III unidirectional deletions and single-strand DNA were prepared by following the manufacturer's instructions (Promega Corp., Madison, WI, U.S.A.). DNA was sequenced by the dideoxy chain termination method [31] with Sequenase (U.S.B., Cleveland, OH, U.S.A.). The sequence was performed on both strands. Multiple protein sequence alignments were performed with the CLUSTAL V program [32].

Transcription start site

The transcription start site was determined by primer extension in accordance with the procedure described in [33], with the following exceptions. The oligonucleotides were labelled with [γ -³²P]ATP by following the procedure described in [29]. Poly(A)⁺ mRNA (2 μ g) was used, and hybridization of oligonucleotides and RNA was conducted by incubation for 90 min at 65 °C followed by 6 h at 60 °C. The following oligonucleotides were used (the numbers in parentheses are their positions in the coding region): *YNT1* (+67, +91), 5'-AGA CGT TCC ATA GGT TAA AAA TTG G-3'; *YNI1* (+61, +85), 5'-ACC ATG CCA AGG CCC ACA ATC ACA A-3'; *YNA1* (+56, +90), 5'-CAT GAG GGA ACT TTG CCA TCG CAG T-3'; *YNR1* (+49, +83), 5'-AAA GGT AGC TCC GTG ATA TCT A-3'.

Other methods

Crude extracts and determination of nitrate reductase activity were performed as described [34].

RESULTS

Clustering of the genes involved in nitrate assimilation in the yeast *H. polymorpha*

DNA sequencing in the λ JA13 phage, isolated from a lambda EMBL3 *H. polymorpha* genomic DNA library by using a nitrate reductase homologous probe synthesized by PCR [2], revealed

Figure 2 Scheme of the cluster containing the genes involved in nitrate assimilation in *H. polymorpha*

(A) The cluster contains the nitrate transporter (*YNT1*, 1524 bp), nitrite reductase (*YNI1*, 3132 bp), a $\text{Zn(II)}_2\text{Cys}_6$ transcriptional factor (*YNA1*, 1587 bp) and nitrite reductase (*YNR1*, 2577 bp) genes. It comprises 9201 bp from the *YNT1* start codon to the *YNR1* stop codon. The regions between *YNT1* and *YNI1*, between *YNI1* and *YNA1*, and between *YNA1* and *YNR1* contain 220, 45 and 113 bp respectively. (B) The primer extension reactions corresponding to each gene, and the transcription start sites.

that *YNT1*, *YNR1* and *YNI1* were clustered [4]. Extensive DNA sequencing of the same phage insert showed the presence of a new open reading frame (ORF) called *YNA1* (yeast nitrate assimilation) located between *YNI1* and *YNR1* to define the cluster *YNT1*-*YNI1*-*YNA1*-*YNR1* (Figure 2A). The *YNA1* start codon is located 45 bp from the stop codon of *YNI1*, whereas the stop codon of *YNA1* is located 113 bp from the ATG of *YNR1*. The cluster consists of 9205 bp from the first codon of *YNT1* to the last codon of *YNR1*.

The genes of the cluster are transcribed independently; the transcriptional start site for each gene was mapped by primer extension (Figure 2B). Heterogeneity in the length of the 5' region of the transcripts was observed, indicating several transcriptional start sites, a feature already described for *Saccharomyces cerevisiae* [35,36]. The *YNT1* TATA box lies at -58 and the main initiation site is located at -21. *YNI1* presents four transcription start sites with one nucleotide difference in a region rich in T and A. For *YNA1* the strongest signal of the primer extension reaction was at -10 and -9; for *YNR1* three initiation sites were observed at -28, -21 and -20 situated in a CT-rich stretch with the main initiation site in the CAAG sequence.

***YNA1* encodes a putative $\text{Zn(II)}_2\text{Cys}_6$ binuclear cluster**

The *YNA1* coding region contains 1587 bp encoding a protein of 529 residues. Search for proteins similar to Yna1p through the gapped BLAST program [37] showed that the *YNA1* product possesses highest similarity with the well-known transcriptional factors characterized by the consensus sequence Cys-Xaa₂-Cys-Xaa₆-Cys-Xaa₆-Cys-Xaa₆-Cys-Xaa₂-Cys in the N-terminal region, such as those encoded by *nirA* [11], *CHA4* [38], *QAI1F* [39], CAT8 [40] and *nit-4* [12]. Alignment of Yna1p, NIT4 and NIRA

with the Clustal V program [32] (Figure 3) reveals the highest similarity at the N-terminus around the six-cysteine cluster region. In the Yna1 protein seven residues are found between the second and third Cys. The six-cysteine region is followed by a basic region Lys-Arg-Arg-Arg-Lys-Xaa-His-(Xaa)₈-Lys-Lys-Xaa-Arg, which is present in the proteins of this family. After this sequence there is a region of approx. two-thirds of the protein that has low similarity with NIT4 and NIRA. The third part of the protein matches well with the two filamentous fungi proteins. The proline- and glutamine-rich regions present in NIRA and NIT4 respectively are not found in the *YNA1* product.

Expression of *YNA1*

When *YNA1* expression in cells incubated in different nitrogen sources was examined by Northern blot analysis with poly(A)⁺ RNA, a transcript of 1.7 kb, which is in good agreement with the *YNA1* gene length, was observed in all the conditions studied. However, the level of the *YNA1* transcript was much lower in the cells incubated in ammonium, even in the presence of nitrate, than in those cells incubated in nitrogen-free or nitrate-containing media (Figure 4). A second transcript of approx. 500 bp, regulated in the same way as the 1.7 kb message, was observed in the Northern blot but its nature remains unknown.

YNA1 is expressed at a very low level in comparison with the structural genes for nitrate assimilation. In Northern blot analysis the *YNA1* transcript was undetected in total RNA samples and a very low, almost undetectable β -galactosidase activity was observed in cells containing the *YNA1* gene promoter fused to the *lacZ* gene (results not shown).

The *YNA1* gene promoter was replaced by the *ADHI* gene promoter from *S. cerevisiae* to test the possibility that high

<i>nirA</i>	-----MGEKLDPELSSDGPHTKSSSKGQGTSTDN--APAS-KRRCVSTACIACR--R	47
<i>nit-4</i>	MNSDDVQMSSQDAPGASAGLAPDNIASSLPSKKSRGADPTNQKRCVSTACIACR--R	58
<i>YNA1</i>	-----MNLQTRICITCKLLR	16
<i>nirA</i>	RKSKCDGNLPSCAACSSVYHTTCVYDPNSDHRKGVYKDDTLTKNSTLLTLIALLN	107
<i>nit-4</i>	RKSKCDGALPSCAACASVYGTETCIYDPNSDHRKGVYREKNSDKAQNATLQILIEATLN	118
<i>YNA1</i>	RK-DCDGVFSCNLCIKRK-RECIYQADSKRRRYHTDITQYLEKKVRLQDFLSKNDP	74
<i>nirA</i>	YEEEDAFDLVRQIRSCDNLEDVQASLVNQEKSSGWSLNAVIHE-----ENDIAQTQDF	161
<i>nit-4</i>	ASEEDVIDIRRICTDDLDEVAESIRREKATATNDNDSDDEPTQGRDDATSSQAVEG	178
<i>YNA1</i>	QLQLGENGEQTQIQFLKRVDEAKLTLLNEMDDLRDN--S-----PEPDFQ	120
<i>nirA</i>	ESELAKGMSNLVLD-GSRKFIGGTSNLIPLFP--GSELNEFKPLGATNGLGVSRTWT	217
<i>nit-4</i>	ERDLARKMGELEIRIENGVSFRFIGGTSNLIPLFP--GSELNEFKPLGATNGLGVSRTWT	236
<i>YNA1</i>	YDDLGLVGVAGLN-----LYYIDP--IFKTELLSLFKKHC--SEITYAI	161
<i>nirA</i>	TVTDDQQLISHLITMYFSWHYFFFTLSKELFYRDSRG--VP--SQYCSLLVNTML	271
<i>nit-4</i>	EVTKNPQLIIHLVNMVFNWHYFFFTLSRSLFYRDFIKGNAAGPRSTVYCSLLVNTML	296
<i>YNA1</i>	TLT-----L-P--TLDTWVSTDTDP--HQLMLTVY	189
<i>nirA</i>	ALGCHFSWPGAREDPNSATAGDHFKEAKRLILNDELVNSKLTQVQALMSVREAG	331
<i>nit-4</i>	ALGCHFTSDGAFAPVPGDSRTKGDFHFAEAKRLIVQNDYEKPLRTVQALMSVREAG	356
<i>YNA1</i>	GVACLYSKNPASKISKILMNQAQIADMRRLFNLD--EY--LQGLLLSCYMG	241
<i>nirA</i>	CGREGKWWVYSGMSFRMAFDLGLNLESSSLRLDSEEEIDARRITFWGCFLFDKCSNYLG	391
<i>nit-4</i>	CGREKWWVYSGMSFRMAQDGLMLDIOGSLD--EKEVDARRITFWGCFVFDKCSNYLG	413
<i>YNA1</i>	MGNDSMSLYISMACSLTQHMLGHSYDEHS--SAKFAPRTTFFQSALLSITCTQDRIT	299
<i>nirA</i>	RQFOFTTANTSVSAVDILPNEESTLWSPYSDMGPSREYAQPSRTRAVADOIS--QLCKISG	451
<i>nit-4</i>	RLFPQLPENTYVNPVYDFPDEDAELWSPYSDMGPSREYAQPSRTRAVADOIS--QLCKISG	472
<i>YNA1</i>	TS-RIGVPSCHFKRIISPVYVQLSRNDPSYLEELCFYITRVWYIMDRFTQICISQ	358
<i>nirA</i>	DLVVFYDLAPKEPKSSKQLELKKLSEIHTREAWKGLPKLEPREGQLPQALLMHMFY	510
<i>nit-4</i>	DLVVFYDLAPKEPKSSKQLELKKLSEIHTREAWKGLPKLEPREGQLPQALLMHMFY	532
<i>YNA1</i>	DLGDTQERK-----LLKTA-ELAFQALKNSLMPFRDE--KHSKNYHIMFQNLN	406
<i>nirA</i>	QLLLHLRPLFLYKTSPLPQHVSPKRLCTQAAAISKLRLYKRTYGFQICNIAVY	570
<i>nit-4</i>	HLQYHLRPLFLYKTSPLPQHVSPKRLCTQAAAISKLRLYKRTYGFQICNIAVY	591
<i>YNA1</i>	EACVLLHLRFLIDPLCRN-----K--CVTSATEITKLRLLETGGIERASYHAF	456
<i>nirA</i>	IAHTALTIIHLNLPEKN--AQRDVIHGLRHEEMGESWLCARRTLRLDISAKWQVQL	627
<i>nit-4</i>	MLHSACTIHLNLPEKN--ARRDITHGVROLEMAEDWPCARRTLGLISVLARKWNVEL	648
<i>YNA1</i>	ISSLAAMIHLVLYSEKLRPEHLNLETCVTALKISQQRWKRGALEMLCEFSQSHGINV	516
<i>nirA</i>	PREAVIVFEQTHARWGSWG--PWOQAASPTTSDS--PPSVSSQSVVATTDLSQPVSO	681
<i>nit-4</i>	PEEAIVLKRTRDEKYGMEFSTSEVPSPNRTAPSLAPSPAPPYTPASLLFTTAARVLEQ	709
<i>YNA1</i>	PILETFSRWTKF-----	529
<i>nirA</i>	SAGN-----QPANFSMG-----TSPNLTQPVASQYSSTSPGSPVSAMRAV--Q	723
<i>nit-4</i>	QYFSFMTLYSHPTFPHLGPESSIDPGMSPHIVTSDLDPSAFIIPQQQNMQAISLSQ	768
<i>YNA1</i>	-----	
<i>nirA</i>	RSFSAQLAHNEARQPEP-----TYLR-----PVSTSYGPVPSQSAQEQWYSPT	765
<i>nit-4</i>	NNMLQHQQHHHLQHQHQPPQHNNHMTYQQQQHLLTHPVSSASSMSDTLATITAWGIP	828
<i>YNA1</i>	-----	
<i>nirA</i>	-----EAQFRAFTAHSMPITTSQAQSP--LTTFTDPENLV	798
<i>nit-4</i>	SSPGNNNNNNIVSQHPHHQKQPQQQQPQQAQRYPTVGSVGTNTVKPPAAATQTFTPAQLH	888
<i>YNA1</i>	-----	
<i>nirA</i>	E-----ESQDWSRDVNALQLGAEDWTQ-----	821
<i>nit-4</i>	ANNLATATRSASNHSVGRHVSFSSIIYADGQWYLKDGVTWQGFQGWLEGGAGTA	948
<i>YNA1</i>	-----	
<i>nirA</i>	-----WNGLPTTADSWRYVDNVP--NIPSTSPADYKPPQPPNMAPNPQYPTDPA	875
<i>nit-4</i>	TSTGGIGDGGGPTGGAGDSMARLAPRNIIGGGGGGGGS--TGQRQOQQQROQQQQQQQ	1007
<i>YNA1</i>	-----	
<i>nirA</i>	NVNSNOTNMI FPGSPQR-----	892
<i>nit-4</i>	QQQQQQQQQQQQQQQQQANMFAYHHGAERGGGIESTGMGMITVGGGGGGFDPIGGSL	1067
<i>YNA1</i>	-----	
<i>nirA</i>	LDDLVLGLDELGSIDGLHPLGLD	1090
<i>nit-4</i>		
<i>YNA1</i>		

Figure 3 Alignment of the putative protein encoded by *YNA1* with those encoded by *nirA* from *A. nidulans* and *nit-4* from *N. crassa*

Identities are indicated by asterisks and conservative changes by dots.

constitutive levels of Yna1p would lead to a modification in the expression pattern of the nitrate assimilation structural genes. However, in the $\Delta yna1::URA3$ strain transformed with the *ADH1_p-YNA1* construct, *YNR1* expression, as well the nitrate reductase activity, remained as in the wild type (results not shown).

YNA1* gene product activates the transcription of *YNT1*, *YNI1* and *YNR1

The similarity of the *YNA1* gene product to those proteins encoded by *nirA* and *nit-4* [11,12] that are involved in the nitrate induction of the nitrate assimilation genes suggested a role for

Figure 4 Levels of *YNA1* transcripts

The amount of poly(A)⁺ RNA loaded in each lane was 2 µg. The cells from which RNA was extracted were grown in ammonium, washed and incubated at 10 mg wet weight of cells/ml for 2 h in the nitrogen sources indicated, each at 5 mM. The actin levels, probed with the actin gene (*ACT1*) from *S. cerevisiae*, were used to verify the amount of poly(A)⁺ RNA loaded in each lane.

the *YNA1* gene in nitrate regulation. To test a possible role for *YNA1* in nitrate induction, a strain bearing the null allele $\Delta yna1::URA3$ was constructed (Figure 1). Growth tests of the $\Delta yna1::URA3$ strain showed its inability to grow in nitrate and nitrite, whereas it was capable of growth in ammonium, glutamic acid, proline or hypoxanthine (results not shown). It was observed that the mutant did not induce nitrate reductase even after incubation for several hours in nitrate. This indicates that the incapacity of the $\Delta yna1::URA3$ strain to grow in nitrate is due at least in part to the absence of nitrate reductase. A conclusive explanation for the lack of growth of the disrupted strain as well as for the function of the *YNA1* gene was obtained by analysis of *YNT1*, *YNI1* and *YNR1* gene expression in the wild-type and disrupted strains. Both strains were grown in ammonium and transferred to ammonium-containing, ammonium plus nitrate-containing, nitrite-containing and nitrogen-free medium for 2 h. Total RNA analysed by Northern blotting showed that, whereas in the wild-type strain the transcripts of *YNT1*, *YNI1* and *YNR1* increased after the cells were transferred from ammonium to nitrate, in the $\Delta yna1::URA3$ strain no such increase was observed (Figure 5). Similarly in the disrupted strain no expression was observed in the nitrogen sources studied.

Because *H. polymorpha* possesses a high rate of non-homologous recombination, it could be argued that insertion or replacement of the $\Delta yna1::URA3$ construct in another location in the genome in addition to that detected by Southern blotting in the *YNA1* locus would lead to a false interpretation of the results. This hypothesis was discarded because the $\Delta yna1::URA3$ mutant that had been transformed with a replicative plasmid containing the *YNA1* gene recovered the wild-type phenotype (results not shown).

Figure 5 Disruption of the *YNA1* gene

(A) To disrupt the *YNA1* chromosomal copy, the *EcoRV*–*NdeI* fragment was replaced by the *H. polymorpha* *URA3* gene in the plasmid pNia2. The resulting plasmid, pYNA1-UR2, was digested with *SalI* and *BamHI*, liberating a fragment $\Delta yna1::URA3$ of approx. 3.9 kb that was used to transform a *H. polymorpha* *ura3, leu1* strain. (B) Southern blot of genomic DNA from the wild-type (WT) and $\Delta yna1::URA3$ strains, probed with the *PstI*–*EcoRV* DNA fragment. The wild type yielded bands of 1.6 and 3.3 kb with *HindIII* and *EcoRI* respectively, whereas the $\Delta yna1::URA3$ strain yielded 2.6 and 2.5 kb bands when the genomic DNA was digested with the same enzymes.

DISCUSSION

In this paper the characterization of a new gene involved in nitrate assimilation in *H. polymorpha* is reported. As far as is known the *YNA1* gene is the first transcriptional factor reported in *H. polymorpha*. Analysis of the structure and function of Yna1p indicates that this protein belongs to the C_6 zinc cluster family of the fungal regulatory proteins. Yna1p contains the six-cysteine cluster domain at the N-terminal region of the protein. Moreover, analysis of Yna1p by Lupas's method [41] predicts the presence of a coiled-coil structure formed by residues 93–114, which could be involved in the dimerization of the Yna1 protein. An acidic region located in the C-terminal part of the protein involved in the activation of these transcriptional factors [42] cannot be clearly delimited in Yna1p. With regard to the function of *YNA1* gene, its deletion to generate the strain $\Delta yna1::ura3$ clearly shows its involvement in the transcriptional activation of the genes encoding the enzymes involved in nitrate assimilation, because none of them was induced by nitrate or nitrite in this strain. The idea that *YNA1* might be involved in the transcriptional regulation of other pathways related to the assimilation of nitrogen was explored. However, the $\Delta yna1::URA3$ strain was able to grow in all the nitrogen sources checked such

as ammonium, glutamic acid, proline and hypoxanthine at the same rate as the wild type, indicating that *YNA1* is specific for the nitrate assimilation pathway, as are the *nit-4* and *nirA* genes in filamentous fungi.

YNA1 was expressed in all the nitrogen sources studied, with a decrease in the presence of ammonium, contrasting with reports for *nirA* and *nit-4*, for which similar expression is shown in all conditions studied [12,43]. The Northern blot analysis revealed the presence of a transcript of approx. 1.7 kb, in agreement with the size of the gene, that was not present in the $\Delta yna1::URA3$ strain (results not shown).

The nitrate assimilation gene cluster region contains the three structural genes involved in nitrate assimilation and a regulatory gene. The cluster contains 9205 bp, of which a 8820-bp length corresponds to the four ORFs, indicating that in this region approx. 95.8 % is coding. As a reference, in *S. cerevisiae* ORFs occupy on average 72 % of the genome [44]. The clustering of the nitrate and nitrite reductase genes involved in nitrate assimilation has been reported in all fungi explored with the exception of *N. crassa* [45,46]. However, one exclusive feature of the *H. polymorpha* gene cluster is the presence of the *YNA1* gene involved in the regulation of the other genes of the cluster. Another peculiarity of the *H. polymorpha* cluster is the fact that nitrite and nitrate reductase genes are convergently transcribed, unlike the divergent transcription of these genes in filamentous fungi. In the cluster all the genes are transcribed independently and in the same direction. The transcription initiation site of *YNR1* possesses a feature of highly expressed genes from *S. cerevisiae*, a CT-rich region followed by initiation at the CAAG sequence [47].

With regard to the regulation of the gene cluster, several facts indicate a hypothetical mechanism. The structural *YNR1* and *YNI1* gene promoters contain the sequences CGGCTCTGAA-AATTTCCG and CGGGACTCATAGTGTCCG, where the CGGX_nCCG palindromic sequences are separated by 12 nucleotides. These sequences could be the target site for the *YNA1* product. The DNA consensus binding-site sequences described for NIT4 and NIRA [14,15] are not the nitrate assimilation *H. polymorpha* gene promoters. Furthermore, analysis of the promoter of the genes in the cluster with the MATIND and MATINSPECTOR programs [48] reveals that several GATA core sequences, targets of the single zinc-finger proteins such as NIT2 and AREA, involved in nitrogen catabolism derepression, are present in the four genes. Therefore a *nit-2* and *areA* counterpart in *H. polymorpha* could be involved in controlling the regulation of the nitrate assimilation genes in yeast.

Nitrate assimilation is rare in yeasts and has been found only in some genera, namely *Hansenula*, *Pachysolen*, *Rhodhorula* and *Candida*. Either most yeasts have lost the capacity to use nitrate or they never possessed it. The clustering of the genes involved in nitrate assimilation in *H. polymorpha* and *H. anomala* (P. Garcia and J. M. Siverio, unpublished work) could be related to the few yeasts able to utilize nitrate. It has been proposed for bacteria that the acquisition of a new metabolic capacity by horizontal transfer of genes is easier if the genes are clustered [49].

The characterization of *YNA1* along with *YNT1*, *YNI1* and *YNR1*, together with the construction of the corresponding strains bearing the null alleles, consolidates *H. polymorpha* as a model for the study of the nitrate assimilation pathway in yeast; it will aid in the characterization of plant genes involved in this pathway.

We thank C. Gancedo (IIB-CSIC, Madrid, Spain) for much advice and for criticisms of the manuscript. This work was supported by grants PB94-0589 from Dirección General de Investigación Científica y Técnica (DGICYT) and 94/817 from Gobierno de Canarias to J. M. S.

REFERENCES

- 1 Guerrero, I. M., Vega, J. M. and Losada, M. (1981) *Annu. Rev. Plant. Physiol.* **32**, 169–204
- 2 Ávila, J., Pérez, M. D., Brito, N., González, C. and Siverio, J. M. (1995) *FEBS Lett.* **366**, 137–142
- 3 Brito, N., Ávila, J., Pérez, M. D., González, C. and Siverio, J. M. (1996) *Biochem. J.* **317**, 89–95
- 4 Pérez, M.D., González, C., Ávila, J., Brito, N. and Siverio, J. M. (1997) *Biochem. J.* **321**, 397–403
- 5 Johnstone, I. L., McCabe, P. C., Greaves, P., Gurr, S. J., Cole, G. E., Brow, M. A. D., Unkles, S. E., Clutterbuck, A. J., Kinghorn, J. R. and Innis, M. A. (1990) *Gene* **90**, 181–192
- 6 Banks, G. R., Shelton, P. A., Kanuga, N., Holden, D. W. and Spanos, A. (1993) *Gene* **131**, 69–78
- 7 Unkles, S. E., Campbell, E. I., Punt, P. J., Hawker, K. L., Contreras, R., Hawkins, A. R., Van den Hondel, C. A. M. J.J. and Kinghorn, J. R. (1992) *Gene* **111**, 149–155
- 8 Quesada, A., Galván, A. and Fernández, E. (1994) *Plant J.* **5**, 407–419
- 9 Crawford, N. M. and Arst, Jr., H. N. (1993) *Annu. Rev. Genet.* **27**, 115–146
- 10 Marzuf, G. A. (1997) *Microbiol. Mol. Biol. Rev.* **61**, 17–32
- 11 Yuan, G.-F., Fu, Y.-H. and Marzluf, G. A. (1991) *Mol. Cell. Biol.* **11**, 5735–5745
- 12 Burger, G., Strauss, J., Scazzocchio, C. and Lang, B. F. (1991) *Mol. Cell. Biol.* **11**, 5746–5755
- 13 Vashee, S., Xu, H., Johnson, S. A. and Kodadek, T. (1993) *J. Biol. Chem.* **268**, 24699–24706
- 14 Punt, P. J., Strauss, J., Smit, R., Kinghorn, J. R., Van den Hondel, C. A. M. J. J. and Scazzocchio, C. (1995) *Mol. Cell. Biol.* **15**, 5688–5699
- 15 Fu, Y. H., Feng, B., Evans, S. and Marzluf, G. A. (1995) *Mol. Microbiol.* **15**, 935–942
- 16 Marmorstein, R., Carey, M., Ptashne, M. and Harrison, S. C. (1992) *Nature (London)* **356**, 408–414
- 17 Marmorstein, R. and Harrison, S. C. (1994) *Genes Dev.* **8**, 2504–2512
- 18 Swaminathan, K., Flynn, P., Reece, R. J. and Marmorstein, R. (1997) *Nature Struct. Biol.* **4**, 751–759
- 19 Strauss, J., Muro-Pastor, M. I. and Scazzocchio, C. (1998) *Mol. Cell. Biol.* **18**, 1339–1348
- 20 Fu, Y.-H. and Marzluf, G. A. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 5331–5335
- 21 Kudla, B., Caddick, M. X., Langdon, T., Martinez-Rossi, N. M., Bennett, C. F., Sibley, S., Davies, R. W. and Arst, Jr., H. N. (1990) *EMBO J.* **9**, 1355–1364
- 22 Xiao, X., Fu, Y.-H. and Marzluf, G. A. (1995) *Biochemistry* **34**, 8861–8868
- 23 Merckelbach, A., Gödecke, S., Janowicz, Z. A. and Hollenberg, C. P. (1993) *Appl. Microbiol. Biotechnol.* **40**, 361–364
- 24 Faber, K. N., Haima, P., Harder, W., Veenhuis, M. and Ab, G. (1994) *Curr. Genet.* **25**, 305–310
- 25 Agaphonov, M. O., Poznyakovski, A. I., Bogdanova, A. I. and Ter-Avanesyan, M. D. (1994) *Yeast* **10**, 509–513
- 26 Rothstein, R. (1985) in *DNA cloning II* (Glover, D. M., ed.), pp. 45–66, IRL Press, Oxford
- 27 Carlson, M. and Botstein, D. (1982) *Cell* **28**, 145–154
- 28 Kingston, R. E. (1987) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. O., Seidman, J. G., Smith, J. A. and Struhl, K., eds.), pp. 4.5.1–4.5.3, John Wiley, New York
- 29 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 30 Engler-Blum, G., Meier, M., Frank, J. and Müller, G. A. (1993) *Anal. Biochem.* **210**, 235–244
- 31 Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- 32 Higgins, D. G. and Sharp, P. M. (1988) *Gene* **73**, 237–244
- 33 Triezenberg, S. J. (1992) in *Current Protocols in Molecular Biology* (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. O., Seidman, J. G., Smith, J. A. and Struhl, K., eds.), pp. 4.8.1–4.8.5, John Wiley, New York
- 34 González, C. and Siverio, J. M. (1992) *J. Gen. Microbiol.* **138**, 1445–1451
- 35 Hahn, S., Hoart, E. T. and Guarente, L. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8562–8566
- 36 MacNeil, J. B. and Smith, M. (1986) *J. Mol. Biol.* **187**, 363–378
- 37 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402
- 38 Holmberg, S. and Schjerling, P. (1996) *Genetics* **144**, 467–478
- 39 Geever, R. F., Huiet, L., Baum, J. A., Tyler, B. M., Patel, V. B., Rutledge, B. J., Case, M. E. and Giles, N. H. (1989) *J. Mol. Biol.* **207**, 15–34
- 40 Hedges, D., Proft, M. and Entian, K. D. (1995) *Mol. Cell. Biol.* **15**, 1915–1922
- 41 Lupas, A., Van Dyke, M. and Stock, J. (1991) *Science* **252**, 1162–1164
- 42 Schjerling, P. and Holmberg, S. (1996) *Nucleic Acids Res.* **24**, 4599–4607
- 43 Fu, H.-F., Kneesi, J. Y. and Marzluf, G. A. (1989) *J. Bacteriol.* **171**, 4067–4070
- 44 Dujon, B. (1996) *Trends Genet.* **12**, 263–270
- 45 Exley, G. E., Colandene, J. D. and Garret, R. H. (1993) *J. Bacteriol.* **175**, 2379–2392
- 46 Okamoto, P. M., Fu, Y.-H. and Marzluf, G. A. (1991) *Mol. Gen. Genet.* **227**, 213–223
- 47 Rathjen, J. and Mellor, J. (1990) *Nucleic Acids Res.* **18**, 3219–3225
- 48 Quandt, K., Frech, K., Karas, H., Wingender, E. and Werenr, T. (1995) *Nucleic Acids Res.* **23**, 4878–4884
- 49 Lawrence, J. G. and Roth, J. R. (1996) *Genetics* **143**, 1843–1860