J. Microbiol. Biotechnol. (2004), **14**(6), 1239–1248

JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY

© The Korean Society for Microbiology and Biotechnology

Fabrication of a Partial Genome Microarray of the Methylotrophic Yeast *Hansenula polymorpha:* Optimization and Evaluation of Transcript Profiling

OH, KWAN SEOK[1], OHSUK KWON[1], YUN WI OH[1], MIN JEONG SOHN[1], SOONGEE JUNG[1], YONG KYUNG KIM[1], MIN-GON KIM[1], SANG KI RHEE[1], GERD GELLISSEN[2], AND HYUN AH KANG[1]*

[1]*Laboratory of Metabolic Engineering, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-333, Korea*
[2]*Rhein Biotech GmbH, Eichsfelder Str. 11, 40595 Dusseldorf, Germany*

Received: February 25, 2004
Accepted: June 27, 2004

Abstract The methylotrophic yeast *Hansenula polymorpha* has been extensively studied as a model organism for methanol metabolism and peroxisome biogenesis. Recently, this yeast has also attracted attention as a promising host organism for recombinant protein production. Here, we describe the fabrication and evaluation of a DNA chip spotted with 382 open reading frames (ORFs) of *H. polymorpha*. Each ORF was PCR-amplified using gene-specific primer sets, of which the forward primers had 5'-aminolink. The PCR products were printed in duplicate onto the aldehyde-coated slide glasses to link only the coding strands to the surface of the slide via covalent coupling between amine and aldehyde groups. With the partial genome DNA chip, we compared efficiency of direct and indirect cDNA target labeling methods, and found that the indirect method, using fluorescent-labeled dendrimers, generated a higher hybridization signal-to-noise ratio than the direct method, using cDNA targets labeled by incorporation of fluorescence-labeled nucleotides during reverse transcription. In addition, to assess the quality of this DNA chip, we analyzed the expression profiles of *H. polymorpha* cells grown on different carbon sources, such as glucose and methanol, and also those of cells treated with the superoxide-generating drug, menadione. The profiles obtained showed a high-level induction of a set of ORFs involved in methanol metabolism and oxidative stress response in the presence of methanol and menadione, respectively. The results demonstrate the sensitivity and reliability of our arrays to analyze global gene expression changes of *H. polymorpha* under defined environmental conditions.

Key words: DNA microarray, transcript profiling, *Hansenula polymorpha*, target labeling, open reading frames

*Corresponding author
Phone: 82-42-860-4378; Fax: 82-42-860-4594;
E-mail: hyunkang@kribb.re.kr

Recent advances of DNA microarray technology provide a way to explore the metabolic and genetic control of genome-wide gene expression [7, 17]. It is now possible to investigate differential gene expression profiling for thousands of genes in a single experiment. The DNA array techniques have been most comprehensive in the traditional yeast *Saccharomyces cerevisiae*, which is the first eukaryotic organism with the whole genome completely sequenced [14]. Analysis of gene expression patterns of *S. cerevisiae* using DNA microarray has contributed to the understanding of physiological aspects of cells, such as cell division and stress response to various environmental conditions [11, 26]. Apart from basic science, genome-wide expression analyses have been applied to obtain valuable information on the metabolic pathways in the industrial *S. cerevisiae* strains under several biotechnology processes, such as winemaking [10, 24]. DNA microarray technology is now employed to study other yeast species such as the fission yeast *Schizosaccharomyces pombe* and the pathogenic yeast *Candida albicans*. Whole-genome DNA microarrays of *S. pombe* [5, 22] and a partial genome array of *C. albicans* [6, 23] have been developed and used for genome expression profiling studies.

The thermotolerant methylotrophic yeast *Hansenula polymorpha* (*Pichia angusta*) has gained increasing interest as a useful system for fundamental research and applied purposes. *H. polymorpha* has been an excellent model system to study the genetic control mechanism of methanol metabolism, nitrate assimilation, and peroxisome biogenesis [1, 4, 28]. Moreover, it has recently become one of the preferred microbial host organisms for the production of recombinant proteins [13, 29]. The partial genome sequence information of *H. polymorpha* has been available to the public as a part of the comparative genome analysis project on 13 hemiascomycetous yeasts with biomedical and biotechnological potentials [3]. In this study, we fabricated

a cDNA microarray spotted with 382 *H. polymorpha* ORFs and optimized procedures for target labeling and hybridization. The partial cDNA array was further evaluated for differential expression profiling of *H. polymorpha* cells cultivated under different environmental conditions.

MATERIALS AND METHODS

Strain, Culture Conditions, and Total RNA Isolation

The yeast strain used for this study was *H. polymorpha* A16 (*leu2*), derived from *H. polymorpha* CBS4732 strain [30]. Yeast cells were grown in YP (1% yeast extract, 2% bacto peptone) medium with 2% glucose (YPD) or 1% methanol (YPM). For menadione treatment, the culture was treated with menadione bisulfite (Sigma, St. Louis, MO, U.S.A.) at concentration of 1 mM for 2 h. Cells were cultured to the exponential phase (OD 0.8- 1.0 at 600 nm) in shake flasks at 37°C, and total RNA was isolated by hot phenol method [9].

PCR Amplification and Microarray Processing

For PCR amplification of *H. polymorpha* ORFs, primer sets of 1,202 ORFs were designed based on the publicly available random sequenced tags (RSTs) information of *H. polymorpha* CBS4732 strain (http://pedant.gsf.de/cgi-bin/wwwfly.pl?Set=Pichia_angusta_CBS_4732_RST&Page=index). The primers were 18 mer in length and designed to contain 5'-aminolink only at the forward primers. The sequence information on the primer sets is available at our web-site (http://www.kribb.re.kr/metabolic/). Genomic DNA of *H. polymorpha* strain A16 was isolated as previously described [16] and used as a template for PCR amplification reactions in 96-well PCR plates. PCR reaction mixture was composed of 8 μg of chromosomal DNA, 100 pmoles of forward and reverse primers, and 50 μl of 2× bulk premix Taq DNA polymerase (Bioneer, Korea) in 100 μl of total volume. PCR products were purified and concentrated to 100 ng/μl in 20 μl 1× spotting buffer (TeleChem International, Sunnyvale, CA, U.S.A.) using 96-well purification system (Qiagen, Chatsworth, CA, U.S.A.). Purified PCR products were separated on agarose gels by electrophoresis to determine relative DNA concentrations. Samples were then transferred to a 384-well plate to print onto super aldehyde slide glasses (CEL associates, Houston, TX, U.S.A.) with a ProSys5510 microarrayer (Cartesian Technologies, Irvine, CA, U.S.A.). Printed slides were treated with sodium borohydride solution and stored in a dust-free and lightproof container until use. Before hybridization reaction, microarray slides were prehybridized in buffer containing 5× SSC, 0.1% SDS, and 1% Bovine Serum Albumin (BSA) at 50°C for 30 min. The slides were washed stepwise by 2× SSC and by 0.2× SSC for 3 min each at room temperature. The slides were centrifuged at 1,000 rpm for 3 min to dry in 50 ml uncapped centrifuge tubes.

Target Labeling and Hybridization

We slightly modified direct-target labeling methods [15] as follows. A mixture of total RNA (100 μg) and oligo-dT (2 μg) (in a total volume of 24 μl) was heated at 70°C for 10 min and quickly chilled on ice. The following were then sequentially added: 5× reverse transcriptase reaction buffer, 0.1 M DTT (Invitrogen, Carlsbad, CA, U.S.A.), 10× dNTP (5 mM d(AGC)TP, 2 mM TTP), fluorescent nucleotide (1 mM Cy3™-dUTP or Cy5™-dUTP) (Amersharm Biosciences, Piscataway, NJ, U.S.A.), RNAsin (RNase inhibitor), and 200 U of Superscript™ reverse transcriptase (Invitrogen) to a total volume of 50 μl. The reaction mixture was incubated at 42°C for 2 h to generate Cy3- or Cy5-labeled target cDNAs. The target cDNAs were purified using MinElute™ PCR purification kit (Qiagen), vacuum dried, and resuspended in 15 μl TE buffer (10 mM Tris-Cl, 1 mM EDTA pH 8.5). After addition of 8.8 μl of hybridization buffer (10% SDS, 20× SSC), the labeled target cDNAs were heated at 99°C for 5 min, cooled at room temperature for 5 min, and applied to the prehybridized slide in a hybridization chamber. The chamber was tightly sealed and incubated in a 45°C water bath for 16- 18 h. The slide was then removed from the chamber, washed sequentially for 10 min each in 2× SSC/0.2% SDS buffer, 2× SSC buffer, and 0.2× SSC buffer, rinsed in distilled water for 2 min, and dried by centrifugation for 3 min at 1,000 rpm in 50 ml uncapped centrifuge tubes.

The labeling of target cDNAs by indirect method was performed using 3DNA™ Submicro EX Expression Array Detection Kit, according to the manufacturer's protocol (Genisphere, Hatfield, PA, U.S.A.). Briefly, total RNA was reverse transcribed using dNTP mix with the reverse transcription primers, which were tagged with either Cy3- or Cy5-specific 3DNA capture sequence. The synthesized cDNAs were incubated at 80°C for 10 min and at 50°C for 20 min, and then applied to the prehybridized slide. After the first hybridization, the fluorescent Cy3- or Cy5-3DNA reagents were added to the array for the subsequent hybridization with the synthesized cDNAs.

Microarray Data Processing and Analysis

The hybridized slides were scanned with ScanArray 5000 scanner (Gsi Lumonics, Kanata, OT, U.S.A.) using appropriate photomultiplier tube (PMT) value to obtain the highest intensity without saturation. A 16 bit TIFF image was generated for each channel of Cy3 and Cy5. Both channels were normalized against each other using GenePix™ Pro 4.0 (Axon Instruments, Foster City, CA, U.S.A.) software. Spot intensity and background signals were quantified for each channel and transferred to Microsoft Excel spreadsheet for more specific analysis. Spots with the signal intensity significantly greater than local background in duplicated arrays were included for further normalization using the program S-PLUS (Insightful, Seattle, WA, U.S.A.) [27].

The data of spots showing more than 2-fold induction or repression were extracted for hierarchical clustering using the program Cluster [8]. The resulting cluster was visualized using the program TreeView, which is available at http://rana.lbl.gov.EisenSoftware.htm. The raw data for microarray analysis presented in this study are also available as Excel files at our web site (http://www.kribb.re.kr/metabolic/).

RT-PCR Analysis

To validate DNA microarray data, we carried out a semi-quantitative reverse transcription PCR (RT-PCR) as described previously [21]. Thus, total RNA (10 μg) was converted into cDNA using the 1st strand cDNA synthesis kit for RT-PCR (Roche Molecular Biochemicals, Mannheim, Germany), and the product was then subjected to PCR with the same primers that were used to amplify the ORF to spot on the microarray. To avoid saturation of amplification, samples were taken and analyzed from the reaction mixture at every 5 cycles up to 35 cycles after the initial 15 cycles of elongation reaction (72°C for 2 min).

Construction of GOD Expression Plasmids for Promoter Analysis

Plasmid pDLMOX-GOD was constructed by subcloning the glucose oxidase (GOD) expression cassette fused with the *Aspergillus oryzae* α-amylase secretion signal [20] between the *Eco*RI-*Hin*dIII sites of pDLMOX-HIR [18]. Plasmid pCAT-GOD was then constructed by exchanging the methanol oxidase (*MOX*) promoter between the *Bam*HI-*Hin*dIII sites of pDLMOX-GOD with the peroxisomal catalase (*CAT*) promoter, which was amplified from the *H. polymorpha* chromosomal DNA by PCR and digested with *Bgl*II and *Hin*dIII.

RESULTS AND DISSCUSSION

Primer Design, PCR Amplification, and Spotting

To construct DNA microarrays, several approaches can be used to immobilize DNA fragments on a solid support. A comparative study on different strategies of covalent attachment of DNA to glass surfaces [33] has suggested

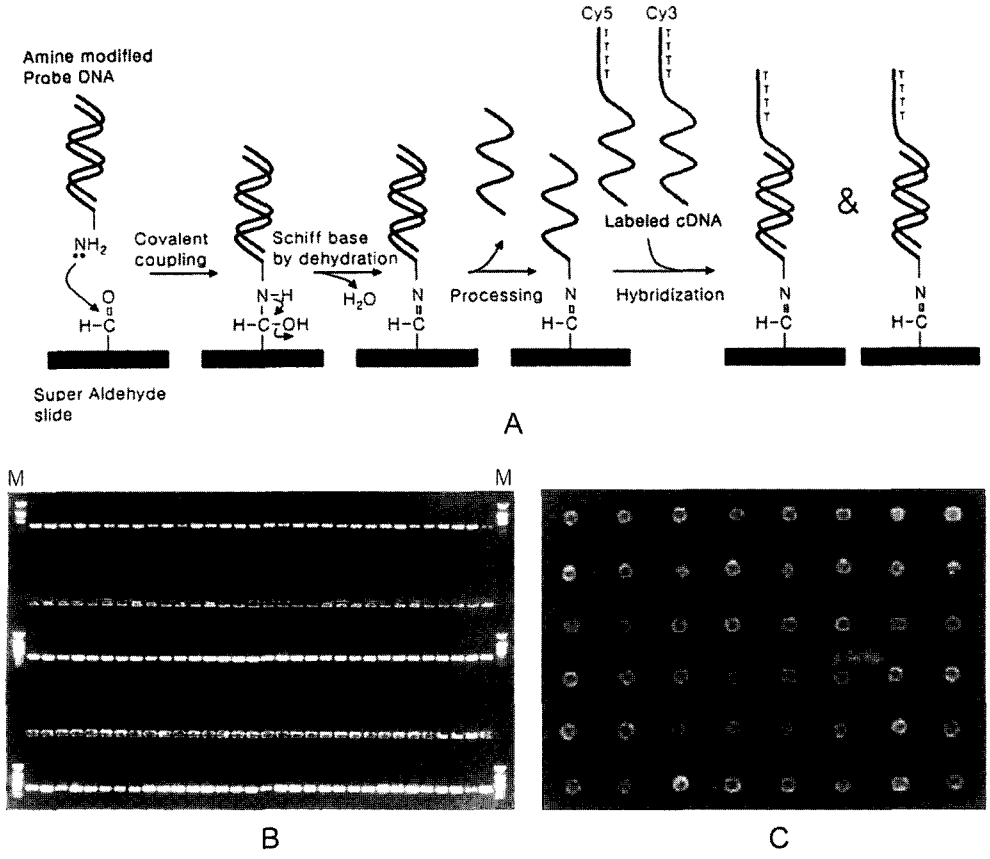

Fig. 1. Construction of *H. polymorpha* partial genome cDNA microarray.
(A) Schematic representation of surface chemistry, processing, and hybridization of *H. polymorpha* cDNA microarray. (B) Agarose gel separation of PCR products. Left and right lines designated as M contain the standard DNA size marker (100 bp DNA ladder, Bioneer). (C) Syber green-staining of microarray slide after spotting. Images for one representative agarose gel containing 96-PCR samples after purification, and a part of DNA array after spotting are presented.

fixation of aminated DNA to an aldehyde-modified surface as a method of choice to build DNA microarrays. Thus, we decided to use this strategy to fabricate a cDNA microarray of *H. polymorpha*. Moreover, we designed gene-specific primers (18-mer) to contain amine modification only at 5'-end of the forward primer to allow strand-specific probe spotting. The resultant PCR products contained 5'-aminolink only at the coding strands, which could be tethered at their ends on the surface of aldehyde-coated glass via covalent coupling between amine and aldehyde groups. This would generate single-stranded array elements, minimizing the interference of complementary strands during hybridization (Fig. 1A). The similar strategy was employed for the construction of *S. pombe* cDNA array with the strand specificity, which can discriminate transcriptional direction. It has been shown that single-stranded array elements gave higher sensitivity than double-stranded elements of the same size and sequence on the glass slides [22].

At the time when we had initiated our work on the development of the *H. polymorpha* DNA microarray, only partial sequence information on about 2,500 predicted open reading frames (ORFs) of *H. polymorpha* were publicly available as random sequenced tags [3]. We selected 1,202 ORFs, for which the coding region over 750 base pairs could be obtained as PCR products from the total chromosomal DNA. The PCR amplification was carried out highly efficiently with a success rate of >90%, as shown in Fig. 1B. Purified PCR products were concentrated in 20 μl spotting buffer at an average concentration of 650 ng/μl. We had initially spotted duplicated sets of 384 PCR products, including 380 samples amplified from randomly selected ORFs and duplicated samples of 2 control ORFs, methanol oxidase (*MOX*) and actin (*ACT*) genes. The resulting microarray chip was named KR-HPM-384 (KRIBB-*H. polymorpha* Microarray-384). The spot quality on the KR-HPM384 chip was confirmed by the homogeneous spot intensity obtained after Syber Green 2 staining (Fig. 1C). The full list of ORFs spotted on the KR-HPM-384 chip is available from our website (http://www.kribb.re.kr/metabolic/).

Optimization of Target Labeling and Hybridization Process

Owing to its high-throughput nature, cDNA microarray technology is vulnerable to several systematic variations introduced during experimental processes. Thus, it is important to set up standard procedures for labeling cDNA targets, prehybridization, hybridization, and slide washing to consistently generate high intensity and low background images. In an effort to optimize the conditions of analysis using KR-HPM-384, we have compared the efficiency of two different cDNA labeling methods by self-against-self hybridization. Thus, cDNA targets were prepared from total RNA isolated from the *H. polymorpha* cells cultivated

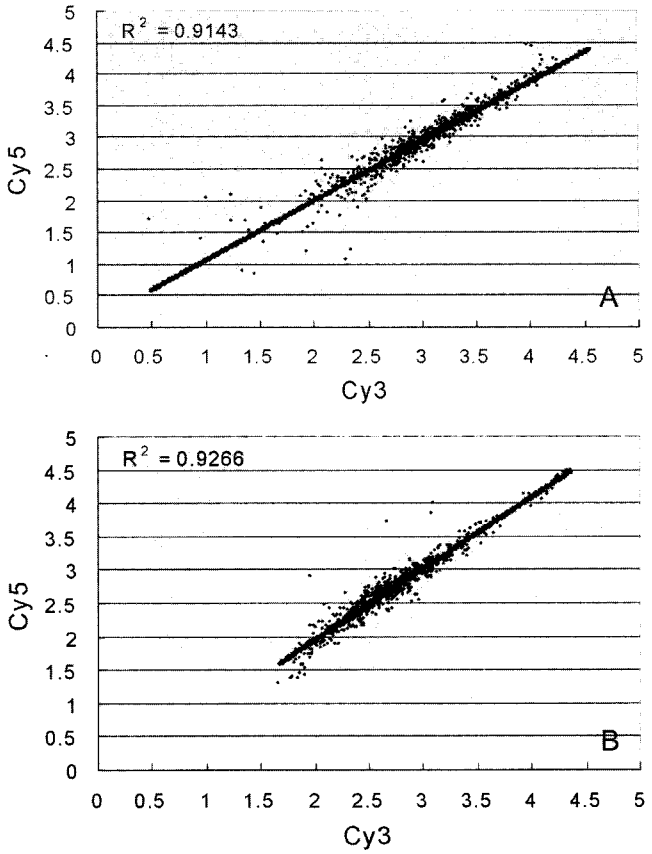

Fig. 2. Scatter plots showing the distribution of Cy5 and Cy3 signal intensities by self-self hybridization experiments.
Total RNAs prepared from *H. polymorpha* grown in YPD media were labeled with Cy-3 and Cy-5 dyes, respectively, and subjected to hybridization on the same arrays. (A) Direct labeling (starting RNA 100 μg), (B) 3DNA-genisphere indirect labeling (starting RNA 25 μg).

in YPD medium, labeled with both Cy3 and Cy5, and then hybridized on the same array. The self-self hybridizations should ideally produce similar signal intensity in both channels for every spot. Global normalization was applied to normalize Cy5 signal intensity against Cy3, on an assumption that total intensity of Cy5 channel is equal to that of Cy3 [2, 31]. We observed that the 3DNA indirect labeling method using fluorescent-labeled dendrimers exhibited a higher hybridization signal-to-noise ratio than the direct-labeling method, in which fluorescent nucleotides were directly incorporated into the cDNA targets (Fig. 2). 3DNA labeling procedure consistently showed R-square values of 0.93 or higher and tight scatter plots, especially even for spots with low hybridization signals. Furthermore, the 3DNA indirect method generated reliable hybridization signals with much lower amount of starting RNA (about 25 μg RNA vs. 100 μg RNA) than the direct labeling method. With the optimized 3DNA labeling method, we were able to detect expression of 362 (94%) among the 382 *H. polymorpha* ORFs spotted on the array.

Analysis of Differential Gene Expression Depending on Carbon Source

To evaluate the performance of our partial genome microarray for differential gene expression profiling, we carried out transcriptome analysis of *H. polymorpha* cells cultivated in different carbon sources, glucose and methanol. Cells were grown up to 0.5 OD_{600} in YPD medium containing 2% glucose as carbon source, harvested, washed, and then reinoculated into YPD medium or YPM medium containing 1% methanol instead of 2% glucose. After 2 h of cultivation, total RNAs were prepared from the cells cultivated on YPD and YPM, respectively. cDNA targets were prepared from each RNA sample by indirected 3DNA labeling method and hybridized to KR-HPM-384 arrays to identify

Table 1. *H. polymorpha* genes regulated by more than two-fold after cultivation on methanol for 120 min.

RST code	Fold regulation (>2)		Gene description
	Induction	Repression	
DNA recombination and repair			
bb0aa004c08t1	3.00		DNA ligase (ATP) precursor
bb0aa009a02d1	2.71		CTF4 protein
Cell budding and cycle			
bb0aa009f07t1	2.60		*Eremothecium gossypii* Ycl016c gene
bb0aa021h11dp1	2.17		Exoribonuclease RAT1
bb0aa007g11t1	2.25		Cytokinesis (cell division), budding, cell growth
Protein synthesis and transport			
bb0aa023g11dp1	2.69		Vacuolar protein-sorting protein YPS33
bb0aa031e12dp1	2.62		Dynein heavy chain, cytosolic (DYHC)
bb0aa024f11tp1		2.49	*Candida glabrata* TEF3 gene for translation elongation factor
C-compound and carbohydrate utilization			
MOX	18.68		Methanol oxidase
xbb0aa001h06t1	4.00		Formaldehyde transketolase
bb0aa004g08t1	2.67		Probable finger protein YKL038w
Amino acid metabolism			
bb0aa011a01tp1	2.65		CAR1-arginase
bb0aa018b01tp1		3.17	Dimethylaniline monooxygenase-like protein
Unclassified proteins			
bb0aa011e01t1	2.35		Probable membrane protein YLR328w
bb0aa023f11tp1	2.15		Phosphoribosyl-ATP diphosphatase
bb0aa011g04tp1	2.23		*Mus musculus* yil097wFYV10-erythroblast macrophage protein EMP
bb0aa024g02tp1	2.95		Hypothetical protein YBL029w
Lipid, fatty acid, and isoprenoid metabolism			
bb0aa008e12d1	2.39		*Pichia pastoris* acyl-coenzyme A oxidase (*POX1*) gene
bb0aa009d03d1		4.24	Threonine aldolase
Transport			
bb0aa012h09dp1	7.22		Related to myo-inositol transport protein ITR1
bb0aa024b08tp1	2.66		Uric acid-xanthine permease (UAPA transporter)
bb0aa013h05tp1		6.63	Iron transport multicopper oxidase, Fet3
Nitrogen and sulfur utilization			
bb0aa025d06dp1	3.80		Probable nitrate assimilation transcription factor
bb0aa025a07tp1	2.14		Aspartate transaminase-alfalfa
bb0aa001b03t1	5.62		Dihydropyrimidinase
bb0aa010d08tp1	2.90		UGA3-transcriptional activator for GABA
xbb0aa001b07t1		3.75	Histidinol dehydrogenase, phosphoribosyl-ATP diphosphatase
Detoxification and stress response			
bb0aa023c06dp1	12.48		Peroxisomal catalase (*CAT*)
bb0aa023c06tp1	5.24		NADH:ubiquinone oxidoreductase 51kDa subunit
bb0aa011e11dp1	2.10		SIT1-Transporter of the bacterial siderophore ferrioxamine B
Transposable elements			
bb0aa010e01tp1	2.51		Hypothetical protein-putative pseudo-TY5
Signal transduction			
bb0aa001d02d1	2.36		Hypothetical protein T19A5.2-*Caenorhabditis elegans*

Annotations for Table 1 are from the PEDANT genome database at MIPS (http://pedant.gsf.de/) and some of them being edited by hand.

A

RT-PCR Glucose	RT-PCR Methanol	Gene	Microarray fold
		Actin	Control (0)
		Related to myo-inositol transport protein ITR1	Induction (7.22)
		Iron transport multicopper oxidase	Repression (6.63)

B

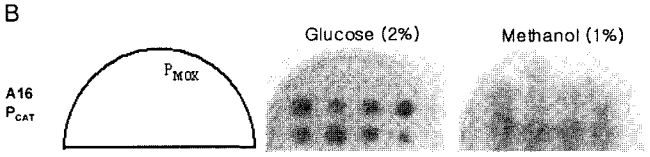

Fig. 3. Validation of methanol-regulated *H. polymorpha* gene expression.

(A) RT-PCR. Ten ng of total RNA were converted into the 1st strand cDNA and then subjected to PCR. (B) Expression of *A. oryzae* GOD reporter gene. Histochemical staining showing the level of GOD activity in the transformants of A16 strain containing P_{MOX}-GOD or P_{CAT}-GOD constructs. The *H. polymorpha* transformants were patched and incubated on SD minimal plates containing 2% glucose or 1% methanol, respectively, for 36 h, and 10 ml of reaction agar [20] were then applied over the plate, followed by incubation at 37°C until color appeared.

differentially expressed genes. As listed in Table 1, among 382 ORFs spotted on the array, 32 ORFs changed their expression levels by two-fold or greater (Table 1). In particular, the expression of genes coding for methanol oxidase (*MOX*), formaldehyde transketolase, and peroxisomal catalase (*CAT*), which are known to be involved in the methanol metabolism, were highly induced after shift to YPM medium, supporting prior knowledge on methylotrophic yeasts [32]. Interestingly, the microarray analysis showed that several genes involved in nitrogen and sulfur utilization were also significantly induced in the cells cultivated on methanol. The reproducibility of the expression pattern was confirmed by almost identical signal ratios in duplicated spots within an array and by reverse signal patterns in the dye swapping experiments (data not shown).

The changes of gene expression levels observed by cDNA microarray analysis were further confirmed by semi-quantitative RT-PCR analysis (Fig. 3A). There was a close correlation between the microarray data and the RT-PCR analysis, validating the results of cDNA microarray. On the other hand, the expression pattern of *CAT*, which showed a high induction ratio comparable to that of *MOX* in the microarray analysis, was confirmed by evaluating

Table 2. *H. polymorpha* genes regulated by more than two-fold after menadione treatment for 60 min.

RST code	Fold repression		Gene description
	Induction	Repression	
Detoxification and stress response			
bb0aa018a06tp1	15.22		Drug transporter of *Mycobacterium tuberculosis* CDC1551
bb0aa023c06dp1	25.24		Peroxisomal catalase
bb0aa011g05tp1	3.34		HAP1-transcription factor
bb0aa011e11dp1	6.61		SIT1-Transporter of the bacterial siderophore ferrioxamine B
bb0aa010h02dp1	3.54		WSC3-cell wall integrity and stress response component 3
Protein modification			
bb0aa022h10tp1	5.55		HOC1 protein
bb0aa012a06tp1	2.48		MNS1-α,1,2-mannosidase
bb0aa022d12dp1		5.55	Probable α-1,2-manosyltransferase
bb0aa010g05tp1		4.48	Similarity to Mnn4p
bb0aa021h11tp1		7.16	TTP1 protein
DNA repair			
bb0aa025f01tp1	3.51		3-Methyladenine DNA glycosylase
Budding, cell polarity, and filament formation			
bb0aa009f07t1	5.22		*Eremothecium gossypii* Ycl016c gene
bb0aa031b05dp1	4.21		Probable membrane protein YIL140w
C-compound and carbohydrate utilization			
bb0aa020f11tp1	5.45		Hypothetical protein; similar to *S. cerevisiae* PSP1 suppressors of temperature sensitive mutations
bb0aa010h01tp1	3.86		PUT3-positive activator of the proline utilization pathway
bb0aa026d09dp1	4.36		Pyruvate carboxylase
bb0aa017e12tp1	4.14		MFS transporter of unknown specificity
bb0aa020f05dp1		6.21	Mannose-6-phosphate isomerase
bb0aa018e11tp1		8.73	*P. angusta* GDP-mannose pyrophosphorylase

Table 2. Continued.

RST code	Fold repression		Gene description
	Induction	Repression	
Ribosome biogenesis, translation			
bb0aa013c09dp1	7.30		Translation initiation factor eIF-4a (TIF2)
bb0aa020h12tp1		34.64	Ribosomal protein S3.e, cytosolic
bb0aa024h11tp1		3.16	*Candida glabrata* TEF3 gene for translation elongation factor3
bb0aa022e02tp1		5.19	Eukaryotic initiation factor 4A eIF-4A (TIF1)
bb0aa003g01t1		10.31	Polyadenylate binding protein II-human
Lipid, fatty-acid, and isoprenoid metabolism			
bb0aa009e08t1		15.98	*Aspergillus nidulans* oleate d-12 desaturase (*odeA*) gene
bb0aa017c08dp1		3.00	Regulatory protein SIN3
bb0aa020f09tp1		3.57	Glutamate synthase (NADH2) glt1 precursor
bb0aa030f07tp1		5.70	Regulatory protein GAL4
Transport			
bb0aa010h11tp1		5.10	SLA1-cytoskeleton assembly control protein
bb0aa017b05dp1		2.95	Probable membrane protein YLR409c
bb0aa028d02tp1		2.93	HD exchanging protein
bb0aa013h05tp1		3.18	Iron transport multicopper oxidase, Fet3
bb0aa026f10tp1		3.96	H+transporting two-sector ATPase α protein precursor
bb0aa002c01t1		9.91	*Pichia angusta* α-COP-like protein gene
xbb0aa002a08t1		5.23	ATP synthase β chain, mitochondrial precursor
Assembly of protein complexes			
bb0aa002d01t1	2.64		β-Adaptin of human
Amino acid metabolism			
bb0aa022f10tp1	4.01		Branched-chain-amino-acid transaminase BAT1 precursor
Metabolism of vitamins, cofactors, and prosthetic groups			
bb0aa027b12dp1		4.61	Farnesyl-pyrophosphate synthetase
Cytoplasmic and nuclear degradation			
bb0aa017c06dp1	2.73		26S proteasome regulatory particle chain RPN2
Mitotic cell cycle and cell cycle control			
xbb0aa002b05t1		7.21	Probable spindle pole body associated protein alp4
bb0aa009d09d1		3.22	Chromosome segregation protein SMC1
Unclassified proteins			
bb0aa026g07tp1	3.25		Probable membrane protein YNL163c
bb0aa023e0dp1	2.26		Probable membrane protein YNR013c
bb0aa015f09tp1	2.55		Probable membrane protein YNL244c
bb0aa022d02tp1		2.41	Hypothetical protein YFR029w
bb0aa018d06tp1		6.51	Hypothetical protein YJR056c
bb0aa019a11dp1		6.39	Hypothetical protein YNL041c
bb0aa008e06d1		3.35	Hypothetical protein T20N10.10-*Arabidopsis thaliana*
bb0aa007a04t1		2.57	Hypothetical protein YIL144w
bb0aa023b09tp1		5.16	Probable serine/threonine-specific protein kinase YKL171w
bb0aa020f02tp1		4.68	Hypothetical protein-*Arabidopsis thaliana*
bb0aa008a06t1		3.82	Hypothetical protein 15E6.120-*Neurospora crassa*
bb0aa018d11tp1		3.96	Hypothetical protein YNL118c
bb0aa001d02d1		2.93	Hypothetical protein T19A5.2-*Caenorhabditis elegans*
bb0aa003h09t1		2.26	Gene J0916 of *S. cerevisiae* ORFs from chromosome X
bb0aa029d04tp1		4.10	Probable membrane protein YNL247w
bb0aa027d09dp1		5.54	Probable membrane protein YLR241w
bb0aa011f07dp1		2.42	ygr090w-similarity to PIR: T40678 hypothetical SPBC776.08c *S. pombe*
bb0aa004f07t1		5.65	Hypothetical transmembrane protein
bb0aa008g07t1		3.54	Probable membrane protein YPR105c

Annotations for Table 2 are from the PEDANT database at MIPS (http://pedant.gsf.de/), and some of them being edited by hand.

the regulated expression of *A. oryzae GOD* under the *MOX* and *CAT* promoters, respectively (Fig. 3B).

Analysis of Differential Gene Expression upon Menadione Treatment

We also analyzed changes in gene expression profile in the response of cells to the superoxide-generating drug menadione using the *H. polymorpha* partial genome microarray. Menadione (2-methyl-1,4-naphthoquinone) is a component of multivitamin drugs and has been used as a therapeutic agent for cancer. However, free radicals generated by menadione were reported to cause a time-dependent cell injury [19]. To examine the effect of menadione on gene expression, exponentially growing *H. polymorpha* cells were treated with 1 mM menadione for 2 h. Total RNAs were prepared from the cells cultivated on YPD with and without menadione, respectively, and were used to prepare cDNA targets to hybridize to KR-HPM-384 arrays. The obtained microarray data indicated that the expression of 61 ORFs among 382 ORFs on the array were reproducibly induced (21 ORFs) or repressed (40 ORFs) at least two-fold in the cells cultivated with menadione (Table 2). As previously reported in the transcriptome analysis of *S. cerevisiae* cultivated in the presence of menadione [11], the response of *H. polymorpha* to menadione was characterized by strong induction of genes known to be involved in the detoxification and stress response to reactive oxygen species, drug transport, cell wall modification, and DNA damage repair. On the contrary, genes involved in protein synthesis and secretion, such as ribosomal protein and αCOP-like protein, were significantly repressed. In particular, the expression of quite a few genes with unknown functions was apparently decreased in the presence of menadione. Another notable feature of the response to menadione was the differential expression of

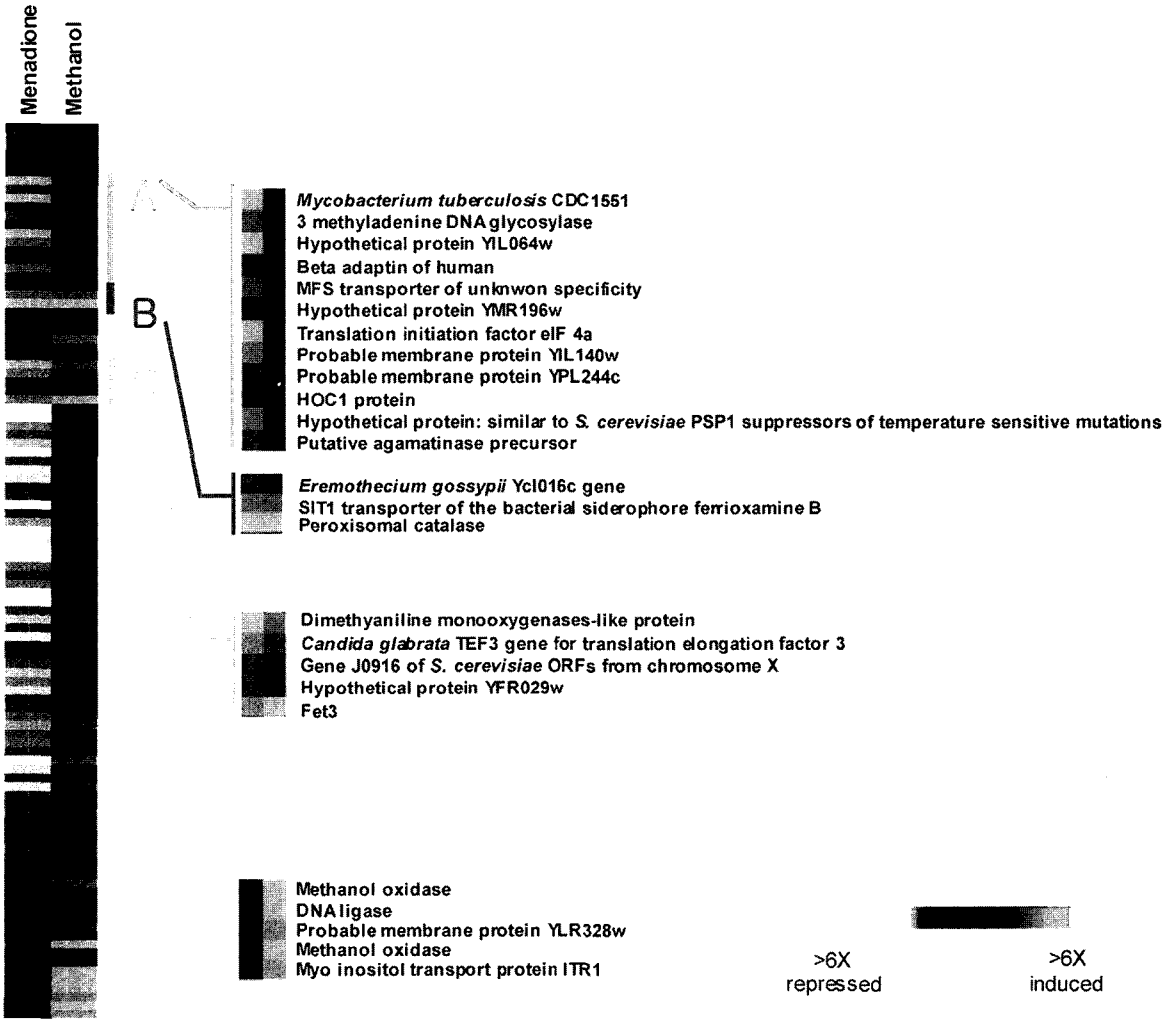

Fig. 4. Cluster analysis of gene expression in the presence of methanol and menadione, respectively.
The ORFs with the change of expression levels by at least two-fold in the presence of methanol or menadione were hierarchically clustered, based on their expression patterns. Some representative genes grouped into A, B, C, and D are listed.

two putative genes coding for eIF 4A proteins (TIF1 and TIF2), which are 78% identical at the amino acid sequence level. These two ORFs with 43% identity at the nucleic acid sequence level were distinguishable by the array hybridization conditions used in this study. This result is in good agreement with previous observations that most isozymes are differentially expressed in response to environmental shock [11].

Compared to the shift of carbon source from glucose to methanol, the menadione treatment appeared to cause more significant change in the expression profile. The ORFs whose expression levels changed significantly by at least two-fold in the presence of methanol or menadione were hierarchically clustered, based on their expression patterns (Fig. 4). Some genes were induced or repressed in response to both reagents, methanol and menadione (groups B and C). There are other genes that responded in a more specific way, upregulated only by menadione (group A) or by methanol (group D) treatments. Interestingly, our analysis of the microarray showed a different expression pattern of *CAT* and *MOX* genes. In contrast to the expression of *CAT*, the expression of *MOX* was induced only in the presence of methanol, but not in the menadione treatment. This confirms earlier findings that *MOX* and *CAT* are not identically regulated, although both promoters have a common USA element [12]. These results, taken together, demonstrated that the partial *H. polymorpha* cDNA microarray constructed in this study is sensitive and reliable for differential gene expression profiling.

DNA microarrays provide "global views" of several biological processes, which facilitates comprehensive description of the metabolic and physiological state of an entire organism under a given condition. The present study showed that interesting information on the expression pattern of several novel genes could be obtained by using even the partial genome cDNA array containing only 382 ORFs. Following completion of whole *H. polymorpha* genome sequencing [25], we have just finished constructing whole *H. polymorpha* genome microarrays by using the same strategy presented in this study. It is expected that the whole *H. polymorpha* genome microarray will provide a more powerful tool for a genome-wide transcriptional profiling under defined genetic and physiological conditions, thus generating invaluable information for pathway engineering and process optimization in exploiting *H. polymorpha* as a cell factory.

Acknowledgments

This study was supported by grants awarded to H. A. Kang from the Korean Ministry of Science and Technology (Microbial Genomics and Applications R& D Program; KRIBB Research Initiative Program; Korea-Ukraine International Cooperation Program).

References

1. Avila, J., C. González, N. Brito, and J. M. Siverio. 1998. Clustering of the *YNA1* gene encoding a $Zn(II)_2Cys_6$ transcriptional factor in the yeast *Hansenula polymorpha* with the nitrate assimilation genes *YNT1*, *YNI1* and *YNR1*, and its involvement in their transcriptional activation. *Biochem. J.* **335:** 647-652.
2. Bilban, M., L. K. Buehler, S. Head, G. Desoye, and V. Quaranta. 2002. Normalizing DNA microarray data. *Curr. Issues Mol. Biol.* **4:** 57-64.
3. Blandin, G., B. Llorente, A. Malpertuy, P. Wincker, F. Artiguenave, and B. Dujon. 2000. Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia angusta*. *FEBS Lett.* **487:** 76-81.
4. Bystrykh, L. V., L. R. Aminova, and Y. A. Trotsenko. 1988. Methanol metabolism in mutants of the methylotrophic yeast *Hansenula polymorpha*. *FEMS Microbiol. Lett.* **51:** 89-94.
5. Chen, D., W. M. Toone, J. Mata, R. Lyne, G. Burns, K. Kivinen, A. Brazma, N. Jone, and J. Bähler. 2003. Global transcriptional responses of fission yeast to environmental stress. *Mol. Biol. Cell* **14:** 214-229.
6. Cowen, L., A. Nantel, M. S. Whiteway, D. Y. Thomas, D. C. Tessiler, L. M. Kohn, and J. B. Anderson. 2002. Population genomics of drug resistance in *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **99:** 9284-9289.
7. DeRisi, J. L., V. R. Iyer, and P. O. Brown. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278:** 680-686.
8. Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95:** 14863-14868.
9. Elion, E. A. and J. R. Warner. 1984. The major promoter element of rRNA transcription in yeast lies 2 kb upstream. *Cell* **39:** 663-673.
10. Erasmus, D. J., G. K. van der Merwe, and H. J. van Vuuren. 2003. Genome-wide expression analyses: Metabolic adaptation of *Saccharomyces cerevisiae* to high sugar stress. *FEMS Yeast Res.* **3:** 375-399.
11. Gasch A. P., P. T. Spellman, C. M. Kao, C. H. Orna, M. B. Eisen, G. Storz, D. Botstein, and P. O. Brown. 2000. Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11:** 4241-4257.
12. Gellissen G., C. P. Hollenberg, and Z. A. Janowicz. 1995. Gene expression in methylotrophic yeasts, pp. 195-239. *In* Smith, A. (ed.). *Gene Expression in Recombinant Microorganisms*. New York, Dekker, U.S.A.
13. Gellissen, G. 2002. *Hansenula polymorpha - Biology and Applications*. Wiley-VCH, Weinheim.
14. Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver. 1996. Life with 6000 genes. *Science* **274:** 546-567.
15. Hegde, P., R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N. Lee, and J. Quackenbush. 2000.

A concise guide to cDNA microarray analysis. *Biotechniques* **29:** 548-550, 552-554, 556-562.
16. Holm, C., D. W. Meeks-Wagner, W. L. Fangman, and D. Botstein. 1986. A rapid, efficient method for isolating DNA from yeast. *Gene* **42:**169-173.
17. Jeoung, D., J.-W. Kim, Y.-H. Lee, M. Baek, S. Lee, N.-I. Baek, and H. Y. Kim. 2002. cDNA microarray analysis of transcriptional response to hyperin in human gastric cancer cells. *J. Microbiol. Biotechnol.* **12:** 664-668.
18. Kang, H. A., J. H. Sohn, E. S. Choi, B. H. Chung, M. H. Yu, and S. K. Rhee. 1998. Glycosylation of human α_1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts. *Yeast* **14:** 371-381.
19. Kim, K. A., J. Y. Lee, K. S. Park, M. J. Kim, and J. H. Chung. 1996. Mechanism of menadione-induced cytotoxicity in rat platelets. *Toxicol. Appl. Pharmacol.* **138:** 12-19.
20. Kim, S. Y., J. H. Sohn, Y. R. Pyun, and E. S. Choi. 2002. A cell surface display system using novel GPI-anchored proteins in *Hansenula polymorpha*. *Yeast* **19:** 1153-1163.
21. Kim, T. W., S. G. Min, D. H. Choi, J. S. Jo, and H. Y. Kim. 2000. Rapid identification of *Lactobacillus plantarum* in Kimchi using polymerase chain reaction. *J. Microbiol. Biotechnol.* **10:** 881-884
22. Lyne, R., G. Burns, J. Mata, C. J. Penkett, G. Rustici, D. Chen, C. Langford, D. Vetrie, and J. Bähler. 2003. Whole-genome microarrays of fission yeast: Characteristics, accuracy, reproducibility, and processing of array data. *BMC Genomics* **4:** 27.
23. Murad, A. M. A., C. dEnfert, C. Gaillardin, H. Tournu, F. Tekaia, D. Talibi, D. Marechal, V. Marchais, J. Cottin, and A. J. P. Brown. 2001. Transcript profiling in *Candida albicans* reveals new cellular functions for the transcriptional repressors CaTup1, CaMig1 and CaNrg1. *Mol. Microbiol.* **42:** 981-993.
24. Perez-Ortin, J. E., J. Garcia-Martinez, and T. M. Alberola. 2002. DNA chips for yeast biotechnology. The case of wine yeasts. *J. Biotechnol.* **98:** 227-241.
25. Ramezani-Rad, M., C. P. Hollenberg, J. Lauber, H. Wedler, E. Griess, C. Wagner, K. Albermann, J. Hani, M. Piontek, U. Dahlems, and G. Gellissen. 2003. The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res.* **4:** 207-215.
26. Roberts, C. J., B. Nelson, M. J. Marton, R. Stoughton, M. R. Meyer, H. A. Bennett, Y. D. He, H. Dai, W. L. Walker, T. R. Hughes, M. Tyers, C. Boone, and S. H. Friend. 2000. Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* **287:** 873-880.
27. Stare, J., J. Stare, F. E. Jr. Harrell, and H. Heinzl. 2001. BJ: An S-Plus program to fit linear regression models to censored data using the Buckley-James method. *Computer Methods Programs Biomed.* **64:** 45-52.
28. Titorenko, V. I., H. R. Waterham, J. M. Cregg, W. Harder, and M. Veenhuis. 1993. A complex set of interacting genes controls peroxisome biogenesis in *Hansenula polymorpha*. *Proc. Natl. Acad. Sci. USA* **90:** 7470-7474.
29. van Dijk, R., K. N. Faber, J. A. K. W. Kiel, M. Veenhuis, and I. van der Klei. 2000. The methylotrophic yeast *Hansenula polymorpha*: A versatile cell factory. *Enzyme Microbiol. Technol.* **26:** 793-800.
30. Veale, R. A., M. L. Giuseppin, H. M. van Eijk, P. E. Sudbery, and C. T. Verrips. 1992. Development of a strain of *Hansenula polymorpha* for the efficient expression of guar alpha-galactosidase. *Yeast* **8:** 361-372.
31. Yang, Y. H., S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngal, and T. P. Speed. 2002. Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30:** e15.
32. Yurimoto, H., T. Komeda, C. R. Lim, T. Nakagawa, K. Kondo, N. Kato, and Y. Sakai. 2000. Regulation and evaluation of five methanol-inducible promoters in the methylotrophic yeast *Candida boidii*. *Biochim. Biophys. Acta* **1493:** 56-63.
33. Zammatteo N., L. Jeanmart, S. Hamels, S. Courtois, P. Louette, L. Hevesi, and J. Remacle. 2000. Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. *Anal. Biochem.* **280:** 143-150.