

# Global response of *Saccharomyces cerevisiae* to an alkylating agent

SCOTT A. JELINSKY AND LEONA D. SAMSON\*

Department of Cancer Cell Biology, Division of Toxicology, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115

Edited by David Botstein, Stanford University School of Medicine, Stanford, CA, and approved December 7, 1998 (received for review September 16, 1998)

**ABSTRACT** DNA chip technology enables simultaneous examination of how  $\approx 6,200$  *Saccharomyces cerevisiae* gene transcript levels, representing the entire genome, respond to environmental change. By using chips bearing oligonucleotide arrays, we show that, after exposure to the alkylating agent methyl methanesulfonate,  $\approx 325$  gene transcript levels are increased and  $\approx 76$  are decreased. Of the 21 genes that already were known to be induced by a DNA-damaging agent, 18 can be scored as inducible in this data set, and surprisingly, most of the newly identified inducible genes are induced even more strongly than these 18. We examined 42 responsive and 8 nonresponsive ORFs by conventional Northern blotting, and 48 of these 50 ORFs responded as they did by DNA chip analysis, with magnitudes displaying a correlation coefficient of 0.79. Responsive genes fall into several expected and many unexpected categories. Evidence for the induction of a program to eliminate and replace alkylated proteins is presented.

Exposure to DNA-damaging agents can increase DNA repair capacity and activate cell-cycle checkpoints. Such exposures may also induce enzymes that metabolize toxicants to facilitate their elimination from the organism or may activate programmed cell death (apoptosis) to eliminate highly damaged cells. Thus, it has long been known that cells induce the expression of a variety of genes after toxic exposure, and gene regulation in response to DNA-damaging agents has been well studied in many organisms (1–5).

Random *lacZ* gene fusions and differential hybridization previously have identified 21 *Saccharomyces cerevisiae* genes whose transcript levels are increased in response to DNA-damaging agents (1, 6–8). Both approaches produced catalogs of genes of known and unknown function, but the lack of redundancy with which they were identified indicates that the search for such inducible genes is far from complete (1, 8).

We previously studied the inducible transcription of an *S. cerevisiae* DNA repair gene (*MAG1*, encoding a 3-methyladenine DNA glycosylase) in response to simple alkylating agents such as methyl methanesulfonate (MMS; refs. 9–13). Upstream *MAG1* regulatory elements were identified, and similar elements are found upstream of numerous DNA repair and metabolism genes, suggesting common transcriptional regulatory mechanisms (12–14). We therefore decided to identify all the genes that are regulated coordinately with *MAG1*. Here, we report that DNA chip analysis has expanded by more than 15-fold the catalog of genes that are inducible by a DNA-damaging agent. In addition, DNA chip analysis has identified a class of genes whose transcripts are repressed. Global responses to a DNA-damaging agent have now come into focus, and we present evidence that exposure to an alkylating agent elicits a program to eliminate and replace alkylated proteins from *S. cerevisiae* cells.

## MATERIALS AND METHODS

**Strains, Media, and Growth Conditions.** *S. cerevisiae* strain DBY747 (*MATa his3- $\Delta$ 1 leu2-3,112 ura3-52 trp1-289a gal<sup>S</sup> can1*

*CUP<sup>r</sup>*) was used in this study and was grown in 1% yeast extract/2% peptone/2% glucose at 30°C. Cells were grown to a density of  $5 \times 10^6$  cells per ml as measured by counting duplicated dilutions. Cultures were split into two; MMS (0.1%) was added directly to one culture, and both cultures were incubated at 30°C for 1 h. Cells were pelleted and washed once in distilled H<sub>2</sub>O and once in AE buffer (50 mM NaOAc, pH 5.2/10 mM EDTA) immediately before RNA extraction.

**RNA Extraction.** Total RNA was isolated by using a hot-phenol method (15). Poly(A)<sup>+</sup> RNA was purified from total RNA with Oligotex oligo(dT) selection step (Qiagen, Chatsworth, CA). Poly(A)<sup>+</sup> RNA was amplified and biotin-labeled as follows. Poly(A)<sup>+</sup> RNA (1  $\mu$ g) was converted into double-stranded cDNA by using a modified oligo(dT) primer with a T7 RNA polymerase promoter sequence at the 5' end and the Superscript Choice system for cDNA synthesis (GIBCO). Double-stranded cDNA was purified by phenol/chloroform extractions, precipitated with ethanol, and resuspended at a concentration of 0.5  $\mu$ g/ $\mu$ l in diethyl pyrocarbonate-treated H<sub>2</sub>O. Phase-Lock Gel (5 Prime  $\rightarrow$  3 Prime) was used for all organic extractions to increase recovery. *In vitro* transcription was performed with T7 RNA polymerase (T7 Megascript kit, Ambion, Austin, TX) and with 0.5–1.0  $\mu$ g of cDNA, 7.5 mM unlabeled ATP and GTP, 5.3 mM unlabeled UTP and CTP, and 1.9 mM biotin-labeled CTP and UTP (biotin-11-CTP, biotin-16-UTP, Enzo Diagnostics). Reactions were carried out for 6 h at 37°C, and cRNA was purified by RNA affinity resin (RNeasy spin columns, Qiagen). A sample was separated on a 1% agarose gel to check the size range, and then 10  $\mu$ g of cRNA was fragmented randomly to an average size of 50 bases by heating at 94°C for 35 min in 40 mM Tris·acetate, pH 8.1/100 mM KOAc/30 mM MgOAc.

**GeneChip Hybridizations.** A set of four oligonucleotide arrays (GeneChip Ye6100 arrays, Affymetrix, Santa Clara, CA) containing probes for 6,218 yeast ORFs were used for hybridizations. Hybridizations were done in 200  $\mu$ l of AFFY buffer (Affymetrix) at 40°C for 16 h with constant mixing. After hybridization, arrays were rinsed three times with 200  $\mu$ l of 6 $\times$  sodium chloride/sodium phosphate/EDTA/Triton (SSPE-T; 1 $\times$  0.15 M NaCl/15 mM phosphate, pH 7.6/1 mM EDTA/0.005% Triton) and then washed with 200  $\mu$ l of 6 $\times$  SSPE-T (pH 7.6) for 1 h at 50°C. The arrays were rinsed twice with 0.5 $\times$  SSPE-T (pH 7.6) and washed with 0.5 $\times$  SSPE-T (pH 7.6) at 50°C for 15 min. Staining was done with 2  $\mu$ g/ml streptavidin-phycoerythrin (Molecular Probes) and 1 mg/ml acetylated BSA (Sigma) in 6 $\times$  SSPE-T (pH 7.6). The arrays were read at 7.5  $\mu$ m with a confocal scanner (Molecular Dynamics) and analyzed with GENECHIP software, version 3.0. The samples were normalized by using the total average difference between the perfectly matched probe and the mismatched probe (16).

**Northern-Blot Analysis.** RNA was isolated from log-phase cells exposed to 0.1% MMS for 0, 30, 60, or 120 min by using a hot-phenol extraction method (15). Total RNA (25  $\mu$ g) was fractionated in a 1% denaturing agarose gel, blotted, and probed

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations: MMS, methyl methanesulfonate; SSPE-T, sodium chloride/sodium phosphate/EDTA/Triton.

\*To whom reprint requests should be addressed. e-mail: lsamson@sph.harvard.edu.

with PCR-amplified labeled ORFs (Research Genetics, Huntsville, AL) by using standard methods (17).

## RESULTS AND DISCUSSION

**Global Expression Monitoring After Alkylation Damage.** The GeneChip methodology developed by Affymetrix was used to monitor global gene expression in *S. cerevisiae*. The 6,218 ORFs of this organism are represented on a set of four high density oligonucleotide arrays (16, 18, 19). Poly(A)<sup>+</sup> mRNA was isolated from untreated cells and from cells exposed for 1 h to 0.1% MMS. These conditions were chosen, because they yield optimal *MAG1* induction with minimal cell death (11). Poly(A)<sup>+</sup> RNA was converted into double-stranded cDNA containing the T7 RNA polymerase promoter, and biotin-labeled cRNA was produced and hybridized to the GeneChip arrays. The hybridization-intensity information was gathered by scanning confocal microscopy and analyzed with GENECHIP software, version 3.0 (16). Typical GeneChip-hybridization intensities for control and MMS-treated cells are shown in Fig. 1. As a guide, one MMS-induced, one MMS-repressed, and one nonresponsive ORF are indicated by arrows. It had been established that differences in hybridization intensity between the same ORFs on corresponding chips are proportional to changes in transcript levels and that intensity changes greater than 2.0-fold are both significant and accurate (16). It is important to note that 18 of the 21 genes previously reported to be induced by a DNA-damaging agent (not necessarily MMS) were found to be induced by 2.5- to 6.9-fold. This fact is remarkable, because our study is limited to a single agent (MMS) and a single time point. These genes (and their fold induction) are *UBI4* (10.3), *RNR3* (6.9), *DDR48* (6.2), *RNR1* ( $\approx 5.7$ ), *MAG1* ( $\approx 5.5$ ), *RAD7* ( $\approx 5.5$ ), *DDR2* (5.3), *HIS4* (4.6), *RNR2* (3.9), *HIS3* (3.4), *CDC8* (2.9), *RAD2* (2.9), *RAD54* ( $\approx 2.8$ ), *RAD23* (2.7), *PHR1* (2.6), *CDC9* (2.5), *RAD51* (2.5), *RNR4* (2.5), *CDC17* (1), *RAD6* (1), and *RAD18* (1). (The complete data set can be found at [www.hsph.harvard.edu/geneexpression](http://www.hsph.harvard.edu/geneexpression).)

Of 6,218 ORFs, 325 (5.2%) showed more than a 4-fold increase in transcript level (Table 1); 32 of these increased by greater than 10-fold, and the greatest increase was 251-fold (YFL061W). For reference, *MAG1* was induced  $\approx 5.5$ -fold in this particular experiment, and 115 ORFs were induced more highly than *MAG1*. MMS treatment also resulted in more than a 3-fold reduction in transcript levels for 75 of 6,218 ORFs (Table 2). Of these, 24

decreased by a factor of  $>5$ -fold; 10 decreased by a factor of  $>7.5$ -fold; 4 decreased by  $>10$ -fold.

To establish that the information obtained from GeneChip analysis is accurate, we chose 50 ORFs to examine by conventional Northern-blot analysis; these included 26 inducible (3.1- to 251-fold), 16 repressible (3.1- to 18.1-fold), and 8 nonresponsive ORFs. Freshly isolated RNA from control and from MMS-treated cells was probed with each of the 50 ORFs. Hybridization signals, as measured by Northern-blot and GeneChip analysis, are shown for six of the ORFs (Fig. 2), and data for all 50 ORFs are compiled in Fig. 3A and B. In terms of whether the ORFs were inducible, repressible, or nonresponsive, the Northern-blot and GeneChip data agreed for 48 of the 50 ORFs. Moreover, the data agreed remarkably well in terms of the extent of induction or repression (Fig. 3A and B), displaying a correlation coefficient of 0.79 for the complete data set (Fig. 3B). For the majority of the ORFs, the fold change in transcript levels differed by no more than a factor of two between the Northern-blot and GeneChip data (Fig. 3A). However, for very highly induced ORFs (see Fig. 3B), the correlation weakens slightly, such that the Northern-blot analysis overestimates induction, the GeneChip analysis underestimates induction, or both. Note that for some highly induced ORFs, basal transcript levels are undetectable, making it difficult to calculate accurate fold-induction values. Further, some quantitative differences between GeneChip and Northern-blot data may not be unexpected given that the GeneChip analysis monitors hybridization of fragmented cRNAs to 20 oligonucleotide pairs per ORF (chosen for their uniqueness relative to the entire *S. cerevisiae* genome) and given that Northern blotting monitors hybridization of a complete ORF to immobilized full-length RNA transcripts separated on the basis of size. GeneChip analysis has the potential to discriminate between closely related genes, whereas Northern-blot analysis suffers from the potential that an ORF probe might hybridize to closely related RNAs. Indeed, the data point marked with an asterisk in Fig. 3A highlights this problem for the *PHO3* and *PHO5* genes. By GeneChip analysis, *PHO3* was repressed 13-fold; *PHO5* was scored as nonresponsive, because it was undetectable in both control and MMS-treated cells. One plausible explanation for *PHO5* being undetectable by GeneChip analysis is that weak hybridization is caused by incorrect primer design. However, by using genomic DNA as a probe, the *PHO5* gene can be detected on these chips (16). Because

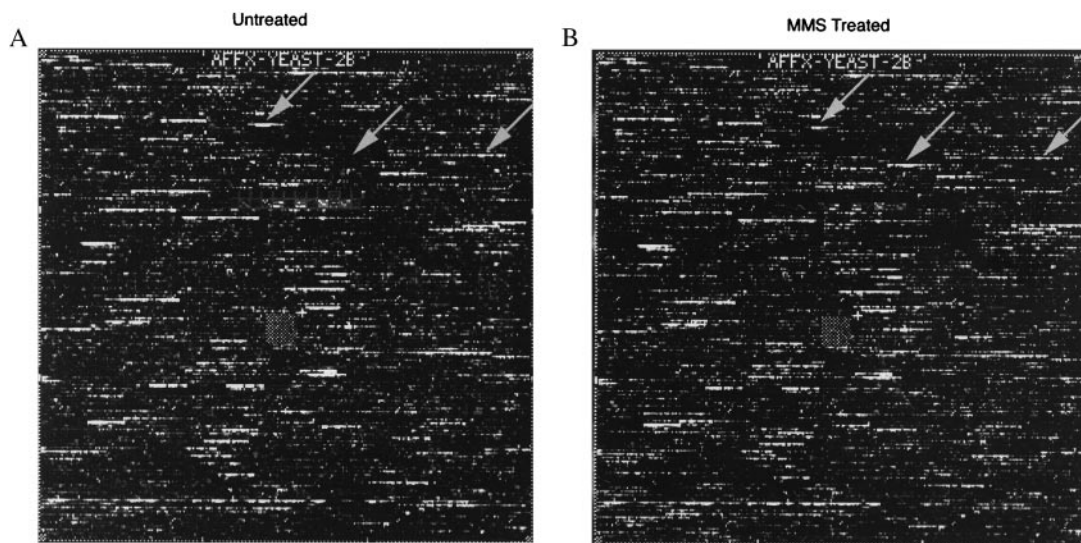


FIG. 1. Fluorescence image of *S. cerevisiae* cRNA hybridization to GeneChip oligonucleotide arrays corresponding to ORFs from YEL001c to YJL088w (chromosomes 5–10; ref. 30) probed with 10  $\mu$ g of biotin-labeled cRNA prepared from *S. cerevisiae* DBY747 log-phase cells untreated (A) or treated (B) with 0.1% MMS for 1 h. At this dose and time, there was a  $>90\%$  survival rate. Differential hybridization between A and B represents ORFs expressed at different levels before and after MMS exposure. Examples of repressed (left arrow), induced (middle arrow), and nonresponsive (right arrow) ORFs are indicated. These ORFs correspond to *RPS26B*, YFL061W, and *ACT1* respectively, and the hybridization differences were 3.6-fold, 251-fold, and 1-fold, respectively.

Table 1. ORFs whose transcripts are induced by >4-fold by MMS ( $n = 325^*$ )

ORF	Gene	Fold	Function <sup>†</sup>	ORF	Gene	Fold	Function <sup>†</sup>
Stress response/detoxification (25 ORFs <sup>‡</sup> )				YFL030W <sup>§</sup>	—	5.0	Transaminases <sup>  </sup>
YLL060C	<i>GTT2</i>	28.8	Glutathione transferase	YGL254W <sup>††</sup>	<i>FZF1</i>	4.9	Sulfite resistance protein
YER143W	<i>DDI1</i>	17.8	DNA damage-inducible	YJL060W	—	4.9	Glutamine transaminase <sup>  </sup>
YPL092W	<i>SSU1</i>	17.8	Sulfite sensitivity	YOR226C	—	4.6	Nitrogen fixation proteins <sup>  </sup>
YBR008C <sup>§¶</sup>	<i>FLR1</i>	15.1	Fluconazole resistance	YDR242W	<i>AMD2</i>	~4.5	Amidase
YLL039C <sup>§¶</sup>	<i>UBI4</i>	10.3	Ubiquitin	Carbohydrate metabolism/fermentation (28 ORFs)			
YBL064C	—	8.5	Antioxidant enzyme <sup>  </sup>	YFL057C	—	14.1	Aryl-alcohol dehydrogenases <sup>  </sup>
YML116W <sup>§¶¶</sup>	<i>ATRI</i>	8.0	Aminotriazole and 4-NQO <sup>R</sup>	YDL243C	—	~13.9	Aryl-alcohol dehydrogenases <sup>  </sup>
YDR059C <sup>§</sup>	<i>UBC5</i>	7.3	E2 ubiquitin-conjugating enzyme	YFL056C	—	13.7	Aryl-alcohol dehydrogenases <sup>  </sup>
YAL005C	<i>SSA1</i>	6.4	HSP70 family, cytoplasmic	YNL241C <sup>§¶</sup>	<i>ZWF1</i>	6.0	Glucose-6-phosphate dehydrogenase
YDL025C <sup>§</sup>	—	6.3	Ser/Thr protein kinase	YAL060W	<i>FUN49</i>	5.8	Alcohol/sorbitol dehydrogenase <sup>  </sup>
YML007W <sup>§¶¶</sup>	<i>YAP1</i>	6.2	Transcriptional activator	YOR120W	<i>GCY1</i>	5.8	Aldo/keto reductase
YMR173W <sup>¶¶</sup>	<i>DDR48</i>	6.2	DNA damage-inducible HSP	YEL020C	—	~5.7	Oxalyl-CoA decarboxylase <sup>  </sup>
YNL241C <sup>§¶¶</sup>	<i>ZWF1</i>	6.0	Glucose-6-phosphate dehydrogenase	YGL062W	<i>PYC1</i>	5.6	Pyruvate carboxylase 1
YOR162C <sup>§¶¶</sup>	<i>YRR1</i>	~5.7**	Transcription factor	YDL099W	<i>CHS6</i>	~5.6	Chitin biosynthesis protein
YML070W <sup>§</sup>	<i>DAK1</i>	5.4	Dihydroxyacetone kinase	YML070W <sup>§</sup>	<i>DAK1</i>	5.4	Dihydroxyacetone kinase
YOL052C-A	<i>DDR2</i>	5.3	Heat-shock protein	YNL331C	—	5.3	Aryl-alcohol reductase <sup>  </sup>
YGR234W <sup>¶</sup>	<i>YHB1</i>	5.2	Flavohemoglobin	YGR244C	<i>LSC2</i>	5.2	Succinate-CoA ligase
YMR174C	<i>PAI3</i>	5.1	Protease A (ysca) inhibitor IA3	YDL066W	<i>IDP1</i>	4.9	Isocitrate dehydrogenase (NADP <sup>+</sup> )
YGL254W <sup>‡</sup>	<i>FZF1</i>	4.9	Sulfite resistance protein	YIR036C	—	4.9	<i>E. coli</i> FabD <sup>  </sup>
YOL025W <sup>§</sup>	<i>LAG2</i>	~4.8	Affects longevity	YDR001C	<i>NTH1</i>	4.8	Neutral trehalase ( $\alpha,\alpha$ -trehalase)
YIR038C	<i>GTT1</i>	4.6	Glutathione transferase	YGL104C <sup>§</sup>	—	4.7	Glucose transport proteins <sup>  </sup>
DNA synthesis/repair (13 ORFs)				YLR164W	—	4.7	Sdh4p <sup>  </sup>
YAR007C <sup>¶</sup>	<i>RFA1</i>	9.5	DNA replication factor A	YPR184W	—	~4.7	Human 4- $\alpha$ -glucanotransferase <sup>  </sup>
YAL015C <sup>¶</sup>	<i>NTG1</i>	~7.6	DNA glycosylase	mRNA processing (15 ORFs)			
YIL066C	<i>RNR3</i>	6.9	Ribonucleotide reductase	YLR136C	<i>TIS11</i>	6.3	Homolog of mammalian TIS11
YER070W	<i>RNR1</i>	~5.7	Ribonucleotide reductase	YML007W <sup>§¶¶</sup>	<i>YAP1</i>	6.2	Transcriptional activator
YKL112W <sup>‡‡</sup>	<i>ABF1</i>	~5.6	ARS1 binding protein	YKL070W	—	~6.1	Transcriptional regulatory
YER142C	<i>MAG1</i>	~5.5	3-Methyladenine DNA glycosylase	YKL112W <sup>‡‡</sup>	<i>ABF1</i>	~5.6	Transcriptional activator
YJR052W <sup>¶</sup>	<i>RAD7</i>	~5.5	Nucleotide excision repair protein	YML112W <sup>§</sup>	<i>CTK3</i>	5.5	Carboxy terminal domain kinase
YMR228W <sup>§</sup>	<i>MTF1</i>	5.2	RNA polymerase specific factor	YGL122C	<i>NAB2</i>	5.4	Nuclear poly(A)-binding protein
YIR008C <sup>¶</sup>	<i>PRI1</i>	4.7	DNA primase	YMR228W <sup>§</sup>	<i>MTF1</i>	5.2	RNA polymerase specific factor
Cell cycle (10 ORFs)				YGL254W <sup>††</sup>	<i>FZF1</i>	4.9	Sulfite resistance protein
YLR299W <sup>††</sup>	<i>CIS2</i>	7.7	$\gamma$ -Glutamyltransferase	YOR185C	<i>GSP2</i>	4.7	GTP-binding protein
YDL132W <sup>§</sup>	<i>CDC53</i>	6.1	Controls G <sub>1</sub> /S transition	Others (57 ORFs)			
YLR178C <sup>§</sup>	<i>TFS1</i>	5.7	Cell-cycle regulator	YMR096W	<i>SNZ1</i>	~65.8	Stationary phase protein
YMR028W	<i>TAP42</i>	5.2	Involved in Tor signaling	YCL026C	<i>FRM2</i>	~27.5	Fatty acid regulation
YKL179C	—	4.5	Kinesin <sup>  </sup>	YML131W	—	15.2	Leukotriene B <sub>4</sub> 12-hydroxydehydrogenase <sup>  </sup>
Signaling/kinases/phosphatases (16 ORFs)				YIL164C	<i>NIT1</i>	12.8	Nitrilase
YPL150W	—	7.2	Ser/Thr protein kinases <sup>  </sup>	YPL171C	<i>OYE3</i>	~10.4	NAPDH dehydrogenase
YDL025C <sup>§</sup>	—	6.3	Ser/Thr protein kinase	YNL335W	—	10.3	Cyanamide hydratase <sup>  </sup>
YOR162C <sup>§¶¶</sup>	<i>YRR1</i>	~5.7	Transcription factor	YLR214W	<i>FRE1</i>	8.9	Ferric (and cupric) reductase
YLR178C <sup>§</sup>	<i>TFS1</i>	5.7	Cell-cycle regulator	YBR256C	<i>RIB5</i>	6.4	Riboflavin synthase a-chain
YML112W <sup>§</sup>	<i>CTK3</i>	5.5	Carboxy-terminal domain kinase	YNL036W	<i>NCE10</i>	6.4	Protein export pathway
YNL183C <sup>§</sup>	<i>NPR1</i>	5.4	Ser/Thr protein kinase	YBR046C	<i>ZTA1</i>	6.2	$\zeta$ -crystallin homolog
YPL152W	<i>RRD2</i>	5.1	Phosphotyrosyl phosphatase <sup>  </sup>	YBR170C	<i>NPL4</i>	6.2	Nuclear protein localization
YGR080W	<i>TWF1</i>	5.0	Twinfilin A	YHR071W	<i>PCL5</i>	5.9	PHO85 cyclin
YLR362W <sup>¶</sup>	<i>STE11</i>	5.0	Ser/Thr protein kinase	YJL068C	—	5.8	Human esterase D <sup>  </sup>
YJL164C	<i>SR43</i>	4.7	cAMP-dependent protein kinase 1	YMR231W	<i>PEP5</i>	~5.8	Vacuolar biogenesis protein
YLL019C	<i>KNS1</i>	~4.5	Ser/Thr protein kinase	YFR010W	<i>UBP6</i>	5.6	Ubiquitin-specific protease
YOL100W	<i>PKH2</i>	4.5	Ser/Thr protein kinase	YGL194C	<i>HOS2</i>	~5.4	Putative histone deacetylase
Cell wall biogenesis (5 ORFs)				YGR218W	<i>CRM1</i>	5.4	Chromosome region maintenance
YKR076W	<i>ECM4</i>	17.7	Cell wall biogenesis	YKL073W	<i>LHS1</i>	~5.3	Chaperone of the ER lumen
YHL030W	<i>ECM29</i>	~8.7	Cell wall biogenesis	YBL033C	<i>RIB1</i>	5.2	GTP cyclohydrolase II
YLR299W <sup>‡</sup>	<i>CIS2</i>	7.7	$\gamma$ -Glutamyltransferase	YHR016C	<i>YSC84</i>	5.2	Hypothetical protein YFR024ca <sup>  </sup>
YKR009C	<i>FOX2</i>	6.5	Hydratase-dehydrogenase-epimerase	YLL063C	<i>AYT1</i>	~5.0	Transacetylase
YOL025W <sup>§</sup>	<i>LAG2</i>	~4.8	Affects longevity	YOR227W	—	5.0	Microtubule-interacting protein
Membrane transport (13 ORFs)				YJL008C	<i>CCT8</i>	4.9	Chaperonin-containing T complex
YBR008C <sup>§¶¶</sup>	<i>FLR1</i>	15.1	Fluconazole resistance protein	YMR004W	<i>MVP1</i>	4.9	Sorting proteins to the vacuole
YOR328W	<i>PDR10</i>	8.8	ABC transporter proteins	YJL041W	<i>NSP1</i>	4.8	Nuclear pore protein
YML116W <sup>§¶¶</sup>	<i>ATRI</i>	8.0	Aminotriazole and 4-NQO <sup>R</sup>	YLL001W	<i>DNM1</i>	4.8	Dynamin-related protein
YOL119C	—	7.6	Monocarboxylate transporters <sup>  </sup>	YNL237W	<i>YTP1</i>	4.8	Mitochondrial electron transport proteins <sup>  </sup>
YOR130C <sup>§</sup>	<i>ARG11</i>	5.8	Integral membrane protein	YCR029C	—	4.7	Hypothetical protein
YMR060C	<i>TOM37</i>	5.6	Outer membrane import receptor	YJL154C	<i>VPS35</i>	4.7	Protein-sorting protein, vacuolar
YCR011C	<i>ADP1</i>	5.5	ATP-dependent permease	YLR163C	<i>MAS1</i>	~4.7	Mitochondrial processing peptidase
YGL104C <sup>§</sup>	—	4.7	Glucose transport proteins <sup>  </sup>	YOR181W	<i>LAS17</i>	~4.7	Actin assembly factor
YGL186C	—	4.5	Purine cytosine permease <sup>  </sup>	YKL173W	<i>SNU11</i>	4.6	U5 snRNP-specific protein
Nitrogen and sulfur metabolism (12 ORFs)				YOR069W	<i>VPS5</i>	4.6	Sorting nexin I homolog
YFL061W	—	251	Cyanamide hydratase <sup>  </sup>	Unknown/unclassified (112 ORFs)			
YOL058W <sup>§</sup>	<i>ARG1</i>	7.0	Argininosuccinate synthetase	These ORFs and their fold induction can be found on the World Wide Web at <a href="http://www.hsph.harvard.edu/geneexpression">www.hsph.harvard.edu/geneexpression</a> .			
YFR030W <sup>§</sup>	<i>MET10</i>	5.8	Sulfite reductase, flavin-binding	Degradation (15 ORFs)			
YKL112W <sup>‡‡</sup>	<i>ABF1</i>	~5.6	ARS1 binding protein	YKL103C	<i>LAP4</i>	13.9	Aminopeptidase yscI precursor
YNL183C <sup>§</sup>	<i>NPR1</i>	5.4	Ser/Thr protein kinase	YLL039C <sup>§¶¶</sup>	<i>UBI4</i>	10.3	Ubiquitin
YDL215C <sup>§</sup>	<i>GDH2</i>	5.2	Glutamate dehydrogenase				



Table 1. (Continued)

ORF	Gene	Fold	Function†
YDR059C§	<i>UBC5</i>	7.3	E2 ubiquitin-conjugating enzyme
YDR092W	<i>UBC13</i>	6.2	E2 ubiquitin-conjugating enzyme
YDL132W§	<i>CDC53</i>	6.1	Controls G <sub>1</sub> /S transition
YMR304W	<i>UPB15</i>	5.7	Ubiquitin-specific protease <sup>  </sup>
YHR027C	<i>RPN1</i>	4.7	26S proteasome regulatory subunit
YOR124C	<i>UBP2</i>	4.6	Ubiquitin-specific protease
YJL001W	<i>PRE3</i>	4.5	20S proteasome subunit (b1)
Amino acid metabolism (41 ORFs)			
YMR189W	<i>GCV2</i>	12.5	Glycine decarboxylase complex
YJR109C	<i>CPA2</i>	12.3	Carbamoyl phosphate synthetase
YJR130C	—	11.5	o-Succinylhomoserine (thiol)-lyase
YKL218C	—	10.6	Threonine dehydratases
YMR062C	<i>ECM40</i>	9.8	Acetylornithine acetyltransferase
YER069W	<i>ARG5,6</i>	9.4	Acetylglutamate kinase
YJL088W	<i>ARG3</i>	9.3	Ornithine carbamoyltransferase
YLR160C	<i>ASP3D</i>	8.4	L-Asparaginase II
YHR018C	<i>ARG4</i>	7.8	Arginosuccinate lyase
YIL116W	<i>HIS5</i>	7.8	Histidinol-phosphate aminotransferase
YLR299W††	<i>CIS2</i>	7.7	γ-Glutamyltransferase
YDR019C	<i>GCV1</i>	7.5	Glycine decarboxylase
YLR158C	<i>ASP3C</i>	7.1	L-Asparaginase II
YOL058W§	<i>ARG1</i>	7.0	Argininosuccinate synthetase
YDR127W	<i>ARO1</i>	6.8	Arom pentafunctional enzyme
YKL215C	—	~6.7	<i>P. aeruginosa</i> hyuA and hyuB <sup>  </sup>
YFR055W	—	~6.6	β-Cystathionases <sup>  </sup>
YLL058W	—	6.4	o-Succinylhomoserine (thiol)-lyase
YIR034C	<i>LYS1</i>	6.2	Saccharopine dehydrogenase
YFR030W§	<i>MET10</i>	5.8	Sulfite reductase, flavin-binding
YHR037W	<i>PUT2</i>	5.8	1-Pyrroline-5-carboxylate dehydrogenase
YOR130C§	<i>ARG11</i>	5.8	Mitochondrial integral membrane protein
YKL112W‡‡	<i>ABF1</i>	~5.6	ARS1 binding protein
YNL104C	<i>LEU4</i>	5.6	2-Isopropylmalate synthase
YDL215C§	<i>GDH2</i>	5.2	Glutamate dehydrogenase
YER052C	<i>HOM3</i>	5.2	L-Aspartate 4-P-transferase
YKL211C	<i>TRP3</i>	5.1	Anthranilate synthase
YFL030W§	—	5.0	Transaminases <sup>  </sup>
YDR354W	<i>TRP4</i>	4.8	Anthranilate phosphoribosyltransferase
YDR035W	<i>ARO3</i>	4.7	2-Dehydro-3-deoxyphosphoheptonate
YER090W	<i>TRP2</i>	4.7	Anthranilate synthase component I
YHR137W	<i>ARO9</i>	4.7	Aromatic amino acid aminotransferase II
YJR025C	<i>BNA1</i>	4.7	3-Hydroxyanthranilic acid dioxygenase
YCL030C	<i>HIS4</i>	4.6	Histidinol dehydrogenase
YLR155C	<i>ASP3A</i>	4.6	L-Asparaginase II

\*Note that 35 ORFs fall into multiple categories.

†Categories derived from MIPS database (Munich Information Center for Protein Sequences) (31).

‡Only ORFs induced >4.5 are listed. The remaining can be found on the World Wide Web at [www.sph.harvard.edu/geneexpression](http://www.sph.harvard.edu/geneexpression).

§ORFs falling into two categories.

¶ORFs known to be involved in resistance to DNA-damaging agents (31).

||Predicted function based on sequence similarity.

\*\*Hybridization intensity in the untreated sample was below a certain threshold and was therefore increased to an arbitrary, low value for the purposes of this calculation. Values are therefore approximate.

††ORFs falling into three categories.

‡‡ORFs falling into four categories.

*PHO3* and *PHO5* have 87% nucleotide-sequence identity and are of similar size, *PHO5* scored as repressible (15-fold) by Northern-blot analysis, presumably because the *PHO5* probe hybridized to the *PHO3* transcript.

**Transcripts Induced by MMS.** Several types of defense mechanism would be expected to protect cells against an alkylating agent such as MMS, namely DNA repair and recombination, cell-cycle checkpoints, and pathways that somehow prevent alkylating agents from reacting with target molecules (e.g., by changing the cell wall, membrane permeability, or drug metabolism). The 325 ORFs that were induced greater than 4-fold by MMS are listed in Table 1; 4-fold was arbitrarily chosen as the cutoff and is more conservative than the cutoff (2-fold) recommended by Affymetrix. The first six groups of genes (stress

Table 2. ORFs whose transcripts are repressed by >3.0-fold by MMS ( $n = 76^*$ )

ORF	Gene	Fold	Function†
Ribonucleotide synthesis (7 ORFs)			
YBL039C	<i>URA7</i>	~11.6‡	CTP synthase 1
YNL141W	<i>AAH1</i>	~9.4	Adenosine deaminase
YDR399W	<i>HPT1</i>	8.5	Hypoxanthine-guanine phosphoribosyltransferase
YML056C	—	8.2	IMP dehydrogenases <sup>¶</sup>
YHR128W	<i>FUR1</i>	4.8	Uracil phosphoribosyltransferase
YKL216W	<i>URA1</i>	3.4	Dihydroorotate dehydrogenase
YMR217W	<i>GUA1</i>	3.4	GMP synthase
rRNA synthesis (11 ORFs)			
YNL112W	<i>DBP2</i>	18.1	ATP-dependent RNA helicase
YGR159C	<i>NSR1</i>	7.9	Nuclear localization sequence binding protein
YDL014W	<i>NOPI</i>	7.1	Fibrillarin
YGL078C	<i>DBP3</i>	4.5	ATP-dependent RNA helicase
YJR063W	<i>RPA12</i>	4.5	RNA polymerase I, 13.7 kDa
YPL211W	<i>NIP7</i>	4.4	60S ribosome subunit biogenesis
YHR089C	<i>GARI</i>	4.0	Nucleolar rRNA processing
YOR310C	<i>NOP5</i>	3.8	Nucleolar protein
YNL248C	<i>RPA49</i>	3.5	RNA polymerase A, 46 kDa
YLR197W	<i>SIK1</i>	3.5	Pre-rRNA processing
YNL113W	<i>RPC19</i>	3.1	RNA polymerase I, III, 16 kDa
Ribosomal proteins (13 ORFs)			
YEL026W§	—	6.2	HMG-like protein Nhp2p <sup>¶</sup>
YDL208W§	<i>NHP2</i>	4.7	HMG-like nuclear protein
YBR048W	<i>RPS11B</i>	4.4	Ribosomal protein S11B
YER131W	<i>RPS26B</i>	3.9	Ribosomal protein S26B
YKL009W	<i>MRT4</i>	3.8	Acidic ribosomal protein PO <sup>¶</sup>
YML026C	<i>RPS18B</i>	3.8	Ribosomal protein S18B
YNL301C	<i>RPL18B</i>	3.8	Ribosomal protein L18B
YLR009W	—	3.5	Ribosomal protein L24.e.B <sup>¶</sup>
YDL130W	<i>RPP1B</i>	3.4	Ribosomal protein P1B
YLR048W	<i>RPS0B</i>	3.4	Ribosomal protein S0B
YPL198W	<i>RPL7B</i>	3.3	Ribosomal protein L7B
YEL054C	<i>RPL12A</i>	3.0	Ribosomal protein L12A
YDR025W	<i>RPS11A</i>	3.0	Ribosomal protein S11A
Phosphate regulation (2 ORFs)			
YBR092C	<i>PHO3</i>	12.4	Acid phosphatase, constitutive
YML123C	<i>PHO84</i>	5.3	Inorganic phosphate transporter
Chromatin arrangement (5 ORFs)			
YBL002W	<i>HTB2</i>	8.1	Histone H2B.2
YBL003C	<i>HTA2</i>	7.2	Histone H2A.2
YEL026W§	—	6.2	HMG-like protein Nhp2p <sup>¶</sup>
YDL208W§	<i>NHP2</i>	4.7	HMG-like nuclear protein
YDR225W	<i>HTA1</i>	4.2	Histone H2A
Others (21 ORFs)			
YJR047C	<i>ANB1</i>	9.5	Translation initiation factor
YOR095C	<i>RKI1</i>	8.9	D-Ribose-5-phosphate ketol-isomerase
YDL037C	—	7.4	Similarity to glucan 1,4-α-glucosidase
YNL111C	<i>CYB5</i>	5.8	Cytochrome b5
YMR290C	<i>HAS1</i>	5.7	RNA-dependent helicase
YLR180W	<i>SAM1</i>	5.0	s-Adenosylmethionine synthetase
YER043C	<i>SAH1</i>	4.4	s-Adenosyl-L-homocysteine hydrolase
YNR053C	—	4.3	Tumor-associated autoantigen <sup>¶</sup>
YDL051W	<i>YLA1</i>	4.2	RNA binding protein
YLR045W	<i>ERG3</i>	4.2	C-5 sterol desaturase
YDR144c	<i>MKC7</i>	~3.9	Aspartyl protease
YDL181W	<i>INH1</i>	~3.8	ATPase inhibitor
YJR048W	<i>CYC1</i>	3.7	Cytochrome c isoform 1
YNL289W	<i>PCL1</i>	~3.5	Cyclin, G <sub>1</sub> /S-specific
YKL081W	<i>TEF4</i>	3.4	Translation elongation factor
YAL025C	<i>MAK16</i>	3.3	Nuclear viral propagation protein
YCR034W	<i>FEN1</i>	3.3	Probable β-1,3-glucan synthase
YJL121C	<i>RPE1</i>	3.3	D-Ribulose-5-phosphate 3-epimerase
YDR044W	<i>HEM13</i>	3.2	Coproporphyrinogen III oxidase
YGL055W	<i>OLE1</i>	3.0	Stearoyl-CoA desaturase
YLR372W	<i>SUR4</i>	3.0	Sterol isomerase
Unknown (19 ORFs)			

These ORFs and their fold induction can be found on the World Wide Web at [www.hsph.harvard.edu/geneexpression](http://www.hsph.harvard.edu/geneexpression).

\*Note that two ORFs fall into multiple categories.

†Categories derived from MIPS database (Munich Information Center for Protein Sequences) (31).

‡Hybridization intensity in the untreated sample was below a certain threshold and was therefore increased to an arbitrary, low value for the purposes of this calculation. Values are therefore approximate.

§ORFs falling into two categories.

¶Predicted function based on sequence similarity.

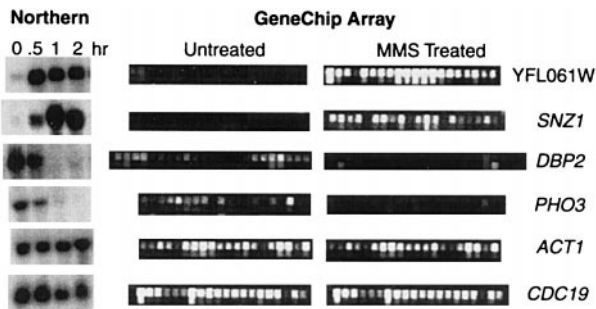


FIG. 2. Verification of GeneChip data by conventional Northern-blot analysis. Northern blots were prepared with total RNA isolated from untreated and 0.1% MMS-treated *S. cerevisiae* DBY747 log-phase cells grown at 30°C in rich media; RNA was from cells exposed to MMS for the indicated times. Blots were probed with 50 ORFs; 6 of them are shown, namely two induced [YFL061W (251-fold) and SNZ1 (~65.8-fold)], two repressed [DBP2 (18.1-fold) and PHO3 (12.3-fold)] and two nonresponsive ORFs [CDC19 and ACT1], as determined by GeneChip analysis. The GeneChip data in this figure are taken from the experiment in Fig. 1. Each ORF is represented on the Ye6100 array by ~20 oligonucleotide pairs. One member of each pair corresponds to a perfectly matched sequence from the ORF (top row); the other pair member contains a single-base mismatch (bottom row). The difference in intensity between the perfectly matched and the mismatched sequences is used to calculate an average intensity difference for each ORF.

response/detoxification, DNA repair/replication, cell cycle, signal transduction, cell wall biogenesis, and membrane transport) are not unexpected, because they could well be providing resis-

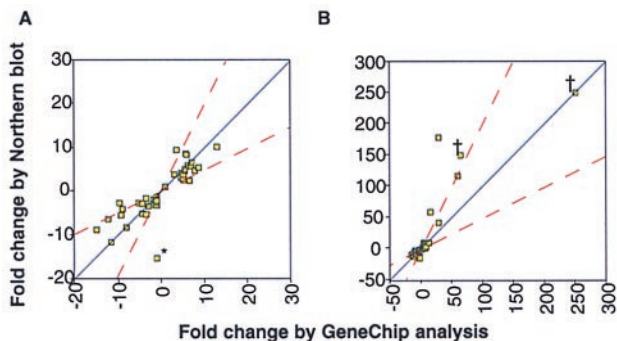


FIG. 3. Correlation of GeneChip and Northern-blot data. The correlation between ORF induction and repression values (in response to MMS) obtained from the GeneChip analysis and Northern-blot data are compared for 50 ORFs. Northern-blot hybridization levels were determined with a Bio-Rad Molecular Imager, and the GeneChip data are from the experiment in Fig. 1. (A) Data plots for genes showing less than a 20-fold change in transcript level (by GeneChip). (B) Data plots for all 50 ORFs. The ORFs were selected to represent the entire range of change (as indicated by the GeneChip data) with changes from a 19-fold decrease to a 251-fold increase. The names and Northern-blot values are YFL061W (>150), SNZ1 (>150), GTT2 (180.0), YKL071W (118.0), SNO1 (60.0), YNR065C (42.0), YIL165C (17.3), SHM1 (9.8), ARG11 (8.7), GIN3 (8.6), ASP3 (6.7), NPL4 (6.1), RNR3 (6.0), ECM29 (5.5), YLR080W (5.3), HIS5 (5.1), YOR227W (5.1), MAG1 (5.0), NTH1 (4.5), YPR1 (4.0), HOM3 (4.0), LAG2 (3.5), CTK3 (3.0), YGR130C (2.8), LHE1 (1.4), YIL131C (1.2), SSA1 (1.0), PHO5 (-15.2), URA7 (-11.4), DBP2 (-8.7), NSB1 (-8.2), PHO3 (-6.2), AAH1 (-5.3), GUA1 (5.1), SAH1 (-5.0), RKI1 (-4.0), YDL213C (-3.1), RPS16A (-3.0), NIP7 (-2.8), YIL158W (2.6), ANB1 (-2.5), TCM1 (-2.3), ENO1 (-2.2), INH1 (-2.0), RPL16A (-2.0), SPE2 (-1.7), CDC19 (-1.6), RPL17 (-1.6), YLR009W (-1.4), and ACT1 (-1.1). The GeneChip values are in Table 1. Dashed red lines represent a factor-of-two difference from a perfect match between Northern-blot and GeneChip analysis. The asterisk in A represents the PHO5 ORF (discussed in Results and Discussion) and † marks transcripts that were undetectable by Northern-blot analysis in untreated cells and whose fold-induction is likely to be less accurate. Note that these points were not included in the calculation of the correlation coefficient.

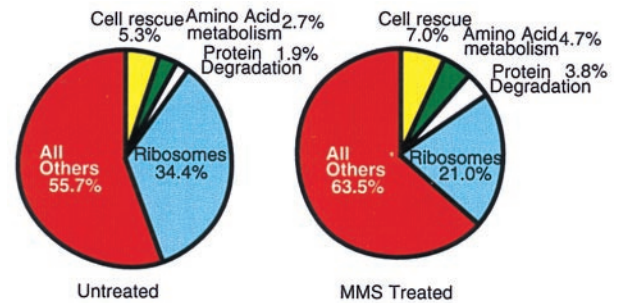


FIG. 4. Changes in the ratios of mRNA subpopulation in response to MMS. The fraction of poly(A)<sup>+</sup> mRNA transcripts from genes related to ribosomes (blue), protein degradation (white), amino acid metabolism (green), cell rescue (yellow), and all others (red) was calculated for untreated and MMS-treated cells by using the GeneChip hybridization data. Hybridization intensities are directly related to absolute poly(A)<sup>+</sup> mRNA level (16). Therefore, hybridization intensities of a given mRNA reflect the abundance of a given mRNA relative to the total amount of mRNA. Genes are grouped into these categories according to the *S. cerevisiae* Genome Database and the Munich Information Centre for Protein Sequences (30, 31).

tance to a chemical that damages DNA. However, with the exception of a few genes (e.g., *MAG1*), it remains to be determined whether and how each gene plays a protective role. Several other groups of genes are more difficult to rationalize, including those for nitrogen and sulfur metabolism, carbohydrate metabolism/fermentation, mRNA processing, "others," and the largest group of MMS-inducible transcripts, 112 ORFs with no known biological function and no sequence similarity to known proteins (see Table 1).

In addition to finding inducible genes potentially involved in repairing, avoiding, or preventing DNA damage, we observed the induced expression of 15 genes involved in protein degradation. In fact, when we relaxed the induction criterion from 4-fold to 2.0-fold (intensity changes greater than 2.0 are considered significant), we observe that 91 of the 143 known protein degradation genes are induced after exposure to this relatively nontoxic MMS dose. Proteins are known substrates for MMS alkylation (20); the fact that genes involved in protein degradation are up-regulated in response to MMS suggests that alkylated proteins may be targeted for degradation and that the elimination of alkylated proteins may be important for cellular recovery. The selective removal of chemically damaged proteins is not unprecedented; recently, it was shown that oxidized proteins are targeted for ubiquitin-mediated degradation in eukaryotic cells (21, 22). It seems logical that cells must replace proteins that were degraded, and evidence for new protein synthesis is suggested by the increased expression of 41 genes involved in amino acid biosynthesis (Table 1). In fact, 91 of the known 194 ORFs involved in amino acid biosynthesis genes are induced greater than 2.0-fold. These data suggest that, in addition to inducing genes to promote recovery from DNA damage, cells also induce genes to promote recovery from protein damage. The relative importance of each pathway remains to be determined, and it should be noted that pathways that prevent DNA alkylation damage may also prevent protein alkylation damage.

**Transcripts Repressed by MMS.** The 76 genes that are repressed after MMS exposure are listed in Table 2. The most notable groups include those involved in nucleotide and RNA synthesis and in the synthesis and assembly of ribosomal proteins. Repression of RNA synthesis and ribosomal genes might suggest that cells down-regulate *de novo* transcription and *de novo* protein synthesis in response to MMS, a suggestion that counters our proposal that MMS induces a program to degrade and replace alkylated proteins. However, an analysis of the global expression of *S. cerevisiae* genes might explain the apparent contradiction. In exponentially growing *S. cerevisiae* cells, ribosomal protein genes produce over one-third of the total poly(A)<sup>+</sup> mRNA (Fig. 4);



indeed, 75 of the 100 most highly transcribed genes in *S. cerevisiae* are ribosomal protein genes. These data are consistent with previous GeneChip and serial analysis of gene expression (16, 23). Overall, MMS treatment induces a modest decrease in the expression of all the ribosomal protein genes (average of 1.7-fold), such that after MMS exposure still one-fifth of all transcripts encode ribosomal proteins (Fig. 4). It seems likely that the modest reduction in ribosomal protein gene expression allows energy to be reshuffled for the increased expression of genes involved in protective responses, while maintaining a basal protein synthesis capacity (Fig. 4). Moreover, a transient but slight decrease in the manufacture of new ribosomes may serve to slow the global production of proteins until alkylation exposure is diminished. There may even be preferential translation of MMS-induced transcripts to promote recovery further, and such preferential translation is not unprecedented (24–26). It should be noted that not all ribosomal protein transcripts are repressed and that the mitochondrial ribosomal protein transcripts actually show a slight increase. (The complete data set can be found at [www.hsph.harvard.edu/geneexpression](http://www.hsph.harvard.edu/geneexpression).) In addition to the repression of transcripts involved in nucleotide, RNA, and protein synthesis, several other genes are repressed (Table 2), and as with the induced transcripts, ORFs of unknown function represent a large group.

Recently, changes in global transcript levels during the cell-cycle progression of *S. cerevisiae* were reported (27). Of the 421 genes determined to have cell-cycle periodicity, only a small fraction (24 genes) are responsive to MMS, suggesting that changes in expression level after MMS treatment are not caused simply by changes in cell-cycle progression. Although over 60% of characterized genes showing cell-cycle periodicity have already been implicated in cell-cycle specific roles, very few of the characterized genes responsive to MMS are known to be involved in MMS resistance. One interpretation is that a majority of the responsive genes are not directly involved in the response to alkylation damage; another is that our study has shown that there are many more genes involved than previously known. The induction by another DNA-damaging agent of a large number of *S. cerevisiae* proteins involved in many different cellular processes was reported recently (28). Evidence obtained from two-dimensional gel electrophoresis indicated that at least 115 proteins are stimulated by H<sub>2</sub>O<sub>2</sub> and that at least 52 are repressed. Of the inducible proteins, 71 have been identified, and previously only 12 were known to act directly in antioxidant defense. The other proteins include heat shock proteins; enzymes for carbohydrate metabolism, amino acid biosynthesis, and protein degradation; and a number of unclassified proteins or proteins with no known function. The induction of a large number of seemingly unrelated proteins after H<sub>2</sub>O<sub>2</sub> exposure parallels the results presented here for transcript induction after MMS exposure.

We have made three major points. (i) The number of yeast genes known to be induced by a DNA-damaging agent has increased by at least 15-fold. (ii) A large number of genes are also repressed. (iii) There is evidence for the initiation of a program to eliminate and replace alkylated proteins from the cell. Whether all of the MMS-inducible and MMS-repressible ORFs listed in Tables 1 and 2 contribute significantly to protecting cells against alkylating agents must now be tested. For a few, their role is already well established, but for most, their role in MMS-resistance remains to be determined. The anticipated availability of thousands of new *S. cerevisiae* strains with null mutations in all nonessential ORFs (29) and with regulated expression of all the ORFs would be essential for such an analysis. Identification and characterization of the regulatory mechanisms for each regulon that contributes to the global response would be particularly important, and it would be important to establish the relative contributions of each type of protective mechanism in ameliorating alkylating-agent toxicity. It would be particularly interesting to determine how critical the turnover of damaged proteins is for the recovery of cells from alkylation damage.

We thank Drs. Armen Tashjian, Evelyn Waldstein, Graham Walker, and John Cairns for useful comments on the manuscript. We also thank Affymetrix for developing such outstanding technology and their Academic User Center Program for making this study possible. This work was supported by National Institutes of Health Grant CA55042 and by Harvard School of Public Health institutional funds. S.A.J. was supported by National Institutes of Health Training Grant CA09078, and L.S. is a Burroughs Wellcome Toxicology Scholar. The Affymetrix User Center is partially supported by National Institutes of Health Grant P01HG01323.

- Friedberg, E., Walker, G. & Siede, W. (1995) *DNA Repair and Mutagenesis* (Am. Soc. Microbiol., Washington, DC).
- Kiser, G. L. & Weinert, T. A. (1996) *Mol. Biol. Cell* **7**, 703–718.
- Herrlich, P., Blattner, C., Knebel, A., Bender, K. & Rahmsdorf, H. J. (1997) *Biol. Chem.* **378**, 1217–1229.
- Zhan, Q., Carrier, F. & Fornace, A. J., Jr. (1993) *Mol. Cell. Biol.* **13**, 4242–4250.
- Fornace, A. J., Jr., Alamo, I., Jr., & Hollander, M. C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8800–8804.
- Bachant, J. B. & Elledge, S. J. (1998) *DNA Repair in Prokaryotes and Lower Eukaryotes*, DNA Damage and Repair, eds. Nickoloff, J. A. & Hoekstra, M. F. (Humana, Totowa, NJ), Vol. 1.
- McClanahan, T. & McEntee, K. (1984) *Mol. Cell. Biol.* **4**, 2356–2363.
- Ruby, S. W. & Szostak, J. W. (1985) *Mol. Cell. Biol.* **5**, 75–84.
- Chen, J., Derfler, B., Maskati, A. & Samson, L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7961–7965.
- Chen, J., Derfler, B. & Samson, L. (1990) *EMBO J.* **9**, 4569–4575.
- Chen, J. & Samson, L. (1991) *Nucleic Acids Res.* **19**, 6427–6432.
- Xiao, W., Singh, K. K., Chen, B. & Samson, L. (1993) *Mol. Cell. Biol.* **13**, 7213–7221.
- Singh, K. K. & Samson, L. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 4907–4911.
- Sweet, D. H., Jang, Y. K. & Sancar, G. B. (1997) *Mol. Cell. Biol.* **17**, 6223–6235.
- Schmitt, M. E., Brown, T. A. & Trumpower, B. L. (1990) *Nucleic Acids Res.* **18**, 3091–3092.
- Wodicka, L., Dong, H., Mittmann, M., Ho, M. H. & Lockhart, D. J. (1997) *Nat. Biotechnol.* **15**, 1359–1367.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
- Cheng, J., Sheldon, E. L., Wu, L., Uribe, A., Gerrue, L. O., Carrino, J., Heller, M. J. & O'Connell, J. P. (1998) *Nat. Biotechnol.* **16**, 541–546.
- Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., et al. (1996) *Nat. Biotechnol.* **14**, 1675–1680.
- Boffa, L. C. & Bolognesi, C. (1985) *Carcinogenesis* **6**, 1399–1401.
- Grune, T., Reinheckel, T. & Davies, K. J. A. (1996) *J. Biol. Chem.* **271**, 15504–15509.
- Shang, F., Gong, X. & Taylor, A. (1997) *J. Biol. Chem.* **272**, 23086–23093.
- Velculescu, V. E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M. A., Bassett, D. E., Jr., Hieter, P., Vogelstein, B. & Kinzler, K. W. (1997) *Cell* **88**, 243–251.
- Kawaguchi, Y., Van Sant, C. & Roizman, B. (1998) *J. Virol.* **72**, 1731–1736.
- Landers, J. E., Cassel, S. L. & George, D. L. (1997) *Cancer Res.* **57**, 3562–3568.
- Schmechel, S., Chute, M., Skinner, P., Anderson, R. & Schiff, L. (1997) *Virology* **232**, 62–73.
- Cho, R. J., Campbell, M. J., Winzler, E. A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T. G., Gabrielian, A. E., Landsman, D., Lockhart, D. J., et al. (1998) *Mol. Cell* **2**, 65–73.
- Godon, C., Lagniel, G., Lee, J., Buhler, J. M., Kieffer, S., Perrot, M., Boucherie, H., Toledano, M. B. & Labarre, J. (1998) *J. Biol. Chem.* **273**, 22480–22489.
- Shoemaker, D. D., Lashkari, D. A., Morris, D., Mittmann, M. & Davis, R. W. (1996) *Nat. Genet.* **14**, 450–456.
- Cherry, J. M., Adler, C., Ball, C., Chervitz, S. A., Dwight, S. S., Hester, E. T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., et al. (1998) *Nucleic Acids Res.* **26**, 73–79.
- Mewes, H. W., Albermann, K., Heumann, K., Lieb, S. & Pfeiffer, F. (1997) *Nucleic Acids Res.* **25**, 28–30.