WestVirginiaUniversity
THE RESEARCH REPOSITORY @ WVU

Graduate Theses, Dissertations, and Problem Reports

2020

The role of genotype in Saccharomyces cerevisiae resistance to 4-methylcyclohexanemethanol

Michael C. Ayers
West Virginia University, mcayers@mix.wvu.edu

Recommended Citation

THE ROLE OF GENOTYPE IN *SACCHAROMYCES CEREVISIAE* RESISTANCE TO 4-METHYLCYCLOHEXANEMETHANOL

Michael C. Ayers

Dissertation submitted to the Eberly College of Arts and Sciences
at West Virginia University

in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Biology

Jennifer Gallagher, Ph.D., Chair
Stephen DiFazio, Ph.D.
Timothy Driscoll, Ph.D.
Jennifer Hawkins, Ph.D.
C. Joel McManus, Ph.D.

Department of Biology

Morgantown, West Virginia

2020

ABSTRACT

The role of genotype in *Saccharomyces cerevisiae* resistance to 4-methylcyclohexanemethanol

Michael C. Ayers

The chemical 4-methylcyclohexanemethanol (MCHM) spilled in the Elk River near Charleston, WV in January 2014, impacting the drinking water of 300,000 residents. Initial studies on MCHM following the spill attempted to characterize its toxicological effects in animals, but largely ignored characterization of its effects on cellular pathways and metabolism. In this study, I used this chemical as a novel stressor for the model eukaryote *S. cerevisiae* to explore the relationship of genotype and stress phenotypes. Initial exploration characterized the stress phenotype of MCHM in yeast through growth assays, cell cycle analysis, and transcriptomic data. MCHM exposure caused cells to arrest growth in G1, activating a well characterized yeast process called the environmental stress response. Further exploration was carried out using a genetic screen of approximately 5000 haploid gene knockout strains, which combined with the transcriptomic data, revealed that the causes of the stress response in yeast were nutrient deprivation related to amino acid biosynthesis and reactive oxygen species production. A QTL analysis of standing variation between two parental strains with variable resistance uncovered the role of zinc homeostasis and its interaction with the hydrotrope chemical properties of MCHM in protein aggregation as a contributor to resistant phenotypes. Finally, an In-Lab evolution study to produce resistant strains for variant analysis showed mutations in the pleiotropic drug response transcription factor *PDR3* cause a reproducible induction of MCHM resistance. These studies combined to characterize cellular changes from MCHM, identify genes required for tolerance, and explore both standing and evolved variation in genotypes that contribute to and produce an MCHM resistant phenotype.

ACKNOWLEDGEMENTS

I could not start to acknowledge the people who made this dissertation possible without thanking my advisor Jen Gallagher above all else. I was out of college for nearly a decade when I applied to WVU. I am so thankful she took a chance on me as one of her first graduate students. She has been there throughout the ups, the downs, the transition from Masters to Ph.D., and the complete overhaul of my project after my original proposal fell through. The greatest thing you can ask for in a mentor is that they want you to succeed just as much as you do. She works as hard as all of her mentees every day to keep us on track. I truly would not have made it through graduate school without a mentor like Jen. I want to say thank you, Jen. I can never repay you.

I also want to thank my committee: Jennifer Hawkins, Stephen DiFazio, Timothy Driscoll, and Joel McManus. Your help has guided me to improve my work and become a better scientist. You have been amazing teachers, and I hope your efforts on me have paid off. I also want to add an extra thanks here to Jennifer Hawkins for the opportunities to work with her as an assistant. It was an amazing experience transitioning from taking her class to teaching with her.

The Gallagher lab has been filled with members since I started in 2014. I often affectionately call our undergrads the undergrad army. I want to thank all of those present and past undergrads for their contributions as scientists and comrades in the lab. Our

post-docs Xiaoqing Rong-Mullins and Amaury Pupo, and our recent addition Tulika Sharma made huge contributions to my development as a scientist. I must especially thank Amaury for both work in the chapters of this dissertation and also my learning experience in bioinformatic analyses.

I want to thank all the graduate students in the Dept. of Biology for creating a great environment for a Ph.D, but I have to acknowledge the other graduate students in my lab who became close friends. I want to thank Suk Lan Ser for being a good friend and labmate. I am so glad we got to teach together this past Spring. Capstone would not have been possible without a friend to teach with. Matt Winans has the distinguished pleasure of being a longtime friend at this point. He, Apoorva, and I went through all the milestones of our Ph.D.s together, and I would not have changed that for anything. Taizina Momtareen is one of my closest friends. She has taken up the enviable mantle of my lab-partner-in-crime over the past year, and for a little while at least will now have our office to herself! I could not have made it through my Ph.D. without the help of such a supportive lab. My best advice at this point is to find a lab like this.

I also have to mention my best friends away from WVU that made it possible to survive graduate school. My best friend since childhood Jerel Culliss has always been there for me. I have missed him in Seattle, but maybe our paths will cross again soon.

My other best friend, Michael Barnes-Reiss in Pittsburgh, got me through one of the roughest times in my life this past year. I will never forget how he took care of me.

Melissa Lehrer joins this list of best friends. She started at WVU in my second year and claimed Apoorva and me as her own. I could not ask for a better friend. I should probably make

iv

some sort of Katy Perry reference here, but it's feeling like this acknowledgements section is Never Really Over.

This brings me to Apoorva Ravishankar. My best friend, labmate, work wife, and every other cliched best friend trope. We started together as the very first graduate students in the Gallagher lab in 2014. I am the luckiest guy in the world to have met my best friend on that day. I have already mentioned how we went through all of our Ph.D. milestones together, but it's more than that. We spent six years in constant contact. We shared a lab bench, daily lunches, advice about life and science, and everything else. I would not be the scientist and person I am without her.

Finally, I want to thank my sister Amy. She has been my rock for decades. The obstacles and celebrations of life are made meaningful by the family we hold closest. I have missed seeing her every day, but we will always be together.

This dissertation is dedicated to my family who are

not here to see its completion:

My brother, Jonathan Ayers

My father, Franklin Ayers

And my mother, Beverly "Sue" Ayers

I Miss you and I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: OBJECTIVES AND OVERVIEW

The relationship between genotypes and phenotypes is central to modern genetics research. Variations in DNA sequence produce variations in phenotypes. Whether the variation results in an alternatively active protein, an inactive protein, or a differentially expressed gene product, changes in sequence can have dramatic effects on an organism's traits. Stress phenotypes, such as loss of viability, cell cycle arrest, DNA damage, and transcriptome reprogramming are studied in yeast to analyze both the toxic effects of stressors and the roles of gene and protein networks in responses. Drugs such as hydroxyurea, identified by screens in the 1960s to have anti-tumor properties and subsequently used to treat diseases such as cancers and sickle cell anemia, have also become useful as sources of chemical stress in cellular biology research (reviewed in Singh & Xu, 2016). For instance, the reversible inhibition of DNA replication by hydroxyurea allows researchers to arrest cells in S phase and study synchronized cell populations. It also damages DNA and causes oxidative stress, so it is used in the study of stress responses and DNA damage repair mechanisms (reviewed in Singh & Xu, 2016). As with much cellular biology research, studying chemical stress phenotypes can impact seemingly unrelated areas of research by revealing intertwined networks of proteins and metabolism in the cell.

The chemical 4-methylcyclohexanemethanol (MCHM) is used in industrial processes to clean coal, and it is most well-known from a spill into the Elk River in January 2014 where it contaminated the drinking water of 300,000 residents in the Charleston, WV area. In the months following the spill, the National Toxicology Program (NTP) embarked on a testing

strategy to determine the potential toxicological dangers of the chemical. Most of the effects the NTP identified were found only at levels 100 or more times higher than those detected from the spill. The most concerning effect was on developmental phenotypes, specifically low fetal birth weight and skeletal abnormalities in rodents (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). Another biological study identified photomotor response changes in zebrafish possibly indicative of neurotoxic effects (Horzmann et al., 2017). Almost no analysis of cellular and molecular changes that occur with MCHM treatment was performed. The NTP studied minimum dosages required to exhibit gene expression changes in rodent kidney and liver cells, approximately 50 times higher than a 1ppm screening level set by the CDC after the spill (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). Another study looked at the expression of a subset of 148 stress proteins in yeast with treatment by MCHM, which found that there was increased expression of chemical stress and protein stress proteins, and DNA damage and oxidative stress in cotreatment of MCHM and a liver extract additive intended to produce MCHM metabolites (Lan et al., 2015). While toxicological outcomes were relatively well studied after the spill, there was very little study of the cellular effects of MCHM prior to this dissertation.

The purpose of the research in this dissertation was to study the genetics of chemical stresses in yeast by using the toxicologically important MCHM. The main goal of this research was to identify genetic underpinnings of MCHM resistant phenotypes. Initial work characterized the phenotypes of MCHM treatment using growth assays and viability assays. I put forth the following specific aims to explore the genotypes that produced MCHM tolerance in yeast.

Specific Aim 1: **Explore the effects of MCHM on eukaryotic cells by identifying stress phenotypes and the genes required for resistance in *S. cerevisiae*.** I characterized the effect of MCHM on the growth of yeast, using viability and cell cycle analysis to determine that the chemical arrests cells in G1 to inhibit growth, as opposed to killing cells outright. We also used a transcriptomic dataset to characterize the cellular response phenotype as the programmed yeast environmental stress response. I utilized a genetic screen of nearly 5000 single gene knockouts to get an unbiased analysis of required genes for MCHM resistance. Elucidating the types of stress phenotypes of the cells with these two genomic datasets, we performed follow-up assays to characterize and confirm the nutrient, oxidative, and DNA damage stresses that were seen.

Specific Aim 2: **Understand the role of existing genetic variation in resistance to MCHM using two strains with variable resistance phenotypes.** I used a segregant collection of 125 haploid offspring of two parental strains with variable resistance to MCHM to perform a QTL on growth in rich liquid media containing MCHM. While variation being spread relatively evenly across the genome seemed was evident, one peak was significant. I used knockouts in multiple strains to narrow down the peak to a zinc transporter gene and confirmed the importance of zinc with supplementation. We combined this data with the analysis of metal ion transporter changes in the transcriptome and changing levels of iron and zinc in MCHM treated yeast to further characterize the role of zinc and metal homeostasis in response to MCHM. This work was

3

combined with transcriptomics and protein aggregation experiments performed by collaborators to explore the interaction between MCHM and zinc in resistance.

Specific Aim 3: **Examine adaptation in the production of resistance to MCHM through production of new genetic variation.** I treated multiple replicates of the sensitive yeast parent from the QTL strains in Aim 2 with MCHM in rich liquid media until resistance developed. Variant analysis was performed to identify the mutations that appeared during the adaptation to MCHM. Variants from control evolution replicates in liquid media without MCHM were used to eliminate regions that accumulate mutations unimportant for resistance to MCHM. Genes containing coding variants unique to the resistant strains were analyzed for shared functions. I filtered any genes that mutated often in resistant strains but rarely in control strains. One gene mutated in every resistant strain, producing mutant alleles including single amino acid changes and large C-terminal truncations due to early stops. I then initiated follow-up functional genetics on this pleiotropic drug response transcription factor to determine if mutations in the parent strain and other genetic backgrounds were sufficient to induce resistance.

References

Horzmann, K. A., de Perre, C., Lee, L. S., Whelton, A. J., & Freeman, J. L. (2017). Comparative analytical and toxicological assessment of methylcyclohexanemethanol (MCHM) mixtures associated with the Elk River chemical spill. *Chemosphere*, *188*, 599–607. https://doi.org/10.1016/j.chemosphere.2017.09.026

Lan, J., Hu, M., Gao, C., Alshawabkeh, A., & Gu, A. Z. (2015). Toxicity assessment of 4-Methyl-1-cyclohexanemethanol and its metabolites in response to a recent chemical spill in west virginia, USA. *Environmental Science and Technology*, *49*(10), 6284–6293. https://doi.org/10.1021/acs.est.5b00371

Singh, A., & Xu, Y.-J. (2016). The Cell Killing Mechanisms of Hydroxyurea. *Genes*, *7*(11). https://doi.org/10.3390/genes7110099

West Virginia Chemical Spill: Collective NTP Findings and Supporting Files. (n.d.). National Toxicology Program U.S. Department of Health and Human Services. Retrieved March 8, 2017, from https://ntp.niehs.nih.gov/results/areas/wvspill/collective.html

Chapter 2: Introduction

Genotypic and phenotypic diversity in yeast:

Saccharomyces cerevisiae is a key eukaryotic model organism for research into cellular function and genomics precisely because it is a diverse species. It has a genome size of approximately 12Mb containing around 6000 genes, yet routinely any two strains may be as variable as 60,000 to 100,000 SNPs and hundreds or thousands of indels (Engel & Cherry, 2013). This genotypic diversity is accompanied by a similarly diverse range of phenotypes. Researchers often explore or exploit phenotypic diversity in culture growth, colony size and morphology, high and low temperature growth, nutrient utilization, chemical and drug resistance, and other cellular traits in yeast strains (J. E. G. Gallagher et al., 2014; Liti et al., 2009; Steinmetz et al., 2002; Strope et al., 2015). The results of such studies inform upon the roles of individual genes within the cell, but also how these genes and the proteins for which they encode form pathways and interrelated networks that create phenotypes and characteristics at the organismal and community level.

Yeast is a useful model to explore relationships between genotype and phenotypes such as resistance to various chemicals. Experiments on specific chemicals do not only reveal the genetic and protein network basis of the resistance to single chemicals. They often reveal how identified gene and protein networks work together in cellular functions. Cellular pathways such as DNA replication or cell cycle analysis through synchronization can be explored through studying stress from chemicals like hydroxyurea (Singh & Xu, 2016). It is sometimes only through stressful perturbations to the cell that processes and mechanisms can be revealed.

6

Previous and continuing work of our group studies several chemical perturbations in yeast cells, including copper, 4-Nitroquinolone-1-oxide (4NQO), glyphosate-based herbicides, and MCHM (J. E. G. Gallagher et al., 2014; Ravishankar, Cumming, et al., 2020; Ravishankar, Pupo, et al., 2020; Xiaoqing Rong-Mullins, Ravishankar, et al., 2017; Xiaoqing Rong-Mullins, Winans, et al., 2017; Xiaqing Rong-Mullins et al., 2017; Winans & Gallagher, 2020).

One of the major ways that resistance to these chemicals can be explored is by exploiting the natural genetic variation in different yeast strains. It is efficient to perform growth assays of these strains in the presence of the chemical to elucidate phenotypic variation in resistance. The resistant and sensitive strains can then be analyzed for the underlying genetic differences behind the phenotypes. For instance, if yeast strain YJM789 is resistant to the chemical 4NQO while strain S288c is sensitive, a collection of haploid progeny of these two parent strains with sufficient segregating markers can be used to perform a quantitative trait locus (QTL) analysis of the resistant phenotype. Indeed, such an analysis, showing the association of a genome locus and the resistant phenotype, revealed that the resistance to 4NQO observed in the YJM789 strain versus the S288c strain maps to a gene called *YRR1* (J. E. G. Gallagher et al., 2014). Further analysis of this strain using genetic manipulations to swap alleles of *YRR1* between the strains allowed for confirmation of this genotype-phenotype relationship. The segregant collection used in this *YRR1* study was also used for a glyphosate-based herbicide study that found mapped resistance to variation in the genes *DIP5* and *PDR5* in these strains. QTL mapping of this collection was used in the studies of this dissertation. The ease of these types of manipulations in yeast allows for unique opportunities for functional genomics.

Other ways that genotypes related to resistance phenotypes can be explored are through laboratory evolution and genetic screen experiments. Yeast as a model organism for genetics research benefits from decades of community resources, including a collection of strains with systematically knocked out ORFs of each of the genome's nearly 5000 nonessential genes (Giaever et al., 2002; Giaever & Nislow, 2014). With each gene knocked out, it is possible to screen for the requirement of each one for any number of phenotypes, including resistance to a chemical. This functional genomics approach of screening the entirety of the yeast knockout collection resulted in functional annotations to thousands of proteins (reviewed in Giaever & Nislow, 2014). Laboratory evolution experiments exploit the ability to quickly evolve yeast strains using their short generation time of approximately 90 minutes in liquid culture to produce strains with a resistant phenotype. Sequencing to determine variants in the resistant strains then allows researchers to test which genotypes are sufficient to produce the phenotype. Evolving yeast in glyphosate-based herbicides revealed copy number variation in genes related to cell wall stability, mitochondrial function, and sterol content that were likely responsible for resistance to additives in the herbicide besides glyphosate (Ravishankar, Pupo, et al., 2020). Both genetic screening of the knockout collection and In-Lab evolutions were used in this dissertation to explore the relationship of genotype to MCHM resistance in yeast.

The Elk River Spill chemical MCHM:

The chemical MCHM was a previously little-studied chemical involved in the processing of coal, until a rusted storage tank resulted in a spill of approximately 10,000 gallons of crude MCHM into the Elk River near Charleston, WV on January 9th, 2014. This spill interrupted the water supply of approximately 300,000 residents. Research on MCHM since the Elk River spill

increased over concerns about the little characterized physical and biological properties of this chemical. In the immediate aftermath of the spill, the chemical was detectable in homes due to its intense and distinct licorice smell. MCHM consists of a cyclohexane ring with a methanol group directly opposite a methyl group on the ring structure. The methyl and methanol groups allow for both cis and trans isomers. The trans isomer of MCHM was found to be detectable due to this smell at approximately 0.06ppb in the air (D. L. Gallagher et al., 2015). Residents reported skin, respiratory, and gastrointestinal issues to poison control and healthcare providers that cleared up without medical intervention (Thomasson et al., 2017). The levels of MCHM at the water intake of the treatment plant for these residents showed levels around 2ppm several days after the spill (Weidhaas et al., 2017). The estimated potential max exposure delivered to homes was approximately 2.6ppm (Weidhaas et al., 2017), and by the January 27th one study detected levels in tap samples of only around 0.15ppm that continued dropping (Foreman et al., 2015). A study conducted well after the spill evaluated the effect of activated carbon on remediation of the chemical in a treatment plant scenario and found it to be the best option as compared to permanganate or chlorine treatment. However, unexpectedly high levels of the activated carbon would have been needed to reach CDC screening levels of 1ppm (Weidhaas et al., 2017). In contrast to chemical water treatment strategies, several studies found that there would potentially have been significant bioremediation from bacterial sources in sediments accumulating the chemical from contact with the water from the spill (Cozzarelli et al., 2017; Yuan et al., 2016b, 2016a).

MCHM toxicological work in broad species:

9

The National Toxicology Program, part of the U.S. Dept. of Health and Human Services, was nominated by the Centers for Disease Control to perform a battery of studies on MCHM in the immediate aftermath of the spill (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). As a chemical with little previous characterization for its effects on biological systems, it was important to run detailed assessments from toxicity in organisms from bacteria to metazoan, to developmental effects on model organism embryos and fetuses. The major goal of the studies was to evaluate whether the screening levels of MCHM in the drinking water imposed by the CDC at the time of the spill, 1ppm, were appropriate to protect the health of the community (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). NTP designed studies to explore developmental toxicity in rats, fish, and worms, dermatologic sensitivity in mice, motor responses indicative of neurological toxicity in fish, and mutagenic effects on the DNA of bacteria and rat blood cells (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). This covered the major concerns of toxicological effects on humans coming into contact with MCHM through the Elk River spill, specifically sensitization to the chemical, a worry due to self-reported rashes among residents following the spill, and developmental and neurotoxic effects, a worry for pregnant women and their children born in the aftermath of the spill (Benson et al., 2018; Johnson et al., 2017; Monnot et al., 2017; *West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.).

The NTP studies on dermatologic sensitization and irritation in mice and rats did show an effect of sensitivity, but at concentrations much higher, on the order of 100,000 fold higher, than the levels the CDC set for screening in the aftermath of the spill (*West Virginia Chemical*

Spill: Collective NTP Findings and Supporting Files, n.d.). The analysis of a photomotor assay

response in zebrafish, potentially indicative of neurotoxicity, saw an effect at 11ppm of MCHM,

approximately 10 times above the CDC levels for water screening (*West Virginia Chemical Spill:*

Collective NTP Findings and Supporting Files, n.d.). This activity was in agreement with the

results of zebrafish work by another lab that detected alterations in motor responses due to

MCHM at between 1 and 25ppm (Horzmann et al., 2017). These responses were at relatively

low dosages compared to the minimum dosages found for other assay effects. NTP assays for

DNA mutagenicity in both bacteria and rat red blood cells detected no DNA damage. An early

yeast study from an outside lab did find evidence that potential metabolites of MCHM from

cotreating yeast and a human cell line with MCHM and S9 liver extract increase expression of

markers related to DNA damage, such as Rad51, a protein involved in repairing DNA double

strand breaks (Lan et al., 2015). The only molecular assays done by the NTP studies was an

expression analysis to determine minimum dosages of MCHM to induce gene expression

changes in liver and kidney of treated rats, presumably a proxy for detection of and attempted

adaptation to the stressor. This assay showed that levels approximately 50 times higher than

the CDC screening level of 1ppm was required to induce changes in liver gene expression, and

kidney gene expression did not change (*West Virginia Chemical Spill: Collective NTP Findings*

and Supporting Files, n.d.).

The most significant finding of the toxicological assessment, according to the NTP, was a

study of developmental toxicity in rats. There was a drop in fetal weight corresponding to a

200mg/kg dose, and significant skeletal deformations of the developing fetus at 400mg/kg

dose. These dosages were calculated to be over 1000 fold higher than any expected exposures

of pregnant woman living in the vicinity of the Elk River spill, but the NTP ordered follow-up analyses of birth weight outcomes for the region in response to this finding (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). The findings of this study did not show any association between adverse birth outcomes, such as decreased birth weight and APGAR scores, for those mothers who may have been exposed to the Elk River spill (Benson et al., 2018). Despite most evidence leading to estimates of little effect of the Elk River spill on health outcomes, this is due to the relatively low dosages of MCHM that was found in the spill. The findings by the NTP and other labs of developmental changes, motor response changes, and DNA damage stress, make further study of this chemical warranted.

MCHM studies in yeast:

Using yeast as a model for studying MCHM effects on cellular processes also began in the aftermath of the Elk River spill. These studies were more focused on characterizing how the chemical produces stress in the cell instead of focusing on toxicological outcomes such as potential for mortality or developmental effects on organisms. The earliest study in yeast was an expression study of a subset of yeast stress proteins tagged with GFP that exposed yeast and a human cell line to MCHM, as well as MCHM + S9, an extract from liver that is intended to provide insight into metabolites of chemicals with which it is cotreated (Lan et al., 2015). This study showed that MCHM increased expression of stress response proteins from their curated collection of proteins known to be important for stress response. Chemical stress and protein stress proteins were upregulated with MCHM alone, and oxidative stress response proteins were upregulated with the MCHM + S9 treatment (Lan et al., 2015). As stated previously, DNA damage response proteins such as Rad51 were also upregulated (Lan et al., 2015). This study

12

was designed to use a stress response collection of GFP tagged proteins, so it was biased to study this subset of 148 proteins from the yeast proteome, out of a genome containing approximately 6000 genes.

Other work in yeast includes the studies of the Gallagher Lab performed concurrently with this dissertation. One study evaluated the metabolome of yeast treated with MCHM by ESI-MS and GC-MS to characterize how MCHM affects yeast metabolism (Pupo et al., 2019). The analysis of polar metabolites found an accumulation of amino acids in the cell that indicates an effect on their production or their depletion through protein production (Pupo et al., 2019). Transcriptomics in the same study pointed to both reduced expression of ribosomal proteins and amino acid biosynthetic gene upregulation, evidence for these causes of amino acid accumulation (Pupo et al., 2019). The nonpolar metabolites found reduction of several phospholipids, such as phosphatidylinositol, that would impact the biophysical properties of membranes in the cell (Pupo et al., 2019). These results agree with and support work presented in this dissertation that describe stress responses in yeast in response to MCHM and MCHM's ability to act as a hydrotrope. The other study conducted by the Gallagher Lab evaluated the effects of the polymorphic polyQ tracts of the Mediator Complex tail protein Med15 on responses to stressors, including MCHM (J. E. G. Gallagher et al., 2020). Genetic variation in this transcription factor, expressed from plasmids, affected the tolerance of MCHM in the same genetic background (J. E. G. Gallagher et al., 2020). This was suggested possibly to be due to the hydrotropic nature of MCHM and the way in which the polymorphic polyQ tracts of the Med15 proteins interact to form protein folding aggregations (J. E. G. Gallagher et al., 2020). The

transcriptomic dataset from this study, analyzed by Amaury Pupo, appears with further downstream analysis by me in the work of this dissertation.

Conclusion:

The work in these studies combines with the studies presented in this dissertation to explore MCHM as a novel stressor of yeast. The work in this dissertation incorporates new genomic datasets and new analyses of existing datasets from earlier studies to characterize the role of genotype in MCHM resistance phenotypes. My work shows that MCHM activates the yeast environmental stress response due to its impacts on nutrient pathways, oxidative stress, metal metabolism, and DNA damage. Genetic variation in a metal homeostatic gene was associated with MCHM resistance, though this phenotype is mostly controlled by small effect loci throughout the genome. Despite the small effects of existing variation, mutations affecting the protein sequence of the one of the transcriptional regulators of drug transporter genes provided a reproducible path for evolved resistance. This work expanded the understanding of MCHM stress phenotypes by focusing on cellular changes. It is also the first to examine the genotypes controlling these phenotypes. The insights from these stress responses are useful to the broader field of yeast stress genetics and the toxicological questions that remain unanswered about this industrial spill chemical.

References

Benson, S. M., Ruestow, P., Keeton, K. A., Novick, R. M., Marsh, G. M., & Paustenbach, D. J. (2018). The 2014 crude 4-methylcyclohexanemethanol chemical release and birth outcomes in West Virginia. *Archives of Environmental & Occupational Health*, *73*(5), 292–301. https://doi.org/10.1080/19338244.2017.1350132

Cozzarelli, I. M., Akob, D. M., Baedecker, M. J., Spencer, T., Jaeschke, J., Dunlap, D. S., Mumford, A. C., Poret-Peterson, A. T., & Chambers, D. B. (2017). Degradation of Crude 4-MCHM (4-Methylcyclohexanemethanol) in Sediments from Elk River, West Virginia. *Environmental Science & Technology*, *51*(21), 12139–12145. https://doi.org/10.1021/acs.est.7b03142

Engel, S. R., & Cherry, J. M. (2013). The new modern era of yeast genomics: Community sequencing and the resulting annotation of multiple Saccharomyces cerevisiae strains at the Saccharomyces Genome Database. *Database*, *2013*(0), bat012–bat012. https://doi.org/10.1093/database/bat012

Foreman, W. T., Rose, D. L., Chambers, D. B., Crain, A. S., Murtagh, L. K., Thakellapalli, H., & Wang, K. K. (2015). Determination of (4-methylcyclohexyl)methanol isomers by heated purge-and-trap GC/MS in water samples from the 2014 Elk River, West Virginia, chemical spill. *Chemosphere*, *131*, 217–224. https://doi.org/10.1016/j.chemosphere.2014.11.006

Gallagher, D. L., Phetxumphou, K., Smiley, E., & Dietrich, A. M. (2015). Tale of two isomers: Complexities of human odor perception for cis- and trans-4-methylcyclohexane methanol from the chemical spill in West Virginia. *Environmental Science & Technology*, *49*(3), 1319–1327. https://doi.org/10.1021/es5049418

Gallagher, J. E. G., Ser, S. L., Ayers, M. C., Nassif, C., & Pupo, A. (2020). The Polymorphic PolyQ

Tail Protein of the Mediator Complex, Med15, Regulates the Variable Response to Diverse

Stresses. *International Journal of Molecular Sciences*, *21*(5).

https://doi.org/10.3390/ijms21051894

Gallagher, J. E. G., Zheng, W., Rong, X., Miranda, N., Lin, Z., Dunn, B., Zhao, H., & Snyder, M. P.

(2014). Divergence in a master variator generates distinct phenotypes and transcriptional

responses. *Genes and Development*, *28*(4), 409–421. https://doi.org/10.1101/gad.228940.113

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A.,

Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R.,

Brachat, S., Campanaro, S., Curtiss, M., Davis, K., … Johnston, M. (2002). Functional profiling of

the Saccharomyces cerevisiae genome. *Nature*, *418*(6896), 387–391.

https://doi.org/10.1038/nature00935

Giaever, G., & Nislow, C. (2014). The yeast deletion collection: A decade of functional genomics.

Genetics, *197*(2), 451–465. https://doi.org/10.1534/genetics.114.161620

Horzmann, K. A., de Perre, C., Lee, L. S., Whelton, A. J., & Freeman, J. L. (2017). Comparative

analytical and toxicological assessment of methylcyclohexanemethanol (MCHM) mixtures

associated with the Elk River chemical spill. *Chemosphere*, *188*, 599–607.

https://doi.org/10.1016/j.chemosphere.2017.09.026

Johnson, V. J., Auerbach, S. S., Luster, M. I., Waidyanatha, S., Masten, S. A., Wolfe, M. S.,

Burleson, F. G., Burleson, G. R., & Germolec, D. R. (2017). Evaluation of 4-

methylcyclohexanemethanol (MCHM) in a combined irritancy and Local Lymph Node Assay

(LLNA) in mice. *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, *105*, 99–105. https://doi.org/10.1016/j.fct.2017.03.034

Lan, J., Hu, M., Gao, C., Alshawabkeh, A., & Gu, A. Z. (2015). Toxicity assessment of 4-Methyl-1-cyclohexanemethanol and its metabolites in response to a recent chemical spill in west virginia, USA. *Environmental Science and Technology*, *49*(10), 6284–6293. https://doi.org/10.1021/acs.est.5b00371

Liti, G., Carter, D. M., Moses, A. M., Warringer, J., Parts, L., James, S. a, Davey, R. P., Roberts, I. N., Burt, A., Tsai, I. J., Bergman, C. M., Bensasson, D., Michael, J. T., Kelly, O., Oudenaarden, A. V., Barton, D. B. H., Bailes, E., Alex, N., Ba, N., … Louis, E. J. (2009). Population genomics of domestic and wild yeasts. *Nature*, *458*(7236), 337–341. https://doi.org/10.1038/nature07743.Population

Monnot, A. D., Novick, R. M., & Paustenbach, D. J. (2017). Crude 4-methylcyclohexanemethanol (MCHM) did not cause skin irritation in humans in 48-h patch test. *Cutaneous and Ocular Toxicology*, *36*(4), 351–355. https://doi.org/10.1080/15569527.2017.1296854

Pupo, A., Ku, K. M., & Gallagher, J. E. G. (2019). Effects of MCHM on yeast metabolism. *PLOS ONE*, *14*(10), e0223909. https://doi.org/10.1371/journal.pone.0223909

Ravishankar, A., Cumming, J. R., & Gallagher, J. E. G. (2020). Mitochondrial metabolism is central for response and resistance of Saccharomyces cerevisiae to exposure to a glyphosate-based herbicide. *Environmental Pollution*, *262*, 114359. https://doi.org/10.1016/j.envpol.2020.114359

Ravishankar, A., Pupo, A., & Gallagher, J. E. G. (2020). Resistance Mechanisms of Saccharomyces cerevisiae to Commercial Formulations of Glyphosate Involve DNA Damage Repair, the Cell Cycle, and the Cell Wall Structure. *G3: Genes, Genomes, Genetics*, *10*(6), 2043–2056. https://doi.org/10.1534/g3.120.401183

Rong-Mullins, Xiaoqing, Ravishankar, A., McNeal, K. A., Lonergan, Z. R., Biega, A. C., Creamer, J. P., & Gallagher, J. E. G. (2017). Genetic variation in Dip5, an amino acid permease, and Pdr5, a multiple drug transporter, regulates glyphosate resistance in S. cerevisiae. *PLOS ONE*, *12*(11), e0187522. https://doi.org/10.1371/journal.pone.0187522

Rong-Mullins, Xiaoqing, Winans, M. J., Lee, J. B., Lonergan, Z. R., Pilolli, V. A., Weatherly, L. M., Carmenzind, T. W., Jiang, L., Cumming, J. R., Oporto, G., & Gallagher, J. (2017). Proteomic and genetic analysis of S. cerevisiae response to soluble copper leads to improvement of antimicrobial function of cellulosic copper nanoparticles. *Metallomics*, *9*(9), 1304–1315. https://doi.org/10.1039/C7MT00147A

Rong-Mullins, Xiaqing, Ayers, M. C., Summers, M., & Gallagher, J. E. G. (2017). Transcriptional Profiling of Saccharomyces cerevisiae Reveals the Impact of Variation of a Single Transcription Factor on Differential Gene Expression in 4NQO, Fermentable, and Non-fermentable Carbon Sources. *G3*, *8*(February), g3.300138.2017. https://doi.org/10.1534/g3.117.300138

Singh, A., & Xu, Y.-J. (2016). The Cell Killing Mechanisms of Hydroxyurea. *Genes*, *7*(11). https://doi.org/10.3390/genes7110099

Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., & Davis, R. W. (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature*, *416*(6878), 326–330. https://doi.org/10.1038/416326a

Strope, P. K., Skelly, D. A., Kozmin, S. G., Mahadevan, G., Stone, E. A., Magwene, P. M., Dietrich, F. S., & McCusker, J. H. (2015). The 100-genomes strains, an S. cerevisiae resource that illuminates its natural phenotypic and genotypic variation and emergence as an opportunistic pathogen. *Genome Research*, *125*(5), 762–774. https://doi.org/10.1101/gr.185538.114

Thomasson, E. D., Scharman, E., Fechter-Leggett, E., Bixler, D., Ibrahim, S., Duncan, M. A., Hsu, J., Scott, M., Wilson, S., Haddy, L., Pizon, A., L. Burrer, S., Wolkin, A., & Lewis, L. (2017). Acute Health Effects After the Elk River Chemical Spill, West Virginia, January 2014. *Public Health Reports*, *132*(2), 196–202. https://doi.org/10.1177/0033354917691257

Weidhaas, J., Lin, L. S., & Buzby, K. (2017). A case study for orphaned chemicals: 4-methylcyclohexanemethanol (MCHM) and propylene glycol phenyl ether (PPH) in riverine sediment and water treatment processes. *Science of the Total Environment*, *574*, 1396–1404. https://doi.org/10.1016/j.scitotenv.2016.08.063

West Virginia Chemical Spill: Collective NTP Findings and Supporting Files. (n.d.). National Toxicology Program U.S. Department of Health and Human Services. Retrieved March 8, 2017, from https://ntp.niehs.nih.gov/results/areas/wvspill/collective.html

Winans, M. J., & Gallagher, J. E. G. (2020). Metallomic and lipidomic analysis of S. cerevisiae response to cellulosic copper nanoparticles uncovers drivers of toxicity. *Metallomics*, *12*(5), 799–812. https://doi.org/10.1039/D0MT00018C

Yuan, L., Zhi, W., Liu, Y., Smiley, E., Gallagher, D., Chen, X., Dietrich, A. M., & Zhang, H. (2016a). Degradation of cis- and trans-(4-methylcyclohexyl) methanol in activated sludge. *Journal of Hazardous Materials*, *306*, 247–256. https://doi.org/10.1016/j.jhazmat.2015.12.016

Yuan, L., Zhi, W., Liu, Y., Smiley, E., Gallagher, D., Chen, X., Dietrich, A., & Zhang, H. (2016b). Aerobic and anaerobic microbial degradation of crude (4-methylcyclohexyl)methanol in river sediments. *The Science of the Total Environment*, *547*, 78–86. https://doi.org/10.1016/j.scitotenv.2015.12.144

CHAPTER 3: OXIDATIVE STRESS RESPONSES AND NUTRIENT STARVATION IN MCHM TREATED *SACCHAROMYCES CEREVISIAE*

Michael C. Ayers, Zachary N. Sherman, and Jennifer E.G. Gallagher[*]

Abstract

In 2014, the coal cleaning chemical 4-methylcyclohexane methanol (MCHM) spilled into the water supply for 300,000 West Virginians. Initial toxicology tests showed relatively mild results, but the underlying effects on cellular biology were underexplored. Treated wildtype yeast cells grew poorly, but there was only a small decrease in cell viability. Cell cycle analysis revealed an absence of cells in S phase within thirty minutes of treatment. Cells accumulated in G1 over a six-hour time course, indicating arrest instead of death. A genetic screen of the haploid knockout collection revealed 329 high confidence genes required for optimal growth in MCHM. These genes encode three major cell processes: mitochondrial gene expression/translation, the vacuolar ATPase, and aromatic amino acid biosynthesis. The transcriptome showed an upregulation of pleiotropic drug response genes and amino acid biosynthetic genes and downregulation in ribosome biosynthesis. Analysis of these datasets pointed to environmental stress response activation upon treatment. Overlap in datasets included the aromatic amino acid genes *ARO1*, *ARO3*, and four of the five *TRP* genes. This implicated nutrient deprivation as the signal for stress response. Excess supplementation of nutrients and amino acids did not improve growth on MCHM, so the source of nutrient deprivation signal is still unclear. Reactive oxygen species and DNA damage were directly detected with MCHM treatment, but timepoints showed these accumulated slower than cells

arrested. We propose that wildtype cells arrest from nutrient deprivation and survive, accumulating oxidative damage through the implementation of robust environmental stress responses.

Introduction

The chemical 4-methylcyclohexane methanol (MCHM) was a previously little-studied chemical involved in the processing of coal, until a rusted storage tank resulted in a spill of crude MCHM into the Elk River near Charleston, WV. The spill's size and location adjacent to a drinking water treatment intake were sufficient to fill homes with an overpowering odor that left many fearful of health consequences (Thomasson et al., 2017). This spill interrupted the water supply of approximately 300,000 residents. Research on MCHM since the Elk River spill has increased dramatically over concerns about the lack of characterized physical and biological properties of this chemical (Weidhaas et al., 2016). For instance, an improved toxicological study on MCHM's effect on model organism viability has been performed (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.), as well as one study on the potential stress responses it may produce in yeast (Lan et al., 2015). However, these studies had a focus on a few specific toxicological outcomes and predetermined stress pathways that may miss other cellular changes. Recently, research has been begun to identify the biochemical and transcriptional changes that MCHM may produce in organisms (Pupo, Ayers, et al., 2019; Pupo, Ku, et al., 2019).

The environmental stress response (ESR) is a shared transcriptional response to multiple stressors, including heat shock, osmotic shock, and hydrogen peroxide treatment, among

22

others (Gasch et al., 2000). The role of paralogous transcription factors Msn2 and Msn4 are important for a large portion of the ESR transcriptional induction (Gorner et al., 1998; Martínez-Pastor et al., 1996). The transcriptional programming of the environmental stress response incorporates signals from diverse stress pathways including nutrient starvation (TOR), osmolarity stress (HOG), and others, into the Msn2/4 transcriptional activators (Capaldi et al., 2008; De Wever et al., 2005; Garmendia-Torres et al., 2007; Gorner et al., 1998; Gutin et al., 2015; Santhanam et al., 2004). Implementation of the stress response involves two waves of Msn2/4 import to the nucleus that modulate the initial intensity and prolonged duration of the response (Gutin et al., 2019). The signaling kinase Mck1 is indispensable for the prolonged response requiring the second import of active Msn2/4 to the nucleus (Gutin et al., 2019), although it is unclear how the signal is maintained as there does not appear to be direct Mck1 interaction with Msn2 (Hirata et al., 2002). The karyopherin Msn5 involved in nuclear import and export activity is important to maintain the prolonged stress response activity of Msn2/4 in the ESR, apparently through the export of initial wave Msn2 from the nucleus, allowing for the second wave of transcriptional activators to function (Gutin et al., 2019).

Amino acid biosynthesis is predominately regulated by the general amino acid transcriptional activator Gcn4. Under amino acid starvation, the levels of Gcn4 increase due to the decreased degradation and increased translation of the protein (reviewed in Hinnebusch, 1997; Meimoun et al., 2000). Gcn4 binds to a consensus promoter sequence upstream of many of the amino acid biosynthetic genes to positively regulate their expression (Arndt & Fink, 1986; Hill et al., 1986; Oliphant et al., 1989). This amino acid transcriptional control includes the aromatic amino acid biosynthetic gene *ARO3* and the specific tryptophan biosynthetic genes

TRP2, TRP3, TRP4, and *TRP5* (Braus, 1991). Tryptophan has been implicated for roles in stress

tolerances other than nutrient starvation. SDS sensitivity, a cell wall and plasma membrane

integrity stress that involves Mck1 effectors in yeast, is dependent on tryptophan biosynthesis

and levels of tryptophan and tyrosine in the media and cell (Schroeder & Ikui, 2019).

Tryptophan biosynthetic mutants, *trp1-5*, were also reported in a special warning for their use

as auxotrophic markers in yeast genetics due to an aberrant sensitivity to many stressors

including rapamycin, high pH, and several metal cations (González et al., 2008). A more recent

report has shown that tryptophan depletion due to a combination of transporter dysfunction at

low temperature and *trp1-5* mutants also confers sensitivity to the DNA damaging agents MMS

and HU (Godin et al., 2016). The role of the aromatic amino acids and their biosynthetic

pathways in stress tolerance other than nutrient starvation is not fully understood, but it is well

established.

Reactive oxygen species (ROS) have previously been implicated as a source of toxicity in

cells treated with MCHM (Lan et al., 2015). Cells contain conserved robust networks to mitigate

the toxic effects of ROS. These include various proteins, from enzymes that detoxify the

reactive species directly, to proteins that repair damage within the cell, such as to DNA (Ayer et

al., 2014). The thioredoxin and glutathione (GSH) pathways have significant roles in the cell's

response to ROS. They perform overlapping functions reducing thiol oxidation that can damage

proteins in the cytosol. Furthermore, GSH has roles in iron homeostasis between the

mitochondria and vacuole, and potentially as a possible buffer for oxidation in disulfide bond

formation during protein folding in the ER (Cuozzo & Kaiser, 1999; reviewed by Toledano et al.,

2013). Mitochondria serve as a major producer of ROS in the cell as oxidative phosphorylation

leaks electrons to molecular oxygen to produce superoxide anions, so these pathways are

activated by normal cellular metabolism (reviewed by all the following, Ayer et al., 2014;

Perrone et al., 2008; Temple et al., 2005). However, they also become important during the

response to toxic chemicals, which can produce ROS directly, or else indirectly through

metabolism and attempted detoxification in the vacuole. Toxicity of chemicals that produce

ROS in the cell may be mitigated through treatment with antioxidants or intensified through

damage to the cellular stress networks (Couto et al., 2016; Sekito et al., 2014).

Much of a yeast cell's response to stress involves the vacuole (Li & Kane, 2010). This

organelle serves as a site for various processes of degradation, detoxification, and metal ion

and pH homeostasis (Li & Kane, 2010). The vacuolar ATPase (v-ATPase) is a structure in yeast

that is highly conserved and has been adapted to perform a wide range of functions in various

eukaryotes. In many animals, including *Drosophila*, homologs of the v-ATPase are known to

contribute to nerve function via vesicular excretion of neurotransmitters in animals including

humans (Hiesinger et al., 2005). In yeast, the v-ATPase is responsible for acidifying the vacuole

interior, creating a proton gradient that is responsible for multiple homeostatic processes

(reviewed by the following, N. Nelson & Harvey, 1999; Nishi & Forgac, 2002). As such, its effects

on metal ion transport and vacuolar acidification result in a set of knockout phenotypes for

many of the v-ATPase's subunits (Hemenway et al., 1995; H. Nelson & Nelson, 1990; Ohya et al.,

1991; Sambade et al., 2005). This *vma-* phenotype includes sensitivities to metal ions, reactive

oxygen species, and pH perturbations in either direction. Inositol depletion through mutation or

chemical treatment negatively impacts the assembly and activity of the v-ATPase in yeast

(Deranieh et al., 2015; Ohya et al., 1991). Any chemical that can inhibit or damage the v-ATPase would likely have distinct consequences for the cell's ability to cope with other stresses.

The goal of this study was to characterize the response of yeast cells to the foreign chemical MCHM. Treatment of yeast with MCHM created pleiotropic effects on networks throughout the cell. We employed methods including viability assays, RNAseq, flow cytometry, and a genetic screen of knockout strains to characterize these effects. We found several expected changes to networks, such as the pleiotropic drug response ABC transporters that remove xenobiotics from the cell. Our data show direct evidence of ROS and DNA damage following MCHM treatment. The stress also revealed a role for the aromatic amino acid biosynthetic pathway outside nutrient availability for the response to this particular stressor. Any toxicity that these cellular changes caused did not result in large changes in cellular viability but instead resulted in the arrest of the cell cycle in G1. Effects on the cell were wide-ranging, but wildtype cells were able to implement the ESR to recover from exposure at levels higher than spill levels.

Materials and Methods:

Yeast strains and media:

The haploid BY4741 strain (MATA, *his3Δ, leu2Δ, ura3Δ, met15Δ)* (Brachmann et al., 1998) and its MATalpha counterpart BY4742 (*his3Δ, leu2Δ, ura3Δ, lys2Δ*), were used for the majority of experiments as denoted. RNAseq, viability, growth, comet and flow cytometry assays used the BY4741 strain. The genetic screen used the BY4742 collection (Giaever et al., 2002). The YJM789 wildtype strain (MATalpha, *lys2*) and its previously generated *aro1*

26

knockout (MATa, *ARO1*::NAT[R])(Rong-Mullins, Ravishankar, et al., 2017) were used in serial

dilution growth assays only as shown. Rich media containing 1% yeast extract, 2% peptone, and

2% dextrose (YPD) or minimal media containing 0.67% yeast nitrogen base and 2% dextrose

(YM) with histidine, leucine, uracil, and methionine supplementation for BY4741 strains were

used in the various experiments as indicated. Experiments where the aromatic amino acids

were supplemented in excess to YPD media were done to a final concentration of 0.02mg/mL

tryptophan, 0.03mg/mL tyrosine, and 0.05mg/mL phenylalanine. The strains used in the *TAT2*

overexpression serial dilution assay were created by transforming BY4742 yeast with either

pRS315 (*CEN, LEU2*) plasmid (Sikorski & Hieter, 1989) or commercially available *LEU2+*

multicopy yeast genomic tiling collection plasmids containing genomic regions from

chromosome XV corresponding to the region surrounding the *TAT2* ORF (Jones et al., 2008).

Supplementations of MCHM, drugs, or other nutrients were added to YPD or YM media as

indicated for each experiment. MCHM for experiments is the same crude MCHM formulation

that made up the primary chemical presence in the Elk River spill. This formulation is

approximately 89% 4-methylcyclohexane methanol, with other cyclohexanes making up the

rest of the material was obtained directly from the manufacturer, Eastman Chemical Copmany

(Kingsport, TN, US). All concentrations of MCHM indicated in figures are in ppm of crude MCHM

formulation, so all conclusions are for this formulation, not pure MCHM.

Growth and viability assays:

Growth and viability assays were performed as previously described (Rong-Mullins,

Winans, et al., 2017). Plating for growth assays was only done at a maximum of five strains per

plate to keep cells in the central portion of the plate, due to noticed position effects on toxicity,

likely from evaporation, a known and observed characteristic of MCHM being its volatility (Phetxumphou et al., 2016; Sain et al., 2015). Furthermore, plates were used within a day to minimize any evaporation of MCHM from the media that would affect concentrations. Concentrations are noted in figure legends. When glutathione was used to attempt to rescue sensitive phenotypes, the concentration on plates was 10mM or 100mM, as noted. Oxidized glutathione was used at a concentration of 5mM corresponding to its structure as a dimer of glutathione. Ammonium sulfate was supplemented at a concentration of 10mM. Viability assays were carried out for 1.5 to 24 hours while kept in log-phase for concentrations indicated in figures.

Flow cytometry:

BY4741 cells were grown to saturation overnight and returned to mid-log phase. Cells were then diluted to a starting OD_{600} of 0.3 in biological triplicate in YPD media containing 550ppm MCHM. Cells were then harvested at the indicated time. Timepoints longer than 90 minutes required dilutions into fresh treatment media two times over the 12-hour experiment to maintain log-phase. The harvesting procedure involved taking 3mL of culture and fixing and preparing cells as previously published (Haase & Reed, 2002). For cell cycle analysis, cell pellets were fixed in 70% ethanol overnight at 4°C. Fixed cells were pelleted and resuspended in 0.5mL of a 2mg/mL RNAse A solution (50mM Tris pH 8.0, 15mM NaCl boiled for 15 minutes and cooled to room temperature) for a 12-hour incubation at 37°C. Cells were then incubated in 0.2mL of a protease solution (5mg/mL pepsin, 4.5µl/mL HCl) for 20 minutes at 37°C. Finally, samples were stored in 0.5mL of 50mM Tris pH 7.5 at 4°C for up to a week before undergoing flow cytometric analysis. Immediately prior to flow cytometry, samples were shortly sonicated at low power on

a Branson model SSE-1 Sonifier for six intervals of one-second bursts, and 50µl of each sample was suspended in 1µM SYTOX Green. Samples were analyzed on a BD LSRFortessa using a FITC channel. Approximately 30,000 events were collected for analysis. FCS Express 5.0 software was used to analyze the DNA content of cells using a multi-cycle DNA histogram and incorporated DNA modeling Statistics. Models were compared via Chi-squared results and the SL S0 model selected for determining the portion of the population in S phase for all samples. All replicates and timepoints fell within good or fair confidence for cell cycle modeling and S phase confidence of the model. The model produced the values for the proportions of the cell population falling within the G1, S, and G2/M phases used for analysis.

For measurement of ROS, live cells were pelleted then suspended in 200µl of 50mM DHE in phosphate-buffered saline (PBS). The dyed cultures were incubated at 30°C for 20 minutes and washed with PBS. A positive control sample of BY4741 cells was treated with 25mM H_2O_2 for 1.5 hours. The DHE dyed samples were then analyzed within 2 hours of harvesting on a BD LSRFortessa using preset propidium iodide detection defaults. Approximately 30,000 events were collected per sample for downstream analysis. A High ROS subpopulation range was determined by gating events in downstream analysis at a DHE fluorescence level of 350. This value gave all three biological BY4741 wildtype strain replicates treated as the H_2O_2 positive controls approximately 95% of their cells falling above this value (94.69% - 96.86% actual percentages of cells). The wildtype and *vma3* cells, both untreated and treated with MCHM, that fell above a fluorescence level of 350 were considered to be part of a High ROS subpopulation. A one-tailed t-test was used to determine significance of increases in subpopulation percentages for each strain.

Genetic screen:

The genetic screen of the BY4742 haploid knockout collection was performed via the phenotyping of serial dilution growth on solid media by adapting a previous screen technique (Bae et al., 2017). In short, frozen 96-well microtiter plates of the knockout collection were thawed and inoculated via pinning into 96-well growth chambers containing YPD in biological triplicate. Chambers were grown at 30°C for 2 days until all wells were saturated. Each well was then serially diluted into 96-well microtiter plates three times at 20-fold concentrations, for three wells at 20-fold, 400-fold, and 8000-fold dilution from saturated. Diluted plates were then pinned onto large YPD solid media plates with or without 300ppm MCHM. Plates were visually scored for growth at 2-3 days, comparing control, and treated plates for each replicate. Knockout strains showing decreased growth of at least one spot, unaccounted for by decreased growth rate in general on control media, were recorded as screen "hits". There was occasional variability in growth between replicates, so knockouts that did not replicate sensitive in all three trials filtered out.

RNAseq:

The transcriptomic dataset for this study was produced previously and published with full raw data output and a partial analysis (Gallagher et al., 2020). The raw data are available at GSE, accession number GSE129898 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE129898). A partial reanalysis for this paper included comparisons of specifically strains of the BY4741 untreated control and the BY4741 MCHM treated cells. The library prep, transcript quantification with salmon (v0.9.1),

bioinformatics pipeline, and differential analysis with DESeq2(1.18.1) were all performed exactly as in the previous study (Gallagher et al., 2020). GO Term analysis was performed with DAVID bioinformatic database as below for this study.

Comet Assay:

DNA damage resulting from MCHM exposure was quantitated using comet tails as previously described (Azevedo et al., 2011; Oliveira & Johansson, 2012). After incubation in MCHM for 30 minutes or hydrogen peroxide for 45 minutes (Heck et al., 2010; Rowe et al., 2008) as a positive control, single cells were solidified in 1.5% low-melting agarose. The cell walls were digested using 80 µl of 2mg/ ml zymolyase to create spheroplasts and immobilized on a glass slide. After cells were then lysed with 10 mM Tris HCl pH 10, 30 mM NaOH, 1 M NaCl, 50 mM EDTA and 0.05% w/v laurylsarcosine, they were rinsed three times with TAE, and nucleic acid was separated electrophoresis through the gel with 14V for 10 minutes. The gels were then neutralized Tris-HCl pH 7.4, and subsequently fixed with 76% and then 96% ethanol solutions to fix the DNA in place. DNA was stained with ethidium bromide for 20 minutes and then rinsed in water. The DNA fragments formed a tail from the rest of the DNA that has remained intact (the head). Images for each sample 17-42 'comets' were captured with a Nikon A1R confocal microscope and measured via pixel intensity and distance (µm) of the DNA migration through the gel using TriTek CometScore 2.0.0.38. Statistical analysis was performed on tail length measurements using a one-way ANOVA and Tukey's HSD test in base package of R version 3.6.3 (aov and TukeyHSD commands).

GO Term analysis:

The knockout screen gene list and the RNAseq differential analysis were each

analyzed further using the GO Term function of the DAVID bioinformatic database at

https://david.ncifcrf.gov (Huang et al., 2009a, 2009b). Default settings were used, and no

optional steps were included. Gene lists were searched for all three ontologies, though process

and function were determined to show the most informative results and included in the main

figures. Full results are available in supplemental files 1-3.

Data and Reagent availability statement

Raw data for transcriptome analysis are available at GSE, accession number GSE129898

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE129898). GO Term analysis results

are available in supplemental files 1-3. The full gene lists of genetic screen hits, significantly

upregulated genes, and significantly downregulated genes are available as Tables S1-S3.

Haploid BY4741 and BY4742 knockout collection strains and YJM789 *aro1* knockout are

available by request. The plasmids and/or resulting strains used in the *TAT2* overexpression

analysis are available by request, or available commercially from DHARMACON. Crude MCHM

used as the main reagent in the experiments was a gift from the Eastman Chemical Company,

but a limited supply is available upon request.

<div align="center">

Results

</div>

Determination of S. cerevisiae sensitivity to MCHM treatment.

S. cerevisiae were treated with a range of MCHM levels to determine minimal

concentrations that reduced growth in rich media (YPD). The levels were initially titrated, and

approximately 400ppm MCHM inhibited the growth of BY4741 yeast (Figure 1A). The

concentrations used on yeast in these growth assays (400-600ppm, 2.8-4.2mM) were

approximately 100 times greater than the highest levels recorded in the water distribution

system of Kanawha County, WV when monitoring began the day after the spill (Whelton et al.,

2014). While the growth assay phenotype itself requires a relatively high dosage of the

chemical, the clear phenotype of decreased growth of yeast at this concentration is informative

to the design of experiments that exploit the many genomic resources of the yeast. Similar

chemical assays of toxic chemicals, such as hydrogen peroxide, use a comparable range of

doses (0.5-5mM) to exploit the yeast growth phenotype as a measurement for cellular changes,

including the original papers establishing the transcriptional programming of the yeast

environmental stress response (0.3mM H_2O_2)(Azevedo et al., 2011; Gasch et al., 2000).

Decreased growth of yeast treated with MCHM is not due to large-scale cell death.

To assess whether MCHM inhibited growth by decreasing cellular viability or by

arresting growth, acute and chronic viability assays were completed (Figure 1B). Mid-log phase

yeast were acutely exposed for 90 min and chronically exposed for 24 hours to the 550ppm

MCHM. Yeast were then washed and plated onto YPD with no MCHM. After growing plates for

two days, there was little if any reduction in colony-forming units with the YPD-MCHM treated

yeast.

To determine if the reduction in growth is due to cell cycle arrest, we analyzed MCHM

treated yeast by flow cytometry (Figure 1C-1F). Untreated cellular populations showed an

asynchronous pattern, for which each of the three phases had upwards of 20% of the total

cellular population (Figure 1C and 1D). However, within 30 minutes, the MCHM treated

33

samples showed that less than 10% of the population was in S phase (Figure 1E and 1F). The cells treated with MCHM appear to arrest almost immediately in G1 phase, leading to a nearly instantaneous loss of S phase population as no new cells leave G1. Over the course of the 12-hour experiment, the population in G2 also decreased to less than 10% of cells (Figure 1E and 1F). MCHM treated yeast failed to grow because they arrested in G1, not because of decreased viability.

Genetic screen reveals cellular pathways and components required for MCHM tolerance

The cell cycle arrest of MCHM treated yeast was hypothesized to be attributable to environmental stress response programming, as many stressors initiate arrest as the cells attempt to ameliorate and recover from damage caused (Gasch et al., 2000). In order to test this hypothesis, we collected genomic datasets to analyze which genes were functionally required for MCHM tolerance in mutant strains and which were differentially expressed in wildtype cells. The nearly 5000 strains of the haploid BY4742 knockout collection were tested for growth phenotypes on 300ppm MCHM in a genetic screen (Supplemental Figure 1). The results revealed that 329 genes were required for the tolerance to MCHM treatment (Supplemental Table 1). Several individually important environmental stress response genes were revealed by the screen. Two genes of note for their role in the control of Msn2/4-regulated stress responses were *MCK1* and *MSN5*. Deletions of these two genes have been shown previously to affect the ability of Msn2 to control stress response programming (Gutin et al., 2019).

GO term analysis of the results of the screen pointed to mitochondrial translation as the most enriched subset of yeast genes (Figure 2A). This particular GO term agrees well with previously published data showing that petite yeast strains are sensitive to MCHM (Pupo, Ayers, et al., 2019). Mitochondrial function is important for dealing with many stressors, especially those involved in reactive oxygen species production. The next most enriched gene subset was the components of the vacuolar ATPase (v-ATPase). This transporter maintains the acidity of the vacuole by pumping H+ ions across the vacuolar membrane. The homeostatic processes that use this ion gradient are evidence that MCHM may be causing ROS stress(N. Nelson & Harvey, 1999; Nishi & Forgac, 2002).

The next most enriched GO terms were for amino acid biosynthesis, especially aromatic amino acid biosynthesis. As the screen was performed on YPD media, it was surprising that these biosynthetic genes, including early precursor producing enzymes (*ARO1* and *ARO3*) and nearly the entirety of the tryptophan biosynthetic pathway (*TRP2-5*), were required when presumably excess amino acids were available from the media.

MCHM treatment significantly affects the transcription of large portions of the yeast genome.

Transcriptomic analysis of the BY4741 strain treated with MCHM also supported the hypothesis that the environmental stress response was activated with exposure to MCHM. There were 592 significantly upregulated genes in response to MCHM (Supplemental Table 2 and Figure 2B), while there were 576 genes significantly downregulated (Supplemental Table 3 and Figure 2C). These lists of genes were analyzed for GO terms and several of the terms were consistent with the activation of the ESR (Figure 2B-C). The upregulated GO terms for drug

35

export, autophagy, and mitophagy point to ESR mechanisms for the removal of toxic substances from the cell and the reapportionment of cellular resources to overcome and adapt to a stressor. The upregulated terms related to amino acid biosynthesis and mitochondrial function (TCA cycle, iron ion homeostasis) were consistent with the genetic screen GO term analysis. The mitochondrial overlap between the datasets is likely due to the importance of ROS stress response, though the importance of the amino acid pathways as mentioned above, especially considering the cells are growing in rich media, is unclear.

The clearest indication for the central role of ESR activation by MCHM exposure was the downregulated GO terms. Ribosome biogenesis and related functions (other ribosomal terms were also significant as seen in Supplemental Table 3 but omitted from the figure for legibility) were the primary GO terms found in this dataset. They are also classic pathways downregulated in the ESR (Gasch et al., 2000). Most ribosomal biogenesis protein genes are essential and do not appear in the haploid knockout collection used in the genetic screen, with 81% of the 79 ribosomal protein genes essential for growth in standard conditions (Steffen et al., 2012). Therefore, the transcriptomic data was necessary supporting evidence for ESR involvement in MCHM response in combination with the stress gene and amino acid biosynthesis knockouts appearing in the screen. Other terms from the downregulated genes, including fatty acid biosynthesis and regulation of cell size, are also likely related to the observed cell cycle arrest and the reapportionment of energy and resources required for overcoming stressors.

Table 1: *List of genes overlapping between the genetic screen and upregulated or downregulated RNAseq results, as shown in Figure 2.* Genes are organized, with subheadings, to highlight functions relevant to environmental stress responses.

Overlap of Screen and Upregulated Genes	Overlap of Screen and Downregulated Genes
Aromatic Amino Acid Biosynthesis:	Aromatic Amino Acid Synthesis/Transport:
ARO1	*PRS1*
ARO3	*PRS3*
TRP2	*TAT1*
TRP3	
TRP4	RNA Processing/Ribosome Biogenesis:
TRP5	*MRT4*
	NRS1
Amino Acid Biosynthesis:	*RNR1*
ILV6	*RPA1*
LYS2	*TSR2*
MET22	
	Translation:
Mitochondrial Localization:	*BUD27*
GGC1	*FUN12*
POS5	
TGL2	Protein Chaperone:
	YDJ1
ABC Multi-drug Transporter:	
SNQ2	Transcriptional Regulator of Stress Response:
	MOT3
Transcription of rDNA:	

RRN10	Ion Transporter/Metal Homeostasis:
	PHO84
Miscellaneous Function:	
MOG1	Miscellaneous Function:
XYL2	BUD30
VVS1 (YBR241C)	GUA1
TVS1 (YCR061W)	PRM7
YHI9	SAM1

Combined genomic datasets implicate nutrient starvation

The genomic datasets individually revealed pathways involved in MCHM response, such as ESR activation from the transcriptome and vATPase and mitochondrial mutant sensitivity in the genetic screen. We decided to compare the genetic screen and transcriptome datasets to identify shared genes that might be especially important to MCHM response. The upregulated genes found in the knockout screen included those involved in mitochondrial function and ROS stress responses (Figure 2D and Table 1). *GGC1* and *TGL2* encode mitochondrial proteins involved in GTP/GDP transport and triacylglycerol lipase activity at the mitochondria respectively (Ham et al., 2010; Vozza et al., 2004). The *GGC1* knockout in particular produces petite yeast which are known to be sensitive to MCHM (Pupo, Ayers, et al., 2019; Vozza et al., 2004). The *POS5* gene, identified originally for conferring peroxide sensitivity, is a mitochondrial protein that acts as an NADH kinase (Krems et al., 1995; Strand et al., 2003). It is required for maintenance of mitochondrial DNA, the loss of which results in the aforementioned petite yeast, as well as detoxification of ROS in the mitochondria (Strand et al., 2003). The presence of

these genes in both datasets points to the importance of mitochondrial function in the response to MCHM stress, most likely due to ROS production.

The other major insight of the combined screen and upregulation datasets is the presence of several amino acid biosynthetic genes. Many of these genes fall under the control of the general amino acid activator Gcn4 (Braus, 1991). Several single genes from pathways for isoleucine/valine (*ILV6*), methionine (*MET22*), and lysine (*LYS2*) were found in the overlap. Furthermore, the aromatic amino acid biosynthetic genes *ARO1* and *ARO3*, as well as four out of five tryptophan biosynthetic genes (*TRP2, TRP3, TRP4,* and *TRP5*), were found in both datasets. As tryptophan pathways in particular, but also tyrosine, have previously been implicated in the response to multiple types of stress including nutrient starvation, pH, metal ions, SDS, and DNA damaging chemicals, the precise role of *TRP* genes in MCHM tolerance remains to be elucidated (Godin et al., 2016; González et al., 2008; Schroeder & Ikui, 2019).

The downregulated dataset overlaps with the screen also revealed genes involved in nutrient homeostasis (Figure 2E). *PRS1*, *PRS3*, and *TAT1* provide more evidence for the significance of tryptophan and amino acid metabolism in MCHM sensitivity. *PRS1* and *PRS3* encode 5-phospho-ribosyl-1(alpha)-pyrophosphate (PRPP) synthetase genes which are necessary for the production of purines and pyrimidines, and histidine and tryptophan amino acids (A. T. Carter et al., 1997; Andrew T. Carter et al., 1994; Hernando et al., 1999). While there are five paralog PRPP synthetases in yeast (Prs1-5), different multimeric complexes of the paralogous proteins are likely involved in producing wildtype levels of PRPP. Disruptions in *PRS1* and *PRS3* have the largest reduction in PRPP pools for nucleotide and amino acid production (Hernando et al., 1999). The protein product of the *TAT1* gene is an amino acid transporter,

characterized originally for its high-affinity import of tyrosine, but also shown to be a low-affinity transporter of tryptophan (Schmidt et al., 1994). The combined screen and transcriptomic profiles indicated that MCHM may cause yeast cells to undergo nutrient starvation, possibly nitrogen specifically, despite growing in rich media. Cells may be responding by initiating cell cycle arrest and reapportioning nitrogen resources until recovered.

The importance of the aromatic amino acid synthetic genes in MCHM tolerance was investigated by additional supplementation of tryptophan, tyrosine, and phenylalanine in rich media containing MCHM. This was insufficient to rescue the growth of either wildtype or pathway gene knockout strains in MCHM (Figure 3A). It is possible that the excess tryptophan was not sufficiently imported into the cells based on the sorting of transporters such as Tat2 to the vacuole instead of the membrane (Schroeder & Ikui, 2019). This is plausible as ergosterol levels in the cell affect this sorting preference, and MCHM is known to alter sterol levels in the cell (Pupo, Ku, et al., 2019; Umebayashi & Nakano, 2003). Overexpression of *TAT2* using a multicopy tiling plasmid did not rescue growth on MCHM with or without supplemental aromatic amino acids (Supplemental Figure 2). The levels of several amino acids increase with MCHM treatment, including the levels of tyrosine (Pupo, Ku, et al., 2019). Unfortunately, the levels of tryptophan were not measured by GC-MS, so it is unknown if levels change with MCHM treatment. Yeast cells seem to be experiencing a nutrient starvation signal upon exposure, reacting by increasing production of amino acids in a way that is required according to knockouts from the screen, and yet failing to adapt to the chemical, based on the presence of excess amino acids alone.

Uncharacterized ORFs both upregulated by and required for MCHM responses.

Of note in the combined datasets is the requirement of *YBR241c* for tolerance to

MCHM (Table 1). Ybr241c is a paralog of the vacuolar sorting protein, Vps73, and is localized to

the vacuolar membrane (Matsumoto et al., 2013). Acidification of the vacuole is critical for

metabolizing chemicals such as MCHM and Ybr241c relocalizes to the cytoplasm when

acidification of the vacuole is blocked (Matsumoto et al., 2013). Ybr241c also physically

interacts with Gtt1, a glutathione transferase of vacuolar proteins (Chandel et al., 2016; Yu et

al., 2008). Like Vps73, Ybr241c shares homology with other sugar transporters such as Hxt1-17

and related proteins. As Ybr241 is a <u>v</u>acuolar protein <u>v</u>ital for <u>s</u>tress response, we named it

Vvs1. Ycr061w localizes to cytoplasmic puncta that do not appear to be vacuoles but instead

vesicles (Huh et al., 2003; Tkach et al., 2012). Ycr061w contains 10-11 predicted

transmembrane domains (Weill et al., 2019). As Ycr061w is a <u>t</u>ransmembrane protein <u>v</u>ital for

<u>s</u>tress response, we named it Tvs1.

Growth rescue by glutathione.

As mitochondrial processes and iron ion homeostasis were found to be significant

results in the genomic datasets, we further explored the role of reactive oxygen species in

MCHM-induced stress. Glutathione serves as an antioxidant for the cell, particularly in the

mitochondria and ER, and is also important in iron-sulfur homeostasis (Bulteau et al., 2012). It

was hypothesized that treatment with glutathione may rescue ROS stress created by exposure

to MCHM, so yeast were co-treated with the chemicals. Glutathione was sufficient to rescue

the growth of wildtype yeast, even at extremely high doses of 600ppm (Figure 3B). Tryptophan

levels are important for resistance to DNA damaging agents (Godin et al., 2016). However,

glutathione was not able to rescue the sensitivity of the *aro1* mutant, so the unknown non-

amino acid role in MCHM tolerance of the aromatic amino acid biosynthetic genes does not seem to be related to an ROS stress mechanism. The sensitivity of the *aro1* knockout mutant was replicated in a divergent yeast strain background, YJM789 (Figure 3B), to control for genetic background specific sensitivity.

The yeast vacuolar ATPase was the second most significant GO term from the genetic screen dataset. The v-ATPase is responsible for acidifying the vacuole of yeast cells by hydrolyzing ATP in order to pump hydrogen ions into the vacuole, creating an electrochemical gradient with multiple homeostatic roles (Deranieh et al., 2015; H. Nelson & Nelson, 1990; N. Nelson & Harvey, 1999; Nishi & Forgac, 2002; Ohya et al., 1991). *vma* mutants show at least three sensitivity phenotypes that may be related to MCHM's effects, including metal ion homeostasis, pH sensitivity, and oxidative stress sensitivity. To test if oxidative stress may be responsible for the sensitivity of *vma* mutants, we co-treated these mutants with MCHM and glutathione (Figure 3C). The observed growth rescue with glutathione treatment supports the hypothesis that the v-ATPase is required to provide a robust level of ROS protection yeast cells need to survive MCHM.

The antioxidant glutathione is a tripeptide of glycine, glutamate, and cysteine, and therefore, it is also potentially a source of nitrogen for cells. This complicates the interpretation of its role as an antioxidant source in rescue assays, given the role of nutrient deprivation upon exposure to MCHM. We tested if glutathione was acting solely as an antioxidant by treating the *glr1* knockout mutant with MCHM and oxidized glutathione. Glr1 is the glutathione reductase required to convert oxidized glutathione to reduced glutathione that can act as an antioxidant (Couto et al., 2016). The oxidized glutathione was able to rescue the wildtype BY4742 strain and

42

glr1 mutant as well as reduced glutathione (Supplemental Figure 3). As the *glr1* mutant was rescued, the role of glutathione as antioxidant is called into question. We also treated cells with ammonium sulfate to test rescue by a general nitrogen source supplement, but the combination with MCHM was surprisingly toxic to cells (Supplemental Figure 3). As supplementation by aromatic amino acids, another nitrogen source, also failed to rescue growth, it is clear that if glutathione is acting as a nitrogen source instead of an antioxidant, it is unique among nitrogen supplements in its ability to alleviate the MCHM growth defect in cells.

Biochemical assays for the presence of ROS in MCHM treated cells.

As genomic datasets and mutant growth assays had implicated ROS production as an effect of MCHM treatment, we attempted to directly detect ROS inside the cell. The dye dihydroethidium (DHE) reacts with intracellular ROS and then creates a fluorescent signal. We dyed yeast treated with MCHM with DHE, then detected the changes in fluorescent signal in the populations via flow cytometry. The flow cytometry results revealed an increase in fluorescence for a subpopulation of wildtype cells (Figure 4A). Within the genetic clonal population of MCHM treated cells represented by the red curve, phenotypic heterogeneity existed. The majority of MCHM treated cells did not have an increase in ROS levels, in contrast to hydrogen peroxide treated yeast (gray line). However, a subpopulation of the red curve, as seen in the portion of the curve under the marker labeled High ROS, did have increased ROS. The High ROS subpopulation of cells increased from 2.96% to 7.82% of the total cellular population in the BY4741 strain (Supplemental Figure 4). It is a known phenomenon that clonal populations in liquid culture can show phenotypic heterogeneity, specifically with respect to ROS sensitivity (Sumner et al., 2003). This may explain why there is little reduction in the

viability of BY4741 cells seen in Figure 1B. If the cells that successfully arrest their cell cycle were able to adapt and minimize their ROS production to avoid MCHM-induced death, they may appear as the large subpopulation with a lower ROS signal in the data. The cells that failed to do so may represent both the few cells that die in the viability assay, as well as those appearing as the small peak in the High ROS range. This assay revealed the first direct evidence that MCHM treatment produces ROS in the cell, in agreement with the indirect evidence originally seen in the sensitivity of *vma* mutants, glutathione rescue, and previous studies involving petite yeast (Pupo, Ayers, et al., 2019).

While the effects of MCHM on yeast point to an increase in ROS stress for the cell, wildtype cells seem to be robust enough, on average, to limit this stress, while mutants in certain pathways cannot. To test this hypothesis, we also performed the DHE assay on one of the most sensitive knockout mutant strains, *vma3*, which previously showed rescued growth with glutathione treatment. This mutant strain showed a similar pattern of DHE fluorescence when treated with MCHM, with the appearance of a High ROS population peak (Figure 4B). The major difference from the wildtype was an increase in the size of the High ROS population peak. The High ROS population in the *vma3* strain rose from 1.08% to 22.19% of all cells (Supplemental Figure 4). The mutant strain also appeared to show a right-shift of the background ROS populations as well. This may indicate *vma* mutants have higher ROS levels in unstressed conditions. Therefore, these strains have an innate sensitivity to MCHM due to their inability to maintain ROS homeostasis. We hypothesize that robust cells were able to detect the increased ROS caused by MCHM stress, then adapt to this by arresting growth and limiting their background ROS production.

MCHM treatment causes DNA damage to wildtype yeast cells.

We also questioned whether the ROS produced in MCHM treated cells may cause DNA damage, a known complication in ROS stressed cells (reviewed in Ayer et al., 2014; Brennan et al., 1994; Lafleur & Retèl, 1993; Perrone et al., 2008; Temple et al., 2005). To address this question, a comet assay was performed to directly measure damage. This assay uses microscopy to measure the length of "tails", resembling comets, released from spheroplasted and lysed cells, and migrated across agarose covered slides via electrophoresis (Figure 4C-F). Longer tails indicate DNA fragments, presumably from damage resulting in double-stranded breaks. Yeast cells treated with MCHM concentrations of 550 ppm showed significantly longer comet tails as compared to no treatment and no spheroplast controls (Tukey's test p adj = 0.0437 and 0.000229 respectively) and approached significance as compared to 5mM H_2O_2 positive controls (p adj = 0.0655). This is the first direct evidence of DNA damage from MCHM treatment, confirming implications of data such as the ROS assay, cell cycle arrest, and other stress response data in previous studies (Lan et al., 2015).

Discussion:

The evidence presented here supports a model where MCHM activates the general environmental stress response programming of the yeast cell via Msn2/4 transcription factors, including specific responses consistent with ROS stress and nutrient starvation. Cell cycle arrest begins within 30 minutes of treatment, likely due to the initial signaling of nutrient starvation. These signals inactivate growth mechanisms soon after the cell is exposed, but before ROS accumulate. The ROS stress accumulates over six to twelve hours and the asynchronous

population of cells continues to arrest in G1 during this time period. The long duration of the G1 arrest in MCHM treated cells is similar to that seen in nutrient starvation (Gasch et al., 2000; Gasch & Werner-Washburne, 2002). It is in contrast to other stresses such as oxidative stress and heat shock, which are more transient, with growth resuming after a short acclimation period to the stress. After a long acclimation period, robust strains such as wildtype were able to manage the stress, recover, and return to proliferation (Figure 5) at dosage levels significantly higher than those found in the Elk River spill. In this model, several signals initiate multiple stress responses, including nutrient starvation, reactive oxygen species, and DNA damage itself, but the nutrient starvation accounts for the most likely primary stress controlling the ESR and G1 arrest.

Nonetheless, there is direct evidence of ROS production and DNA damage, as well as rescue by treatment with glutathione which may be acting as either an antioxidant or a nitrogen source. Other data have previously indicated that petite yeast, known to be sensitive to ROS, are also sensitive to MCHM, and this agrees with the sensitivity of the mitochondrial translation mutants revealed in the genetic screen data of this study (Pupo, Ayers, et al., 2019). The vacuolar ATPase mutants, which are sensitive to metal and pH stress, as well as ROS, were one of the other most enriched GO terms from the screen. This study, combined with previous data, points to MCHM as a source of perturbation for metal ions and ROS (Lan et al., 2015; Pupo, Ayers, et al., 2019).

The list of overlapping genes from the screen and the upregulated datasets included *SNQ2*, an ABC multidrug transporter associated with the removal of foreign substances from the cell (Figure 2D) (Decottignies et al., 1995; Decottignies & Goffeau, 1997; Rogers et al., 2001;

Servos et al., 1993). While there are several such transporters (and related transcription factors) found in the transcriptomics, this is the only one found in the screen to be required for the tolerance of MCHM. We hypothesize that this transporter is responsible for the removal of MCHM or a related toxic byproduct from the cell. Future work on MCHM tolerance should include exploration of the relative roles of different ABC transporters in export of the chemical.

Another possible avenue for study is control of Gcn4 and its related genes in the nutrient starvation signal detected during MCHM exposure. As previously noted, the amino acid biosynthetic genes *ILV6, MET22, LYS2, ARO1, ARO3, TRP2, TRP3, TRP4,* and *TRP5* are all under Gcn4 control and are required for and upregulated in MCHM tolerance (Figure 2D and Table 1). The screen overlap with the downregulated dataset also includes several Gcn4-related genes (Figure 2E and Table 1). *BUD27* is related to translation initiation and may be involved in the expression of genes controlled by Gcn4 (Deplazes et al., 2009; Gstaiger et al., 2003; Mirón-García et al., 2013). *GUA1* is negatively regulated by nutrient starvation and its mutants have been shown to be impaired in *GCN4* translation (Escobar-Henriques et al., 2003; Escobar-Henriques & Daignan-Fornier, 2001; Iglesias-Gato et al., 2011). *PRM7* also contains Gcn4 binding elements, connecting it to control by this transcriptional activator (Schuldiner et al., 1998). *GCN4* was not upregulated upon MCHM treatment at the transcriptional level. However, it is also controlled translationally, so it could have been upregulated at the protein level (Hinnebusch, 1997; Meimoun et al., 2000). It was also not required for tolerance in the genetic screen despite the importance the amino acid biosynthetic genes. Exploration of the background levels of expression of these genes in the *gcn4* knockout mutant could reveal if that is sufficient to perform their required functions in MCHM resistance.

47

Evidence of nutrient starvation is supported by the overlapping genomic datasets, as well as previous work which revealed that several amino acid levels increase in the cell (Pupo, Ku, et al., 2019). Glutathione, in both its reduced and oxidized forms, were able to rescue growth, indicating that it may be acting as a nitrogen source instead of an antioxidant. However, treatment with excess aromatic amino acids nor general nitrogen sources such as ammonium sulfate could rescue growth. The cell's programmed response to nutrient starvation is initiated by the cell to respond to MCHM, as wildtype cells upregulate the cellular and aromatic amino acid biosynthetic pathways. Furthermore, mutants lacking genes in these pathways, those in aromatic amino acids and tryptophan, in particular, are more susceptible to MCHM than wildtype cells, even in rich media with excess aromatic amino acids available.

The major question remains why the aromatic amino acid biosynthesis genes, including all but one of the tryptophan pathway genes, are required for resistance to the chemical. Supplementation with excess tryptophan in YPD did not rescue growth on MCHM, but it is not clear if this supplementation increased intracellular levels of tryptophan. It is possible the sorting of the transporter Tat2 to the plasma membrane was insufficient to import excess tryptophan. Furthermore, GC-MS failed to measure intracellular levels of tryptophan. While overexpression of *TAT2* was previously sufficient to rescue SDS sensitivity in *mck1* knockout mutant cells(Schroeder & Ikui, 2019), a multicopy plasmid expressing the *TAT2* region of chromosome XV was not able to rescue growth on MCHM. Another hypothesis is that the biosynthetic genes themselves are important for signaling recovery from stress so that cells may resume proliferation, regardless of the actual levels of aromatic amino acid products related to the pathway. The *trp1-5* mutants are sensitive to rapamycin (González et al., 2008),

so it is possibly a recovery phenotype. The chemical glutathione can rescue MCHM treatment in wildtype cells, but it cannot do so in these mutants. Whether glutathione is ameliorating ROS stress or providing a specific nitrogen supplement when it rescues wildtype cells, it cannot do the same for these mutants. They may be required to perform a signaling role in the recovery after stress acclimation. Another possibility is that there is a separate source of stress besides nutrient availability or ROS that the genes are required to address, such as production or conversion of other small molecules. It would be useful to determine if there are other stressors for which these genes are required, regardless of amino acid levels, to address the possibility that this is specific to MCHM exposure.

The Elk River MCHM spill revealed that there are many things we do not know about the industrial chemicals that we live among. This and other studies since have revealed that levels of the chemical may not be acutely toxic, but there are immense and still only partially characterized effects on the biochemical pathways that control eukaryotic metabolism. This chemical causes the production of reactive oxygen species and DNA damage in yeast and alters the transcriptome and metabolome in extensive ways. While the effects on possible nutrient starvation and associated amino acid biosynthesis are not fully understood, cells with robust homeostatic processes are generally able to activate the environmental stress responses programmed in their genes to recover from MCHM exposure and resume proliferation.

Acknowledgments

Author Contributions

MCA designed and carried out yeast experiments and wrote the paper. ZNS performed comet assay. JEGG designed the study, supervised the project, and co-wrote the paper.

Competing Financial Interests

The authors declare no competing interests.

References:

Arndt, K., & Fink, G. R. (1986). GCN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5' TGACTC 3' sequences. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(22), 8516–8520.

Ayer, A., Gourlay, C. W., & Dawes, I. W. (2014). Cellular redox homeostasis, reactive oxygen species and replicative ageing in Saccharomyces cerevisiae. *FEMS Yeast Research*, *14*(1), 60–72. https://doi.org/10.1111/1567-1364.12114

Azevedo, F., Marques, F., Fokt, H., Oliveira, R., Orn Johansson, B., & Johansson, B. (2011). Measuring oxidative DNA damage and DNA repair using the yeast comet assay. *Yeast*, *28*, 55–61. https://doi.org/10.1002/yea.1820

Bae, N. S., Seberg, A. P., Carroll, L. P., & Swanson, M. J. (2017). Identification of Genes in *Saccharomyces cerevisiae* that Are Haploinsufficient for Overcoming Amino Acid Starvation. *G3&#58; Genes|Genomes|Genetics*, *7*(4), 1061–1084. https://doi.org/10.1534/g3.116.037416

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, *14*(2), 115–132. https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2

Braus, G. H. (1991). Aromatic amino acid biosynthesis in the yeast Saccharomyces cerevisiae: A model system for the regulation of a eukaryotic biosynthetic pathway. *Microbiological Reviews*, *55*(3), 349–370.

Brennan, R. J., Swoboda, B. E. P., & Schiestl, R. H. (1994). Oxidative mutagens induce intrachromosomal recombination in yeast. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *308*(2), 159–167. https://doi.org/10.1016/0027-5107(94)90151-1

Bulteau, A. L., Planamente, S., Jornea, L., Dur, A., Lesuisse, E., Camadro, J. M., & Auchère, F. (2012). Changes in mitochondrial glutathione levels and protein thiol oxidation in Δyfh1 yeast cells and the lymphoblasts of patients with Friedreich's ataxia. *Biochimica et Biophysica Acta*, *1822*(2), 212–225. https://doi.org/10.1016/j.bbadis.2011.11.003

Capaldi, A. P., Kaplan, T., Liu, Y., Habib, N., Regev, A., Friedman, N., & O'Shea, E. K. (2008). Structure and function of a transcriptional network activated by the MAPK Hog1. *Nature Genetics*, *40*(11), 1300–1306. https://doi.org/10.1038/ng.235

Carter, A. T., Beiche, F., Hove-Jensen, B., Narbad, A., Barker, P. J., Schweizer, L. M., & Schweizer, M. (1997). PRS1 is a key member of the gene family encoding phosphoribosylpyrophosphate synthetase in Saccharomyces cerevisiae. *Molecular and General Genetics MGG*, *254*(2), 148–156. https://doi.org/10.1007/s004380050402

Carter, Andrew T., Narbad, A., Pearson, B. M., Beck, K.-F., Baum, B., Logghe, M., Contreras, R., & Schweizer, M. (1994). Phosphoribosylpyrophosphate synthetase (PRS): A new gene family in Saccharomyces cerevisiae. *Yeast*, *10*(8), 1031–1044. https://doi.org/10.1002/yea.320100805

Chandel, A., Das, K. K., & Bachhawat, A. K. (2016). Glutathione depletion activates the yeast vacuolar transient receptor potential channel, Yvc1p, by reversible glutathionylation of specific cysteines. *Molecular Biology of the Cell*, *27*(24), 3913–3925. https://doi.org/10.1091/mbc.E16-05-0281

Couto, N., Wood, J., & Barber, J. (2016). The role of glutathione reductase and related enzymes on cellular redox homoeostasis network. *Free Radical Biology and Medicine*, *95*, 27–42. https://doi.org/10.1016/j.freeradbiomed.2016.02.028

Cuozzo, J. W., & Kaiser, C. A. (1999). Competition between glutathione and protein thiols for disulphide-bond formation. *Nature Cell Biology*, *1*(3), 130–135. https://doi.org/10.1038/11047

De Wever, V., Reiter, W., Ballarini, A., Ammerer, G., & Brocard, C. (2005). A dual role for PP1 in shaping the Msn2-dependent transcriptional response to glucose starvation. *The EMBO Journal*, *24*(23), 4115–4123. https://doi.org/10.1038/sj.emboj.7600871

Decottignies, A., & Goffeau, A. (1997). Complete inventory of the yeast ABC proteins. *Nature Genetics*, *15*(2), 137–145. https://doi.org/10.1038/ng0297-137

Decottignies, A., Lambert, L., Catty, P., Degand, H., Epping, E. A., Moye-Rowley, W. S., Balzi, E., & Goffeau, A. (1995). Identification and Characterization of SNQ2, a New Multidrug ATP Binding Cassette Transporter of the Yeast Plasma Membrane. *Journal of Biological Chemistry*, *270*(30), 18150–18157. https://doi.org/10.1074/jbc.270.30.18150

Deplazes, A., Möckli, N., Luke, B., Auerbach, D., & Peter, M. (2009). Yeast Uri1p promotes translation initiation and may provide a link to cotranslational quality control. *The EMBO Journal*, *28*(10), 1429–1441. https://doi.org/10.1038/emboj.2009.98

Deranieh, R. M., Shi, Y., Tarsio, M., Chen, Y., McCaffery, J. M., Kane, P. M., & Greenberg, M. L. (2015). Perturbation of the vacuolar ATPase: A novel consequence of inositol depletion. *Journal of Biological Chemistry*, *290*(46), 27460–27472. https://doi.org/10.1074/jbc.M115.683706

Escobar-Henriques, M., Collart, M. A., & Daignan-Fornier, B. (2003). Transcription Initiation of the Yeast IMD2 Gene Is Abolished in Response to Nutrient Limitation through a Sequence in Its Coding Region. *Molecular and Cellular Biology*, *23*(17), 6279–6290. https://doi.org/10.1128/MCB.23.17.6279-6290.2003

Escobar-Henriques, M., & Daignan-Fornier, B. (2001). Transcriptional Regulation of the Yeast GMP Synthesis Pathway by Its End Products. *Journal of Biological Chemistry*, *276*(2), 1523–1530. https://doi.org/10.1074/jbc.M007926200

Gallagher, J. E. G., Ser, S. L., Ayers, M. C., Nassif, C., & Pupo, A. (2020). The Polymorphic PolyQ Tail Protein of the Mediator Complex, Med15, Regulates the Variable Response to Diverse Stresses. *International Journal of Molecular Sciences*, *21*(5). https://doi.org/10.3390/ijms21051894

Garmendia-Torres, C., Goldbeter, A., & Jacquet, M. (2007). Nucleocytoplasmic Oscillations of the Yeast Transcription Factor Msn2: Evidence for Periodic PKA Activation. *Current Biology*, *17*(12), 1044–1049. https://doi.org/10.1016/j.cub.2007.05.032

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., &

Brown, P. O. (2000). Genomic Expression Programs in the Response of Yeast Cells to

Environmental Changes☐D. *Molecular Biology of the Cell*, *11*, 17.

Gasch, A. P., & Werner-Washburne, M. (2002). The genomics of yeast responses to

environmental stress and starvation. *Functional & Integrative Genomics*, *2*(4), 181–192.

https://doi.org/10.1007/s10142-002-0058-2

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A.,

Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R.,

Brachat, S., Campanaro, S., Curtiss, M., Davis, K., … Johnston, M. (2002). Functional profiling of

the Saccharomyces cerevisiae genome. *Nature*, *418*(6896), 387–391.

https://doi.org/10.1038/nature00935

Godin, S. K., Lee, A. G., Baird, J. M., Herken, B. W., & Bernstein, K. A. (2016). Tryptophan

biosynthesis is important for resistance to replicative stress in Saccharomyces cerevisiae. *Yeast*

(Chichester, England), *33*(5), 183–189. https://doi.org/10.1002/yea.3150

González, A., Larroy, C., Biosca, J. A., & Ariño, J. (2008). Use of the TRP1 auxotrophic marker for

gene disruption and phenotypic analysis in yeast: A note of warning. *FEMS Yeast Research*, *8*(1),

2–5. https://doi.org/10.1111/j.1567-1364.2007.00315.x

Gorner, W., Durchschlag, E., Martinez-Pastor, M. T., Estruch, F., Ammerer, G., Hamilton, B.,

Ruis, H., & Schuller, C. (1998). Nuclear localization of the C2H2 zinc finger protein Msn2p is

regulated by stress and protein kinase A activity. *Genes & Development*, *12*(4), 586–597.

https://doi.org/10.1101/gad.12.4.586

Gstaiger, M., Luke, B., Hess, D., Oakeley, E. J., Wirbelauer, C., Blondel, M., Vigneron, M., Peter, M., & Krek, W. (2003). Control of Nutrient-Sensitive Transcription Programs by the Unconventional Prefoldin URI. *Science*, *302*(5648), 1208–1212. https://doi.org/10.1126/science.1088401

Gutin, J., Joseph-Strauss, D., Sadeh, A., Shalom, E., & Friedman, N. (2019). Genetic screen of the yeast environmental stress response dynamics uncovers distinct regulatory phases. *Molecular Systems Biology*, *15*(8). https://doi.org/10.15252/msb.20198939

Gutin, J., Sadeh, A., Rahat, A., Aharoni, A., & Friedman, N. (2015). Condition-specific genetic interaction maps reveal crosstalk between the cAMP/PKA and the HOG MAPK pathways in the activation of the general stress response. *Molecular Systems Biology*, *11*(10), 829. https://doi.org/10.15252/msb.20156451

Haase, S. B., & Reed, S. I. (2002). Improved Flow Cytometric Analysis of the Budding Yeast Cell Cycle. *Cell Cycle*, *1*(2), 117–121. https://doi.org/10.4161/cc.1.2.114

Ham, H. J., Rho, H. J., Shin, S. K., & Yoon, H.-J. (2010). The TGL2 Gene of Saccharomyces cerevisiae Encodes an Active Acylglycerol Lipase Located in the Mitochondria. *The Journal of Biological Chemistry*, *285*(5), 3005–3013. https://doi.org/10.1074/jbc.M109.046946

Heck, D. E., Shakarjian, M., Kim, H. D., Laskin, J. D., & Vetrano, A. M. (2010). Mechanisms of oxidant generation by catalase. *Annals of the New York Academy of Sciences*, *1203*, 120–125. https://doi.org/10.1111/j.1749-6632.2010.05603.x

Hemenway, C. S., Kara, D., Cardenas, M. E., Hiller, M. A., Jones, E. W., & I, J. H. (1995). *Require Calcineurin for Growth*. *844*, 833–844.

Hernando, Y., Carter, A. T., Parr, A., Hove-Jensen, B., & Schweizer, M. (1999). Genetic Analysis and Enzyme Activity Suggest the Existence of More Than One Minimal Functional Unit Capable of Synthesizing Phosphoribosyl Pyrophosphate in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *274*(18), 12480–12487. https://doi.org/10.1074/jbc.274.18.12480

Hiesinger, P. R., Fayyazuddin, A., Mehta, S. Q., Rosenmund, T., Schulze, K. L., Zhai, R. G., Verstreken, P., Cao, Y., Zhou, Y., Kunz, J., & Bellen, H. J. (2005). The v-ATPase V0 subunit a1 is required for a late step in synaptic vesicle exocytosis in Drosophila. *Cell*, *121*(4), 607–620. https://doi.org/10.1016/j.cell.2005.03.012

Hill, D. E., Hope, I. A., Macke, J. P., & Struhl, K. (1986). Saturation mutagenesis of the yeast his3 regulatory site: Requirements for transcriptional induction and for binding by GCN4 activator protein. *Science*, *234*(4775), 451–457. https://doi.org/10.1126/science.3532321

Hinnebusch, A. G. (1997). Translational Regulation of Yeast GCN4 A WINDOW ON FACTORS THAT CONTROL INITIATOR-tRNA BINDING TO THE RIBOSOME. *Journal of Biological Chemistry*, *272*(35), 21661–21664. https://doi.org/10.1074/jbc.272.35.21661

Hirata, Y., Andoh, T., Asahara, T., & Kikuchi, A. (2002). Yeast Glycogen Synthase Kinase-3 Activates Msn2p-dependent Transcription of Stress Responsive Genes. *Molecular Biology of the Cell*, *14*(1), 302–312. https://doi.org/10.1091/mbc.e02-05-0247

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. https://doi.org/10.1038/nprot.2008.211

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

Huh, W., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., & Shea, E. K. O. (2003). Global analysis of protein localization in budding yeast. *Nature*, *425*, 686–691.

Iglesias-Gato, D., Martín-Marcos, P., Santos, M. A., Hinnebusch, A. G., & Tamame, M. (2011). Guanine Nucleotide Pool Imbalance Impairs Multiple Steps of Protein Synthesis and Disrupts GCN4 Translational Control in Saccharomyces cerevisiae. *Genetics*, *187*(1), 105–122. https://doi.org/10.1534/genetics.110.122135

Jones, G. M., Stalker, J., Humphray, S., West, A., Cox, T., Rogers, J., Dunham, I., & Prelich, G. (2008). A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nature Methods*, *5*(3), 239–241. https://doi.org/10.1038/nmeth.1181

Krems, B., Charizanis, C., & Entian, K.-D. (1995). Mutants of Saccharomyces cerevisiae sensitive to oxidative and osmotic stress. *Current Genetics*, *27*(5), 427–434. https://doi.org/10.1007/BF00311211

Lafleur, M. V. M., & Retèl, J. (1993). Contrasting effects of SH-compounds on oxidative DNA damage: Repair and increase of damage. *Mutation Research/DNAging*, *295*(1), 1–10. https://doi.org/10.1016/0921-8734(93)90006-O

Lan, J., Hu, M., Gao, C., Alshawabkeh, A., & Gu, A. Z. (2015). Toxicity Assessment of 4-Methyl-1-cyclohexanemethanol and Its Metabolites in Response to a Recent Chemical Spill in West Virginia, USA. *Environmental Science & Technology*.

Li, S. C., & Kane, P. M. (2010). NIH Public Access. *Biochim Biophs Acta.*, *1793*(4), 650–663. https://doi.org/10.1016/j.bbamcr.2008.08.003.The

Martínez-Pastor, M. T., Marchler, G., Schüller, C., Marchler-Bauer, A., Ruis, H., & Estruch, F. (1996). The Saccharomyces cerevisiae zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *The EMBO Journal*, *15*(9), 2227–2235.

Matsumoto, R., Suzuki, K., & Ohya, Y. (2013). Organelle acidification is important for localisation of vacuolar proteins in Saccharomyces cerevisiae. *Protoplasma*, *250*(6), 1283–1293. https://doi.org/10.1007/s00709-013-0510-2

Meimoun, A., Holtzman, T., Weissman, Z., McBride, H. J., Stillman, D. J., Fink, G. R., & Kornitzer, D. (2000). Degradation of the Transcription Factor Gcn4 Requires the Kinase Pho85 and the SCFCDC4 Ubiquitin–Ligase Complex. *Molecular Biology of the Cell*, *11*(3), 915–927.

Mirón-García, M. C., Garrido-Godino, A. I., García-Molinero, V., Hernández-Torres, F., Rodríguez-Navarro, S., & Navarro, F. (2013). The Prefoldin Bud27 Mediates the Assembly of the

Eukaryotic RNA Polymerases in an Rpb5-Dependent Manner. *PLoS Genetics*, *9*(2). https://doi.org/10.1371/journal.pgen.1003297

Nelson, H., & Nelson, N. (1990). Disruption of genes encoding subunits of yeast vacuolar H(+)-ATPase causes conditional lethality. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(9), 3503–3507. https://doi.org/10.1073/pnas.87.9.3503

Nelson, N., & Harvey, W. R. (1999). Vacuolar and plasma membrane proton-adenosinetriphosphatases. *Physiological Reviews*, *79*(2), 361–385.

Nishi, T., & Forgac, M. (2002). the Vacuolar (H+)-Atpases—Nature'S Most Versatile Proton Pumps. *Nature Reviews Molecular Cell Biology*, *3*(2), 94–103. https://doi.org/10.1038/nrm729

Ohya, Y., Umemoto, N., Tanida, I., Ohta, A., Irda, H., & Anraku, Y. (1991). Calcium-sensitive cls mutants of Saccharomyces cerevisiae showing a Pet- phenotype are ascribable to defects of vacuolar membrane H+-ATPase activity. *Journal of Biological Chemistry*, *266*(21), 13971–13977.

Oliphant, A. R., Brandl, C. J., & Struhl, K. (1989). Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: Analysis of yeast GCN4 protein. *Molecular and Cellular Biology*, *9*(7), 2944–2949.

Oliveira, R., & Johansson, B. (2012). *Quantitative DNA Damage and Repair Measurement with the Yeast Comet Assay* (pp. 101–109). https://doi.org/10.1007/978-1-61779-998-3_8

Perrone, G. G., Tan, S.-X., & Dawes, I. W. (2008). Reactive oxygen species and yeast apoptosis. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1783*(7), 1354–1368. https://doi.org/10.1016/j.bbamcr.2008.01.023

Phetxumphou, K., Dietrich, A. M., Shanaiah, N., Smiley, E., & Gallagher, D. L. (2016). Subtleties of human exposure and response to chemical mixtures from spills. *Environmental Pollution*, *214*, 618–626. https://doi.org/10.1016/j.envpol.2016.04.056

Pupo, A., Ayers, M. C., Sherman, Z. N., Vance, R. J., Cumming, J. R., & Gallagher, J. E. G. (2019). MCHM Acts as a Hydrotrope, Altering the Balance of Metals in Yeast. *Biological Trace Element Research*. https://doi.org/10.1007/s12011-019-01850-z

Pupo, A., Ku, K. M., & Gallagher, J. E. G. (2019). Effects of MCHM on yeast metabolism. *PLOS ONE*, *14*(10), e0223909. https://doi.org/10.1371/journal.pone.0223909

Rogers, B., Decottignies, A., Kolaczkowski, M., Carvajal, E., Balzi, E., & Goffeau, A. (2001). The pleitropic drug ABC transporters from Saccharomyces cerevisiae. *Journal of Molecular Microbiology and Biotechnology*, *3*(2), 207–214.

Rong-Mullins, X., Ravishankar, A., McNeal, K. A., Lonergan, Z. R., Biega, A. C., Creamer, J. P., & Gallagher, J. E. G. (2017). Genetic variation in Dip5, an amino acid permease, and Pdr5, a multiple drug transporter, regulates glyphosate resistance in S. cerevisiae. *PLOS ONE*, *12*(11), e0187522. https://doi.org/10.1371/journal.pone.0187522

Rong-Mullins, X., Winans, M. J., Lee, J. B., Lonergan, Z. R., Pilolli, V. A., Weatherly, L. M., Carmenzind, T. W., Jiang, L., Cumming, J. R., Oporto, G., & Gallagher, J. (2017). Proteomic and genetic analysis of S. cerevisiae response to soluble copper leads to improvement of antimicrobial function of cellulosic copper nanoparticles. *Metallomics*, *9*(9), 1304–1315. https://doi.org/10.1039/C7MT00147A

Rowe, L. A., Degtyareva, N., & Doetsch, P. W. (2008). DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. *Free Radic Biol Med*, *45*(8), 1167–1177. https://doi.org/10.1016/j.freeradbiomed.2008.07.018

Sain, A. E., Dietrich, A. M., Smiley, E., & Gallagher, D. L. (2015). Assessing human exposure and odor detection during showering with crude 4-(methylcyclohexyl)methanol (MCHM) contaminated drinking water. *Science of The Total Environment*, *538*, 298–305. https://doi.org/10.1016/j.scitotenv.2015.08.050

Sambade, M., Alba, M., Smardon, A. M., West, R. W., & Kane, P. M. (2005). A genomic screen for yeast vacuolar membrane ATPase mutants. *Genetics*, *170*(4), 1539–1551. https://doi.org/10.1534/genetics.105.042812

Santhanam, A., Hartley, A., Düvel, K., Broach, J. R., & Garrett, S. (2004). PP2A Phosphatase Activity Is Required for Stress and Tor Kinase Regulation of Yeast Stress Response Factor Msn2p. *Eukaryotic Cell*, *3*(5), 1261–1271. https://doi.org/10.1128/EC.3.5.1261-1271.2004

Schmidt, A., Hall, M. N., & Koller, A. (1994). Two FK506 resistance-conferring genes in Saccharomyces cerevisiae, TAT1 and TAT2, encode amino acid permeases mediating tyrosine and tryptophan uptake. *Molecular and Cellular Biology*, *14*(10), 6597–6606. https://doi.org/10.1128/MCB.14.10.6597

Schroeder, L., & Ikui, A. E. (2019). Tryptophan confers resistance to SDS-associated cell membrane stress in Saccharomyces cerevisiae. *PLOS ONE*, *14*(3), e0199484. https://doi.org/10.1371/journal.pone.0199484

Schuldiner, O., Yanover, C., & Benvenisty, N. (1998). Computer analysis of the entire budding yeast genome for putative targets of the GCN4 transcription factor. *Current Genetics*, *33*(1), 16–20. https://doi.org/10.1007/s002940050303

Sekito, T., Sugimoto, N., Ishimoto, M., & Kawano-kawada, M. (2014). Tributyltin induces cell cycle arrest at G1 phase in the yeast Saccharomyces cerevisiae. *The Journal of Toxicological Sciences*, *39*(2), 311–317. https://doi.org/10.2131/jts.39.311

Servos, J., Haase, E., & Brendel, M. (1993). Gene SNQ2 of Saccharomyces cerevislae, which confers resistance to 4-nitroquinoline-N-oxide and other chemicals, encodes a 169 kDa protein homologous to ATP-dependent permeases. *Molecular and General Genetics MGG*, *236*(2), 214–218. https://doi.org/10.1007/BF00277115

Sikorski, R. S., & Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics*, *122*(1), 19–27.

Steffen, K. K., McCormick, M. A., Pham, K. M., MacKay, V. L., Delaney, J. R., Murakami, C. J., Kaeberlein, M., & Kennedy, B. K. (2012). Ribosome Deficiency Protects Against ER Stress in Saccharomyces cerevisiae. *Genetics*, *191*(1), 107–118. https://doi.org/10.1534/genetics.111.136549

Strand, M. K., Stuart, G. R., Longley, M. J., Graziewicz, M. A., Dominick, O. C., & Copeland, W. C. (2003). POS5 Gene of Saccharomyces cerevisiae Encodes a Mitochondrial NADH Kinase Required for Stability of Mitochondrial DNA. *Eukaryotic Cell*, *2*(4), 809–820. https://doi.org/10.1128/EC.2.4.809-820.2003

Sumner, E. R., Avery, A. M., Houghton, J. E., Robins, R. A., & Avery, S. V. (2003). *Cell cycle- and age-dependent activation of Sod1p drives the formation of stress resistant cell subpopulations within clonal yeast cultures*. *50*, 857–870. https://doi.org/10.1046/j.1365-2958.2003.03715.x

Temple, M. D., Perrone, G. G., & Dawes, I. W. (2005). Complex cellular responses to reactive oxygen species. *Trends in Cell Biology*, *15*(6), 319–326. https://doi.org/10.1016/j.tcb.2005.04.003

Thomasson, E. D., Scharman, E., Fechter-Leggett, E., Bixler, D., Ibrahim, S., Duncan, M. A., Hsu, J., Scott, M., Wilson, S., Haddy, L., Pizon, A., Burrer, S. L., Wolkin, A., & Lewis, L. (2017). Acute Health Effects After the Elk River Chemical Spill, West Virginia, January 2014: *Public Health Reports*. https://doi.org/10.1177/0033354917691257

Tkach, J. M., Yimit, A., Lee, A. Y., Riffle, M., Costanzo, M., Jaschob, D., Hendry, J. A., Ou, J., Moffat, J., Boone, C., Davis, T. N., Nislow, C., & Brown, G. W. (2012). Dissecting DNA damage response pathways by analyzing protein localization and abundance changes during DNA replication stress. *Nature Cell Biology*, *14*(9), 966–976. https://doi.org/10.1038/ncb2549

Toledano, M. B., Delaunay-Moisan, A., Outten, C. E., & Igbaria, A. (2013). Functions and cellular compartmentation of the thioredoxin and glutathione pathways in yeast. *Antioxidants & Redox Signaling*, *18*(13), 1699–1711. https://doi.org/10.1089/ars.2012.5033

Umebayashi, K., & Nakano, A. (2003). Ergosterol is required for targeting of tryptophan permease to the yeast plasma membrane. *Journal of Cell Biology*, *161*(6), 1117–1131. https://doi.org/10.1083/jcb.200303088

Vozza, A., Blanco, E., Palmieri, L., & Palmieri, F. (2004). Identification of the Mitochondrial GTP/GDP Transporter in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *279*(20), 20850–20857. https://doi.org/10.1074/jbc.M313610200

Weidhaas, J. L., Dietrich, A. M., DeYonker, N. J., Ryan Dupont, R., Foreman, W. T., Gallagher, D., Gallagher, J. E. G., Whelton, A. J., & Alexander, W. A. (2016). Enabling Science Support for Better Decision-Making when Responding to Chemical Spills. *Journal of Environment Quality*, *45*(5), 1490. https://doi.org/10.2134/jeq2016.03.0090

Weill, U., Cohen, N., Fadel, A., Ben-Dor, S., & Schuldiner, M. (2019). Protein Topology Prediction Algorithms Systematically Investigated in the Yeast Saccharomyces cerevisiae. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *41*(8), e1800252. https://doi.org/10.1002/bies.201800252

West Virginia Chemical Spill: Collective NTP Findings and Supporting Files. (n.d.). National Toxicology Program U.S. Department of Health and Human Services.

Whelton, A. J., McMillan, L. K., Connell, M., Kelley, K. M., Gill, J. P., White, K. D., Gupta, R., Dey, R., & Novy, C. (2014). Residential Tap Water Contamination Following the Freedom Industries Chemical Spill: Perceptions, Water Quality, and Health Impacts. *Environmental Science & Technology*, *49*(2), 813–823. https://doi.org/10.1021/es5040969

Yu, H., Braun, P., Yildirim, M. A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.-F., Dricot, A., Vazquez, A., Murray, R. R., Simon, C., Tardivo, L., Tam, S., Svrzikapa, N., … Vidal, M. (2008). High-quality binary protein interaction

map of the yeast interactome network. *Science (New York, N.Y.)*, *322*(5898), 104–110.

https://doi.org/10.1126/science.1158684
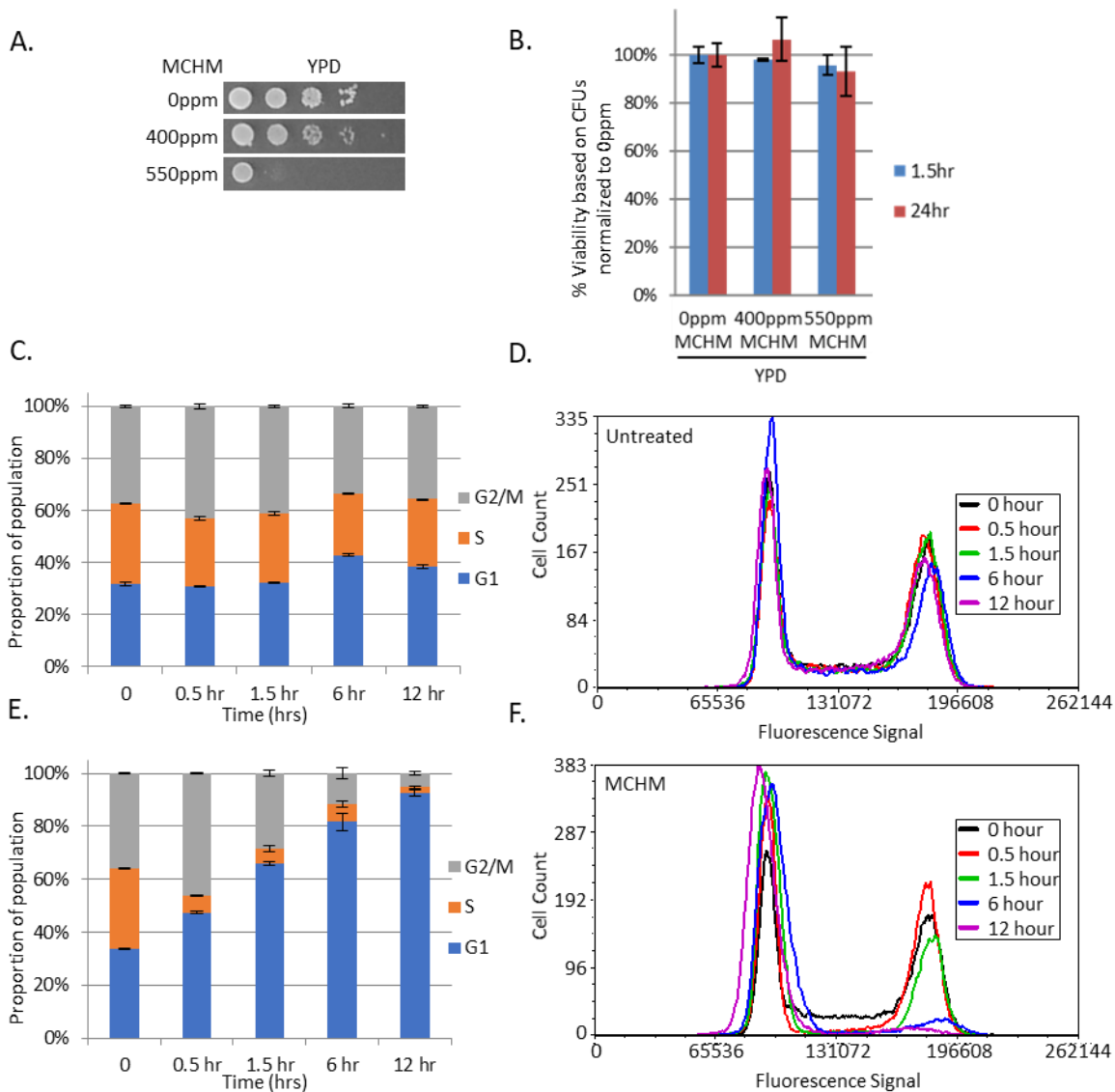
Figures:

Figure 1: *Growth and viability of yeast under MCHM treatment.* **A.** Growth of 10-fold dilutions of yeast on YPD media with and without MCHM treatment. BY4741 strain is derived from the S288c background. **B.** Acute viability assay of the MCHM sensitive strain BY4741 treated with MCHM for 1.5 hours and 24 hours. Colony-forming units for three biological replicates were averaged for each treatment then normalized to media without MCHM treatment. In YPD yeast were exposed to 400ppm and 550ppm crude MCHM. **C.** Proportion of log-phase BY4741 in G1, S, or G2/M phases of the cell cycle based on flow cytometric analysis grown in YPD. **D.** Histograms of DNA content in BY4741 population grown in YPD shown in part C, as based on fluorescence intensity of sytox green. **E.** The proportion of log-phase BY4741 in G1, S, or G2/M phases of the cell cycle based on flow cytometric analysis grown in YPD with 550ppm MCHM. **F.** Histograms of DNA content of BY4741 population grown in YPD treated with MCHM shown in part E, as based on fluorescence intensity of sytox green. All samples were analyzed for 0, 0.5,

1.5, 6, and 12-hour time points to monitor cell cycle changes over time. (All parts were performed by Michael C. Ayers)
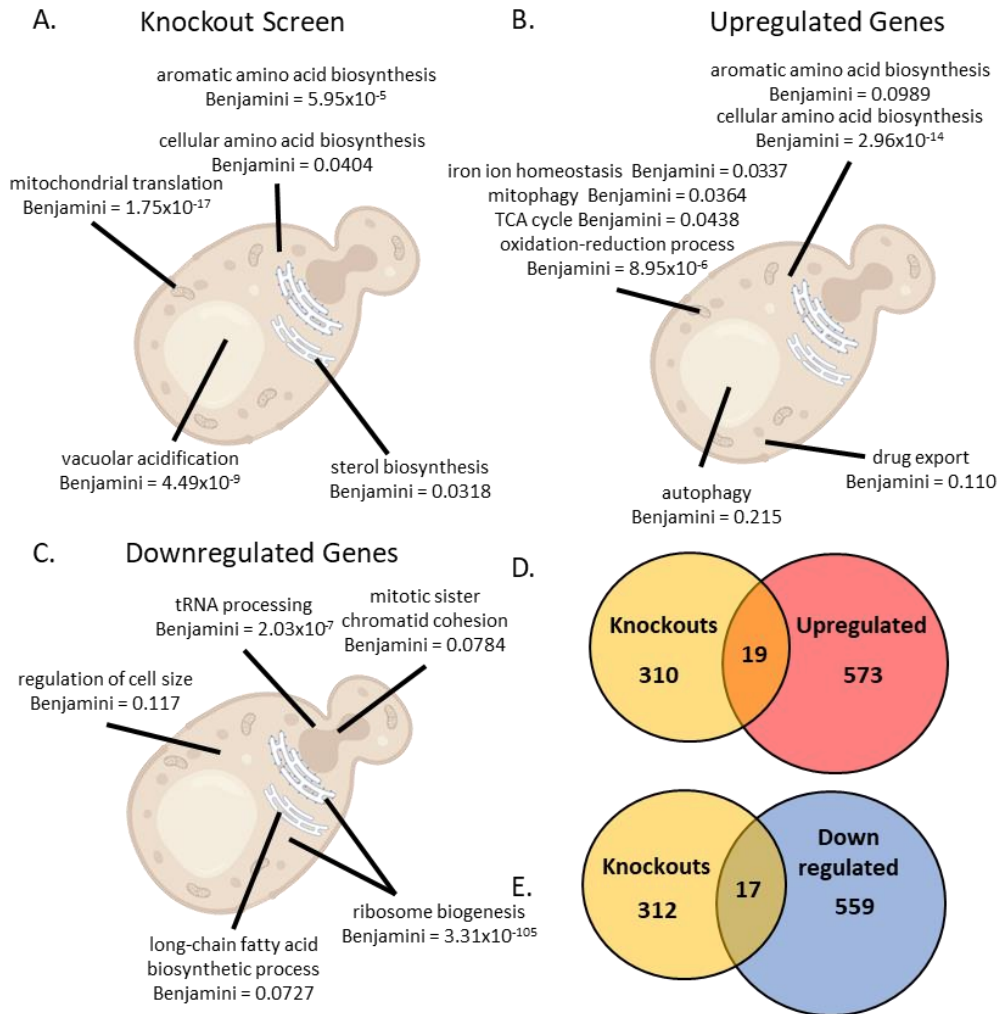
Figure 2: *Genetic screen of knockout collection and RNAseq of BY4741 gene lists.* **A.** Select top biological process GO Terms for the list of 329 genes found to be hits in the genetic screen of BY4742 knockout collection. Benjamini-Hochberg adjusted p-values reported by DAVID software are shown for each term. **B.** Select top biological process GO Terms for the 592 significantly upregulated genes from BY4741 treated with MCHM versus untreated YPD. FDR is shown for each term. **C.** Select top biological process GO Terms for the 572 significantly downregulated genes from BY4741 treated with MCHM versus untreated YPD. FDR is shown for each term. **D.** Venn diagram of the overlap between the 329 genes from the genetic screen and the upregulated genes. **E.** Venn diagram of the overlap between the 329 genes from the genetic screen and the downregulated genes. (Michael Ayers performed the knockout screen and analysis in part A and all downstream analysis of transcriptome and knockout GO term and intersection analyses in parts B through E. Jennifer Gallagher performed RNA experiment and Amaury Pupo performed differential analysis for transcriptome in Parts B through E. from previous study Gallagher *et al.,* 2020.)
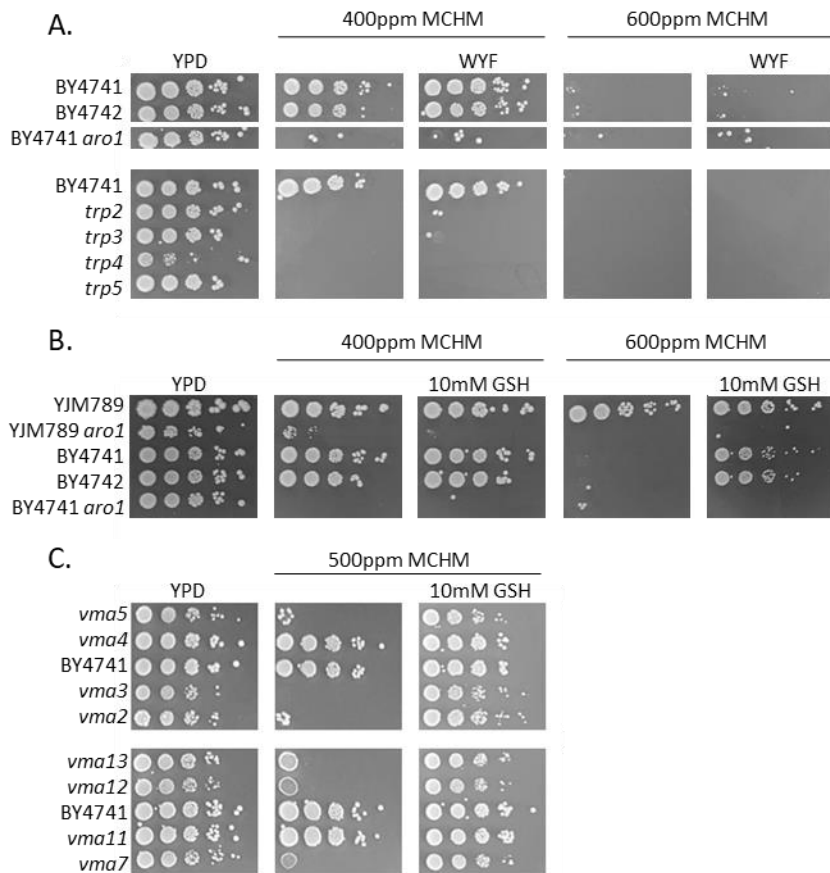
Figure 3: *Sensitive mutants from gene lists cotreated with possible rescuing chemicals* **A.** 10-fold serial dilution growth assay of knockouts of aromatic amino acid biosynthesis genes found in the genetic screen. Strains were cotreated with MCHM and aromatic amino acids Tryptophan (W), Tyrosine (Y), and Phenylalanine (F) in rich media, testing the extra supplementation of these amino acids beyond normal levels in YPD. **B.** 10-fold serial dilution growth assay of wildtype strains and knockout of aromatic amino acid biosynthetic gene *ARO1* with and without cotreatment of 10mM antioxidant glutathione (GSH) and MCHM. **C.** 10-fold serial dilution growth assay of knockouts in the vacuolar acidification pathway as found in the genetic screen. Strains were cotreated with and without 10mM antioxidant GSH and MCHM. (All experiments and analysis were performed by Michael C. Ayers)
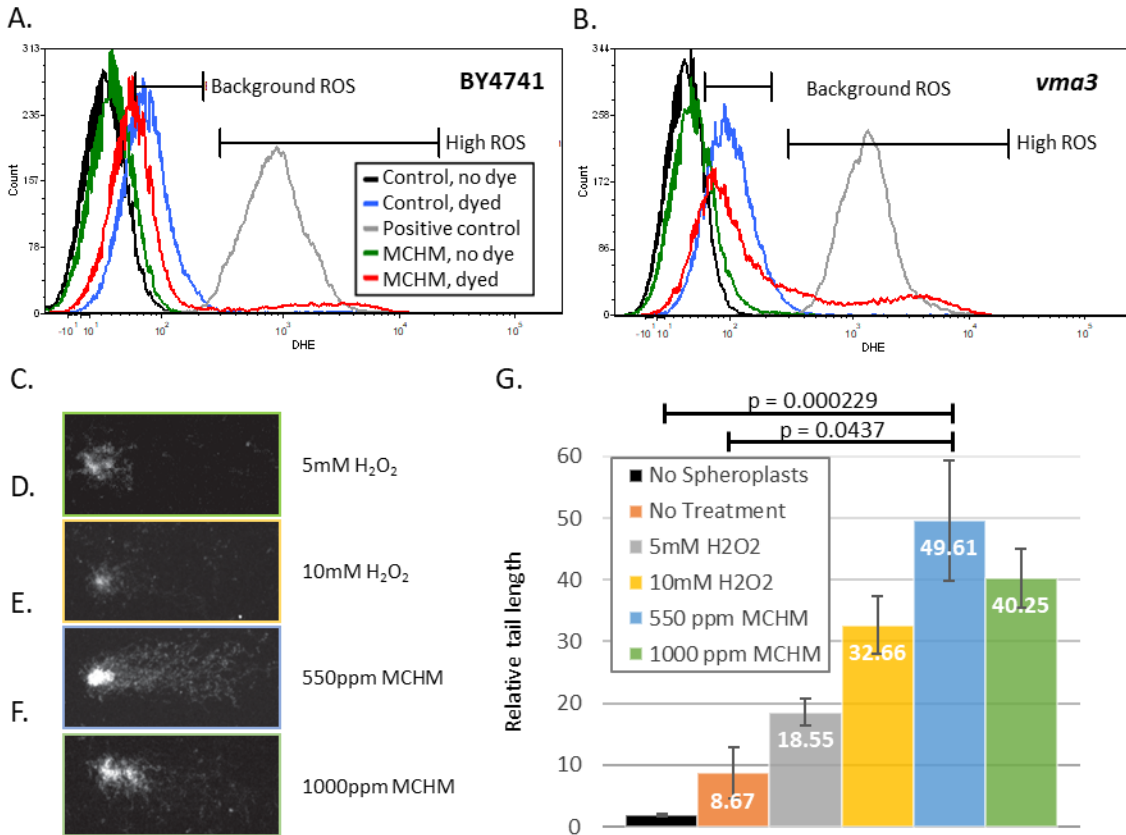
Figure 4: *Flow cytometry and comet assay of yeast strains reveals the production of Reactive Oxygen Species and DNA damage.* Levels of ROS in strains of yeast exposed to MCHM, based on fluorescence of the ROS-reactive dye DHE. Yeast were incubated for 12 hours with or without MCHM then stained with DHE for 20 minutes before cells were sorted using flow cytometry. Hydrogen peroxide treatment for 1.5 hours was used as a positive control to generate ROS in control cells (grey line). Background ROS of untreated yeast in blue measures endogenous ROS compared to unstained yeast with no DHE (black). MCHM treated yeast are in green while MCHM treated yeast stained with DHE are in red. Markers were inserted on the histogram to indicate a range of fluorescence corresponding to background or high ROS based on peaks for untreated control cells. Histograms are for the strains **A.** BY4741, and **B.** *vma3* knockout mutant. **C-F.** Examples of cells visualized by microscopy for the comet assay. Cells underwent treatment prior to spheroplasting, electrophoresis on a slide, and visualization via microscopy and ethidium bromide staining. Lengths of tails were measured with TriTek CometScore 2.0.0.38 software. The samples shown are examples of the following treatments **C.** 5mM H_2O_2, **D.** 10mM H_2O_2, **E.** 550ppm MCHM, and **F.** 1000ppm MCHM. **G.** Graph of tail lengths of cells determined via comet assay. (Michael C. Ayers performed the flow cytometry and follow-up analyses in parts A and B. Zachary Sherman performed the comet assay and analysis in parts C-G.)

71
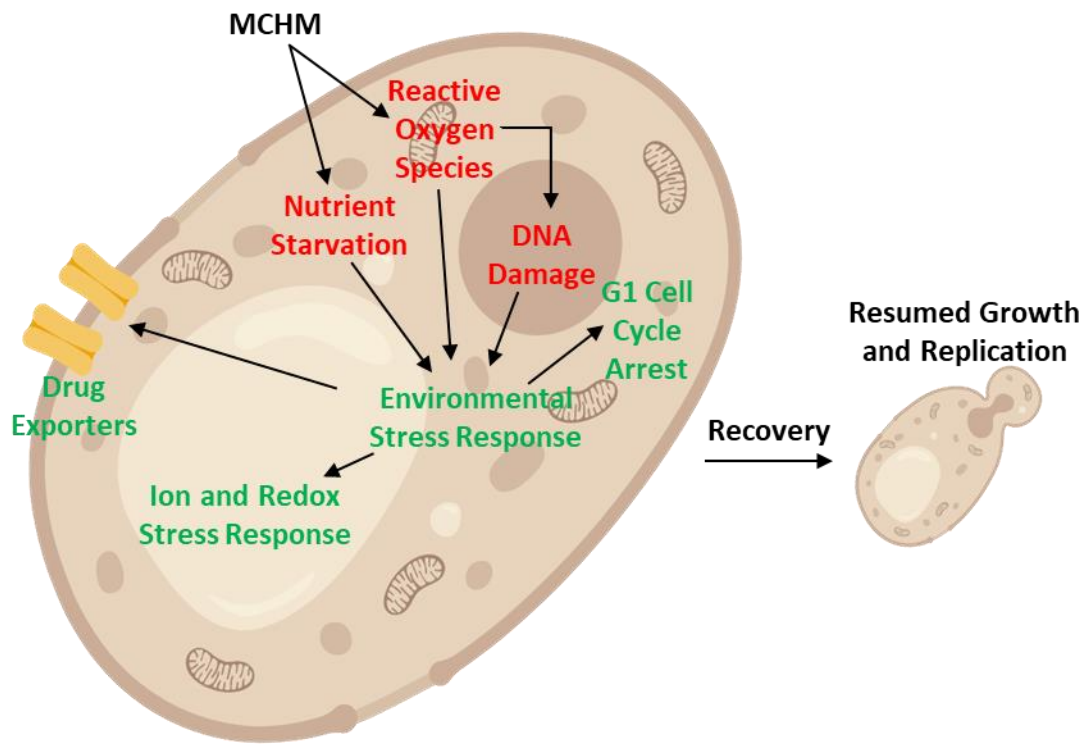
Figure 5: *Model of response and tolerance of yeast cells to MCHM treatment.* Diagram of yeast cell showing MCHM entering cell and causing ROS, DNA damage, and possible nutrient starvation. These signals are shown triggering environmental stress response leading to cellular responses that lead to recovery and resumption of replication. (Model was made by Michael C. Ayers)
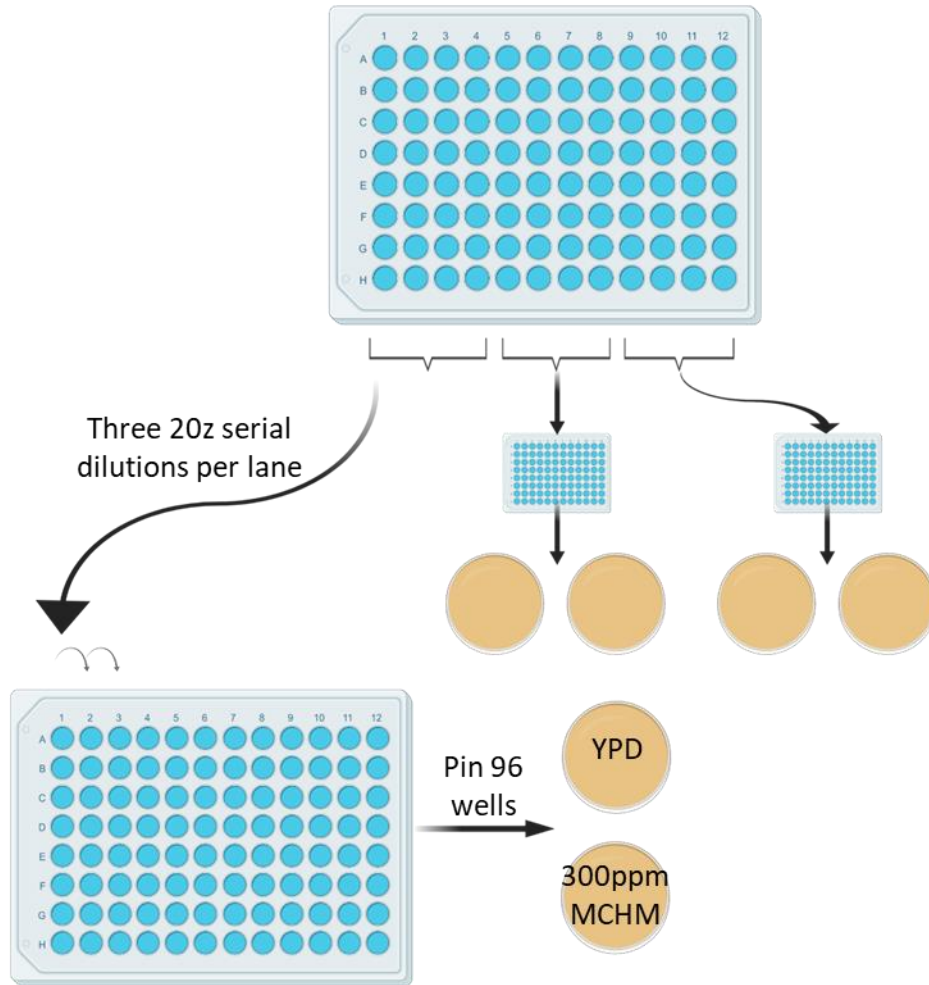
Supplemental Figure 1: *Diagram of the genetic screen of haploid BY4742 knockout collection.* The 96-well plates of the knockout collection were each inoculated into growth chambers in triplicate by pinning and grown for two days to saturation. These chambers were then serially diluted 20-fold three times into new 96 well plates, before being pinned onto YPD agar plates with or without 300ppm MCHM. Pictures were taken after 3 days of growth and scored for growth defects consisting of decreased growth of at least a full spot on MCHM plates compared to control strain (wt BY4742) and control plate growth. (Genetic screen was designed by Michael C. Ayers)
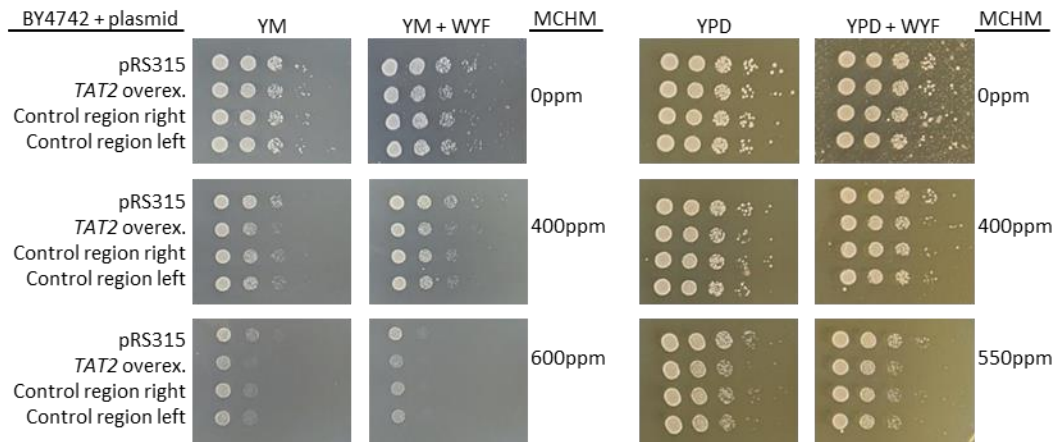
Supplemental Figure 2: *TAT2 overexpression does not rescue the MCHM growth defect.* 10-fold serial dilution growth assay of BY4742 with indicated plasmids on minimal and rich media with indicated treatment with MCHM and supplemental aromatic amino acids tryptophan (W), tyrosine (Y), and phenylalanine (F). pRS315 is a control CEN plasmid expressing the *LEU2* marker to confer growth on leu- media to the wild-type BY4742 without overexpression of any genomic regions. *TAT2* overex. contains a tiling array multicopy plasmid with the genomic region of chromosome XV containing *TAT2* for its overexpression (from the *IFM1* ORF to the left of *TAT2* to the YOL019W-A ORF to the right, six complete ORFs and one partial). The control region right plasmid is the same multicopy tiling array plasmid but with the genomic region flanking the *TAT2* ORF on the right side, from within the *TAT2* ORF to the YOL015W ORF (six complete ORFs and two partials). The control region left plasmid is the same multicopy tiling array plasmid but with the genomic region flanking the *TAT2* ORF on the left side, from the *IFM1* ORF to the *DIS3* ORF directly to the left of *TAT2* (two complete ORFs and two partials). (All experiments were performed by Michael C. Ayers)
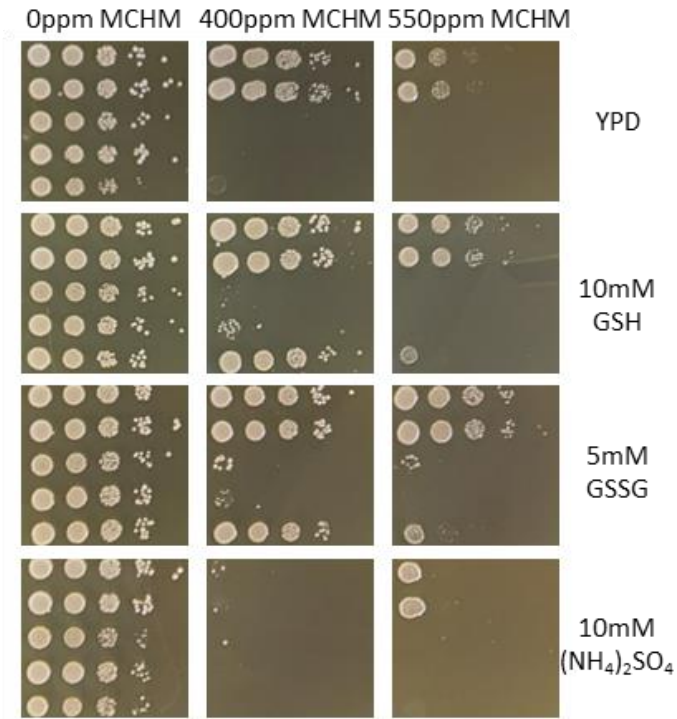
Supplemental Figure 3: *glr1 mutant rescued by oxidized glutathione, but not general nitrogen supplement.* 10-fold serial dilution growth assay of BY4742 and knockout mutants for *glr1, aro1, trp5,* and *vma3*. The mutants were grown on rich media containing 0ppm, 400ppm or 550ppm MCHM, as well as cotreatments with 10mM reduced glutathione (GSH), 5mM oxidized glutathione (GSSG), or 10mM ammonium sulfate ($(NH_4)_2SO_4$). (Michael C. Ayers performed this experiment)
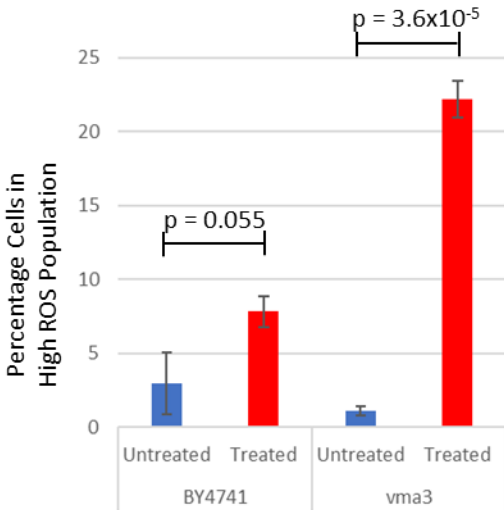
Supplemental Figure 4: *Increased percentage of High ROS cells with MCHM treatment.*
Percentages of High ROS subpopulation of cells from the BY4741 wiildtype and *vma3* mutant
strains, based on fluorescence of the ROS-reactive dye DHE are shown. Yeast were incubated
for 12 hours with or without MCHM then stained with DHE for 20 minutes before cells were
sorted using flow cytometry. Fluorescence level (350) for determining High ROS subpopulation
was based on approximately 95% (94.69%-96.86%) of H_2O_2 positive control cells in all three
BY4741 replicates being above this fluorescence level. Bar graph shows the averages of three
biological replicates for each strain and treatment with standard error bars. One-tailed t-test
was used to determine p-values. (The flow cytometry and this analysis were performed by
Michael C. Ayers)

All Supplemental Tables and Files are available at https://github.com/wvumayers/MCHM-Genetic-Screen.

Table S1: *Full list of knockout strains found as hits in the genetic screen of the BY4742 collection.* Includes ORF and gene names of all 329 hits from the screen. 4983 strains from the haploid knockout collection were assayed for growth in the presence of 300ppm MCHM on solid YPD media. While approximately 900 strains showed some noticeable growth differences on at least one replicate, the 329 strains showed clear growth differences on all 3 biological replicates, accounting for a stringent threshold for inclusion as a "hit" for the knocked out gene being required for MCHM resistance.

Table S2: *Full results of significantly upregulated genes from RNAseq data in MCHM/YPD using BY4741 wildtype strain.* Upregulated transcripts from the BY4741 RNAseq are listed in full. Included are the ORF names, as well as log2fold change and adjusted p-values supplied by the differential analysis. All genes are at least greater than a two-fold increase in expression in MCHM than YPD. Gene names, locations, and putative functions are concatenated to the rightmost columns of the table for further information on included ORFs.

Table S3: *Full results of significantly downregulated genes from RNAseq data in MCHM/YPD using BY4741 wildtype strain.* Downregulated transcripts from the BY4741 RNAseq are listed in full. Included are the ORF names, as well as log2fold change and adjusted p-values supplied by the differential analysis. All genes are at least greater than 2fold decreased in expression in MCHM than YPD. Gene names, locations, and putative functions are concatenated to the rightmost columns of the table for further information on included ORFs.

Supplemental File 1: *Full GO terms for the genetic screen of BY4742 knockout collection.* Excel file containing 3 tabs. Tab 1 gives DAVID GO term results for the biological process. Tab 2 gives DAVID GO term results for molecular function. Tab 3 gives DAVID GO term results for the cellular component.

Supplemental File 2: *Full GO terms for significantly upregulated genes of BY4741 yeast treated with MCHM.* Excel file containing 3 tabs. Tab 1 gives DAVID GO term results for the biological process. Tab 2 gives DAVID GO term results for molecular function. Tab 3 gives DAVID GO term results for the cellular component.

Supplemental File 3: *Full GO terms for significantly downregulated genes of BY4741 yeast treated with MCHM.* Excel file containing 3 tabs. Tab 1 gives DAVID GO term results for the biological process. Tab 2 gives DAVID GO term results for molecular function. Tab 3 gives DAVID GO term results for cellular component.

CHAPTER 4: MCHM ACTS AS A HYDROTROPE, ALTERING THE BALANCE OF METALS IN YEAST

Amaury Pupo**, Michael C. Ayers**, Zachary N. Sherman, Rachel J. Vance, Jonathan R.

Cumming, and Jennifer E.G. Gallagher[*]

** Co-first authors

From the Department of Biology, West Virginia University, 53 Campus Drive, Life Sciences

Building Rm 5105, Morgantown, WV 26506 USA

Running title: *MCHM and metals in yeast*

Abstract

While drugs and other industrial chemicals are routinely studied to assess risks, many widely-used chemicals have not been thoroughly evaluated. One such chemical, 4-methylcyclohexane methanol (MCHM), is an industrial coal-cleaning chemical that contaminated the drinking-water supply in Charleston, WV, USA in 2014. While a wide range of ailments was reported following the spill, little is known about the molecular effects of MCHM exposure. We used the yeast model to explore the impacts of MCHM on cellular function. Exposure to MCHM dramatically altered the yeast transcriptome and the balance of metals in yeast. Underlying genetic variation in the response to MCHM and transcriptomics and mutant analysis uncovered the role of the metal transporters, Arn2 and Yke4, to MCHM response. Expression of Arn2, involved in iron uptake, was lower in MCHM-tolerant yeast and loss of Arn2 further increased MCHM tolerance.

Genetic variation within Yke4, an ER zinc transporter, also mediated response to MCHM and loss of Yke4 decreased MCHM tolerance. The addition of zinc to MCHM-sensitive yeast rescued growth inhibition. *In vitro* assays demonstrated that MCHM acted as a hydrotrope and prevented protein-interactions, while zinc induced the aggregation of proteins. We hypothesized that MCHM altered the structures of extracellular domains of proteins, and the addition of zinc stabilized the structure to maintain metal homeostasis in yeast exposed to MCHM.

Background

The potential for significant human exposure to toxic substances is increasing as thousands of chemicals in routine use have had little safety testing [1–3]. 4-Methylcyclohexanemethanol (MCHM) is an alicyclic primary alcohol used as a cleaning agent in the coal industry. Although health and safety information for this compound is limited, its widespread use in the coal-producing regions of the United States represents a potential hazard to humans and ecosystems. In January 2014, a large quantity of MCHM was spilled into the Elk River in West Virginia, USA and contaminated the drinking water supply of 300,000 people, exposing them to unknown health risks [4]. People exposed to MCHM through the contaminated drinking water reported a variety of significant ill effects [5].

MCHM is not easily degraded biologically because of its low reactivity [6]. In contrast to other well-studied hydrocarbons, such as cyclohexane and benzene, the effects of MCHM on metabolism are under studied [7]. Yeast lines exposed to MCHM exhibited increased expression

of proteins associated with the membrane, cell wall, and cell structure functions, while MCHM metabolites mainly induced proteins related to antioxidant activity and oxidoreductase activity [3]. With human A549 cells, MCHM mainly induced DNA damage-related biomarkers, which indicates that MCHM is related to genotoxicity due to its DNA damage effect on human cells [3].

Yeast provide an ideal model system to understand the interplay between metabolic pathways involved in the transport, toxicity, and detoxification of MCHM. Further, the use of yeast strains with mutations in various metabolic pathways allows direct evaluation of targeted pathways on the fate and toxicity of MCHM in cells. "Petite yeast", lines with mutations that disrupt the electron transport chain that produces ATP in the mitochondria, grow more slowly and have smaller cell size than "grande yeast" (wild type). Because these yeast mutants can generate sufficient energy through glycolysis, however, these are not lethal mutations, but provide a slow-growing line to evaluate the metabolic rate and stress response. In addition to their roles in energy transformations, mitochondria are central for the synthesis of amino acids, nucleotides, and heme and Fe-S cluster proteins. Thus, such yeast lines are ideal models to assess the role of mitochondrial function in response to stress.

Petite yeast have different tolerances to chemicals, which may be related to the production of reactive oxygen species (ROS) and mitochondrial function. For example, petite yeast are more tolerant to 4-nitroquinoline 1-oxide (4NQO) than grande yeast when grown on non-

fermentative carbon sources that favor respiration [8, 9]. 4NQO is metabolized to the active

form only in cells with functional mitochondria, and petite yeast, without mitochondria and

favoring fermentation, are more resistant than wild type yeast. However, petite yeast have

higher levels of endogenous ROS and are sensitive to compounds that also generate ROS [10],

such as MCHM. Petite yeast are additionally more sensitive to H_2O_2 [10, 11] but more resistant

to copper [12]. Sod1 is the main dismutase that neutralizes ROS in the cytoplasm, and in the

mitochondria, Sod2 also neutralizes ROS. Thus, the petite yeast-grande yeast pair represent an

ideal system to evaluate the role of ROS systems, metal homeostasis, and MCHM toxicity.

The hydrophobicity of MCHM alters membrane dynamics, which changes how cells can respond

to the environment, including the import and subcellular localization of metals. Metal

homeostasis is critical in that metals, such as iron (Fe) and zinc (Zn), play important roles in

metabolism as co-factors for enzymes and other proteins, yet, if in excess, induce broad lesions

to cell biology through the generation of ROS and by binding to a variety of biomolecules [13].

The coordinated activity of metal uptake and sequestration transporters function to maintain

metals at optimal levels (Figure 1). For example, there are two Zn transporters located on the

cell membrane of yeast. Zrt1 is the high-affinity transporter that transports Zn when

extracellular levels are low [14], and Zrt2 is the low-affinity transporter [15]. Zrt3 transports Zn

from storage in the vacuole to the cytoplasm when needed [16], while Zrc1, the Zn/H^+

antiporter, and its paralog Cot1 [16, 17] transport Zn into the vacuole from the cytoplasm. Izh1

and its paralog Izh4 are both involved in Zn homeostasis by altering membrane sterol content

or by directly altering cellular Zn levels [18, 19].

81

In the current study, we evaluated the impacts of MCHM on petite and grande yeast strains, focusing on metal ion homeostasis and divergent respiratory pathways in these lines as potential mechanisms of MCHM sensitivity. We integrate transcriptomics, ionomics, and QTL to identify Fe and Zn homeostasis as central to MCHM toxicity, and suggest that loss of metal homeostasis underlies ROS damage and MCHM toxicity. While the MCHM has low solubility in aqueous solutions, we propose that MCHM acts as a hydrotrope, altering membrane dynamics and changing how cells responded to nutrients, including import and subcellular localization of metals via transmembrane ion transporters.

Experimental Procedures

Yeast strains and media

S96 petite yeast were previously generated from S96 (MATa *lys5*) with a six hour incubation with ethidium bromide. The petite phenotype was validated by failure to grow on glycerol and loss of *COX2* from the mitochondrial genome [9]. Yeast knockout strains were previously generated in the BY4741 background [20]. The entire coding regions of *YKE4* were knocked out with hygromycin resistance marker in S96 and YJM789 [21]. S96-derived strains were grown in minimal media supplemented with lysine, while BY4741 strains grown in minimal media were supplemented with histidine, methionine, uracil, and leucine. S96 and BY4741 are considered in the S288c genetic background while YJM789 is a clinical isolate [22]. Crude MCHM provided directly from Eastman Chemical.

QTL analysis

Isolates of the recombinant haploid collection between S96 and YJM789 [23] were utilized to perform a QTL analysis for MCHM tolerance. A genetic map was constructed and combined with phenotypes collected by growth curves of the segregants in YPD containing MCHM using a TECAN M200 plate reader. Briefly, cultures of segregants were grown overnight then diluted to 0.1 OD_{600} starting concentrations in 200 µl of YPD and either 0 ppm (0 mM) control or 400 ppm (2.8 mM) MCHM. Each segregant was grown in biological triplicate for both control and MCHM treatments in 96-well plates. Both parent strains were grown on every plate to normalize plate-to-plate variation. Plates were grown with constant shaking and OD data was collected every 10 minutes for 24 hours. Differences in control and MCHM saturated concentrations from hours 14-19 of the growth curves were averaged into a single data point to serve as the phenotype for each segregant.

The S96 × YJM789 segregant collection used for genetic analysis contains 126 segregants genotyped at 55,958 SNPs identified by physical location. To perform the QTL analysis as previously described [24, 25], the genetic map was estimated for use in place of the existing physical map. Computational efficiency was also improved by collapsing the 55,958 markers into only 5,076 markers. The R/QTL package based this reduction of markers on by finding regions of markers that did not recombine and segregate amongst the offspring, which it collapsed into individual randomly selected markers within the linked region. The genetic map was created using R version 3.4.3 and the QTL package version 1.41.6. The scripts contained in

83

Karl W. Broman's genetic map construction guide were modified for the haploid yeast genome and used to output the map. The known order and location of markers on the physical map were used to validate that the genetic map output was ordered correctly. The physical map was also used to identify linkage names as established yeast chromosomes I to XVI. The QTL analysis was performed using the maximum likelihood EM algorithm method to calculate LOD scores at all loci. Significance thresholds of alpha = 0.05 were applied using 1000 permutations to determine the significance of LOD scores.

Serial dilution assays

Yeast were serially diluted onto solid media as previously described [24]. MCHM and $ZnSO_4$ were added to media, autoclaved and cooled to 65°C, which was then poured. Specific experimental conditions varying the concentrations of MCHM, zinc, and yeast strains are outlined in the results.

Transcriptomics

The RNA-seq of S96 and S96 petite yeast was carried out on hot phenol extracted RNA [9]. The raw reads from sequencing are from 16 samples GSM2915204 through GSM2915219, including normal and petite cells. "GSE108873_mchm_fC1_count_table_clean.txt.gz" is the count data generated via Rsubread. "GSE108873_mchm_fC1_DESeq_c2.tsv.gz" is the differential expression data generated via DESeq2. The accession number is GSE108873

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108873). GO term analysis was undertaken with clusterProfiler [26], an R package that implements methods to analyze and visualize functional profiles (GO and KEGG) of gene and gene clusters. For this, the ORF names from genes up- or downregulated in each condition were translated to the corresponding Entrez id using the function bitr and the package org.Sc.sgd.db. The resulting gene clusters were processed with the compareCluster function, in mode enrichGO, using org.Sc.sgd.db as database, with Biological Process ontology, cutoffs of p-value = 0.01 and q value = 0.05, adjusted by "BH" [27], to generate the corresponding GO profiles, which were then simplified with the function simplify. The simplified profiles were represented as dotplots, providing up to 15 most relevant categories.

Elemental analysis

Yeast were grown to mid-log phase in YPD at which time MCHM was added to a final concentration of 550 ppm (3.9 mM). 1.2×10^8 yeast cells were harvested following 0, 10, 30, and 90 min exposure to MCHM. Samples were split and washed, one set twice with water and the other washed once with 10 mM EDTA to remove metals adsorbed to the extracellular matrix. Four biological replicates for each sample were frozen in liquid nitrogen and stored at –80˚C. Cell pellets were digested in 300 µl of 30% H_2O_2 and followed by 700 µl of concentrated nitric acid. Samples were heated to 85˚C until clear. Sample volume was brought up to 10 ml with HPLC grade water. Digested samples were analyzed in technical triplicate by inductively coupled plasma emission spectrometry (Agilent 5110 ICP, Agilent, Santa Clara, CA, USA) at the following

wavelengths for each metal: Ca 393.366, Fe 238.204, Mg 279.553, Na 589.592, P 213.618, and

Zn 206.200. Concentrations of elements in digests (mg l^{-1}) were normalized to protein levels as

determined by Bradford assay.

Hydrotrope assay

Hydrotrope assays were carried out as previously described [28], with the following

modifications. Eggs were purchased and used within one day. Egg whites were separated and

diluted 1:6 in 50 mM Tris-HCl pH 7.4. In glass tubes, 3 ml of diluted egg whites were mixed with

different concentrations of ATP, MCHM, and zinc sulfate. Samples in the glass tubes were read

at 450 nm after 45-60 seconds of incubation at 60˚C water bath. All treatments were done in 3-

4 replicates and averaged together with the standard deviation shown. Statistical differences

were determined using a student t-test.

Spheroplasting

Spheroplasts of the BY4741 yeast strain were prepared as described [29] with the

following modifications. HEPES was used as the buffer and 100 mg 20T zymolase was added per

OD unit (OD$_{600}$ multiplied by the volume of culture); spheroplasts were incubated for 1 hour at

30˚C. In a 96-well plate, 5 replicates of each treatment were recorded: empty well, spheroplast

media only, no treatment, 0.1% SDS, 0.5% sterile distilled water, 10 mM sorbitol, 1 mM ATP, 10

mM NaXS, 10 µM $ZnSO_4$, 550-1000 ppm (3.9 to 7.1 mM) MCHM. Using a spectrophotometer, the absorbance at 600 nm was recorded every 15 hours at room temperature.

Microscopy

Yeast with proteins tagged at the N-terminus with mCherry under the *TEF2* promoter [30] were grown to log phase and then split into 8 different cultures (4 for treated and 4 for untreated). Once in log-phase, samples were treated with 550 ppm (3.6 mM) MCHM for 30 minutes followed by a 20-minute incubation time of 25 µM calcofluor white (Biotium catalog number 29067) on a shaker in a dark room. Then 30 µl of each sample onto a microscope slide pretreated with 40 µL of 250 µg/ ml concanavalin A via pipette onto microscope slides and let sit under a hood for 30 min to dry. The coverslip was placed on top and nail polish was applied around the edges to hold coverslip in place. Cells were imaged on Nikon A1R confocal microscope using a FITC Laser and DAPI laser. Quantitative analysis of pixel intensity to measure the change in expression of the proteins of interest after exposure to MCHM was done with ImageJ on 17-20 cells for each condition. The signal was normalized to untreated yeast and statistical differences were determined using a student t-test.

Results

Petite yeast have different responses to chemicals because the metabolism of the yeast is shifted away from respiration. Indeed, compared to wild-type (grande) yeast, petite yeast were

more sensitive to MCHM (Figure 2). Growth inhibition of petite yeast was affected at 125 ppm

MCHM while the growth of wild-type yeast was only affected at 500 ppm MCHM. Yeast have

higher tolerances to MCHM when grown in minimal media (YM); however, petite yeast grew

less on YM in general (Figure 2).

To assess the transcriptional response of petite yeast to MCHM, yeast were grown in both YPD

and YM and then treated for 90 minutes with 550 ppm MCHM. 949 genes were differentially

regulated across strains and conditions (Supplemental Table 1). Gene expression levels

between petite and isogenic grande yeast were similar when grown in YPD media (only six

upregulated genes: cell wall components and iron transporters, and one downregulated gene:

putative mitochondrial protein, Figure 3A), but they were clearly different when grown in YM

(131 up and 117 downregulated genes, Supplemental Figure 1A, Supplemental Table 1). In YM,

petite yeast exhibited downregulated cell wall component genes, and there were also

significant changes on small molecule metabolism-related genes (Supp. Figure 1A, 2 and 3,

Supplemental. Table 1). MCHM treatment elicited the upregulation of genes involved in small

molecule and sulfur compound biosynthesis, among others, in both petite and grande yeast,

while the regulation of other genes differed between strains and depended on the media. For

example, genes related to nucleotide and nucleoside metabolism were upregulated in the

petite strain treated with MCHM only when grown in YM (Figure 3, Supplemental Figure 1 and

2, Supplemental Table 1). Genes downregulated due to MCHM treatment also depended on the

media (Figure. 3, Supplemental Figure 1 and 3, Supplemental Table 1). Among the genes with

variable expression due to MCHM treatment and/or the use of petite vs. grande yeast were

88

several involved in zinc homeostasis (*COT1*, *IZH1*, *IZH2* and *IZH4*, *ZRT1* and *ZRT2*) and iron

homeostasis (*ARN1 ARN2*, *ENB1*, *FIT1*, *FIT2, FIT3*, *FTR1*, *GGC1* and *SIT1*) (Figure 3, Supplemental

Figure 1 and Supplemental Table 2 and 3).

The increased expression of iron transporters was further explored given that mitochondria,

the presence of which differs between the two strains, are the site of iron-sulfur cluster protein

biogenesis. Strain ionomic profiles were evaluated for yeast strains grown in YPD. YPD was

chosen to minimize the differences in growth between grande and petite yeast. Because there

were differences in expression of cell wall genes such as *CWP1*, yeast were also washed in EDTA

to determine if increased iron or other metals were associated with the cell wall (water wash)

or internalized (EDTA wash). There was no difference in metal levels from yeast washed with

water alone or EDTA, indicating that the levels of reported ions were absorbed into the cells

and were not associated with the cell walls (Supplemental Figure 4 and Supplemental Table 4).

Iron levels were three times higher in petite yeast than grande yeast, while zinc was 60% lower

in the petite strain (Figure 4A). Other elements were typically higher in grande compared to

petite yeast except for calcium (Figure 4A, Supplemental Table 4), as yeast mitochondria do not

store calcium (reviewed [31]). Levels of sodium, phosphorus, and magnesium were lower in

petite yeast (Figure 4A). Copper was below the limit of detection in this analysis.

To determine if the levels of metals in yeast were changed when exposed to MCHM, yeast were

grown in YPD and then MCHM was added to a final concentration of 550 ppm. The levels of iron

89

did not change for either grande or petite yeast over 90 minutes, although the strain-specific

differences noted above, were still notable (Figure 4B). mRNAs encoding siderophore

transporters such as Arn1 and Arn2 were expressed at higher levels, as were the Fit

mannoproteins, which first bind siderophores in petite yeast in YPD and MCHM compared to

wild-type yeast (Supplemental Table 3). In contrast, the levels of zinc increased two-fold in

grande yeast but did not change in petite yeast over 90 min (Figure 4C). Calcium and sodium

increased with MCHM treatment in both strains, with sodium increasing at a slower rate in the

petite yeast (Figure 4D and 4E). Potassium, phosphorus, and magnesium also increased in

MCHM treatment for grande, but not in petite, yeast (Figure 4F, 4G, and 4H). The levels of

these ions are comparable to other studies [32]. After 90 minutes of MCHM exposure, five out

of the seven ions measured were significantly higher in the grande yeast (Table 1).

Table 1. Ionomic profiles (μg/mg protein) of wildtype (S96) and petite yeast strains grown for 90 min in YPD.

Element	Wildtype	Petite
P*	104.4 ±7.5	65.8 ± 2.2
K*	19.0 ± 0.45	12.3 ± 0.32
Ca	0.554 ± 0.048	0.582 ± 0.027
Mg*	11.35 ± 0.79	7.87 ± 0.25
Na	1.64 ± 0.13	1.77 ± 0.11
Fe*	0.123 ± 0.014	0.313 ± 0.014
Zn*	2.89 ± 0.24	1.17 ± 0.04
Cu	ld	ld

*Denotes significant differences between yeast strains (P<0.05).

ld = level of detection.

There is a significant variation of growth among genetically-diverse yeast strains in response to MCHM. In particular, YJM789, a yeast isolated from a human lung [22], was more sensitive than S96 at 500 ppm MCHM (Figure 5A). Using a segregant collection of YJM789 and S96 that has been used to map genes contributing to differences in phenotypes between strains [8, 23, 24], quantitative trait loci (QTL) analysis was carried out. The growth rate in MCHM of segregant yeast strains was used to assess the association of various parts of the genome with increased growth in MCHM (Figure 5B). Several peaks were noted, but only one broad peak on chromosome nine passed the 95% confidence threshold. Within that peak, we identified *YKE4*, a polymorphic gene that is a ZIP family zinc transporter [33] that plays a role in zinc homeostasis by transporting zinc between the cytoplasm and the secretory pathway [33] and is localized to the ER [30]. Yke4^{YJM789} contained two SNPs that change the protein's amino acid sequence (H5Q and F86L) compared to Yke4^{S288c} (Figure 5C). The H5Q was located in the cytoplasmic signal peptide at the N-terminus, while the F86L is at the C-terminal end of the first transmembrane domain using TMHMM [34]. To further characterize the role of Yke4, *YKE4* was deleted from S96 and YJM789. Deletion of *YKE4* did not alter growth in the presence of MCHM in these strains (Figure 5D). The zinc levels were measured in these strains. The percentage of zinc was normalized to wild type S96. Both YJM789 and the isogenic *yke4* knockout strains had twice as much zinc compared to S96 (Figure 5E). Yke4 is an ER-localized zinc transporter and plays a role in intracellular trafficking of zinc and did not appear to regulate total zinc levels. There were other genomic peaks in the QTL linked to genetic variation of MCHM response and likely contribute to differences seen between these strains.

To assess the contribution of other proteins involved in metal transport, we utilized the yeast

knockout collection to determine the impact of deleting MCHM differentially-regulated genes

on growth. This collection is in an S96-related strain background BY4741. In contrast to S96 and

YJM789 yeast, the *yke4* mutant in this background was sensitive to MCHM (Figure 6A). There

were no significant changes in expression of *YKE4* induced by MCHM (Supplemental Table 1).

However, *ARN2* expression was higher in petite yeast and, from ICP-MS analysis, the

endogenous levels of iron were also higher. The BY4741 *arn2* knockout was more tolerant to

MCHM (Figure 6A). The growth on MCHM of *izh1* and *izh2* knockouts, genes involved in zinc

transport that also were differently regulated, was not altered. Iron levels did not change with

the addition of MCHM (Figure 4B). However, zinc levels increased in the wild-type grande yeast

but not in the petite yeast with MCHM exposure (Figure 4C). We tested whether additional zinc

could alleviate the growth inhibition by MCHM. Growth improved with the addition of 10 µM of

zinc sulfate in MCHM in both BY4741 and the *yke4* knockout (Figure 6B). However, at higher

zinc concentrations (100 µM), all growth was inhibited of all yeast when MCHM was added,

while zinc sulfate at this concentration alone did not alter growth (Supplemental Figure 5A).

Curiously, when zinc was added to YPD without MCHM, the media became slightly opaque.

However, after several days, media cleared around yeast colonies. YPD is an undefined media

composed of yeast extract, peptone, and agar. Zinc could have induced the precipitation of an

unknown compound or compounds that are solubilized by the growth of yeast on solid media;

the precipitation of these media components may limit yeast growth at this higher zinc

concentration. Therefore, we tested whether yeast knockouts of several known zinc

transporters would change the response to MCHM. First, the zinc tolerance of the *zrt1, zrt2,*

zrt3, and *zrc1* knockout yeast were tested. Only at the highest levels of zinc sulfate did the *zrc1*

mutant grow less than the other strains (Supplemental Figure 5A). Then, the addition of 5μM of

zinc sulfate completely rescued reduced growth in the presence of MCHM (Supplemental

Figure 5B). However, increasing zinc levels to 100μM further suppressed the growth of most of

the yeast tested grown on MCHM. The *zrc1* mutant did not grow better at 5 μM and was

completely inhibited at 100 μM of zinc in the presence of MCHM. The suppression of growth in

higher concentrations of zinc was not due to the toxicity of zinc alone, as all yeast tolerated

levels up to 500 μM with only *zrc1* mutant displaying reduced growth.

MCHM is composed of a saturated hexane ring with a methyl group and a methanol group at

opposite carbons (Figure 7A). The methanol and methyl group can be in the *cis* or *trans*

conformation. These characteristics of MCHM allow it to act as a hydrotrope, a compound that

can solubilize hydrophobic substances in aqueous environments. We thus considered MCHM's

role in altering protein-membrane and protein-protein interactions, which may explain the

impacts of MCHM on the transcriptome and ionome. *In vitro* protein aggregation was carried

out with sodium xylene sulfate (NaXS), an industrial hydrotrope, ATP, a biologically relevant

hydrotrope [28], and MCHM. Compared to no treatment (aggregation was set at 100% for no

treatment at 45 seconds), NaXS reduced aggregation to 48% (p= 0.0064), ATP reduced

aggregation to 3% (p=0.00025), and 550 ppm MCHM treatment reduced aggregation to 60%

(p=0.02). However, at one minute of incubation, MCHM allowed full protein aggregation and

was not distinguishable from untreated controls (p=28), while NaXS and ATP continued to

prevent protein aggregation (Figure 7B). Levels of zinc sulfate that rescued MCHM-induced

growth inhibition increased aggregation by 60% (Figure 7B). Zinc sulfate on its own caused nearly immediate aggregation of protein which was not prevented by the addition of MCHM. We tested if adding MCHM before zinc changed the rate of aggregation. When MCHM was added first followed by zinc sulfate, protein aggregation showed no difference at 45 seconds but was at the highest of all treatments tested at one minute of incubation.

Yeast have cell walls that protect cells from osmotic stress, and the cell wall can be easily removed to produce spheroplasts. However, unlike plants and fission yeast, spheroplasted yeast continue to divide their nuclei but don't undergo cytokinesis [35–37], leading to large multinucleated yeast. In this way, we can determine if hydrotropes can cause yeast to lyse when the cell wall is not providing rigid support. SDS, a commonly used detergent, reduced the optical density as yeast cells lysed. Sorbitol provides osmotic support and did not affect yeast growth (Figure 7C). Treatment of spheroplasts with known hydrotropes (ATP and NaXS) did not alter the growth of spheroplasted yeast over 15 hours, while the growth of spheroplasts treated with MCHM was arrested but did not cause lysis (Figure 7C). The dose-dependent reduction of spheroplasted growth likely mirrors the growth inhibition on plates with MCHM. The growth arrest is reversible in MCHM, as cells will continue growing after MCHM is removed (data not shown). The subcellular localization of Zrt1, Zrt2, Zrt3, and Yke4 was measured as cells were exposed to MCHM. Proteins were tagged at the N-terminus with mCherry [30], and cells were stained with calcofluor white to highlight the cell wall. Fluorescence of each protein remained diffuse and no foci appeared after 90 minutes of exposure (Supplemental Figure 6A). The endogenous promotors were replaced with a common constitutive promoter (*TEF2*), so any

increased expression of Zrt1, Zrt2, and Yke4 would likely be at the protein level rather than the mRNA level when exposed to MCHM. The protein levels of the Zrt transporters in YPD change by no more than 20% under the *TEF2* promoter, while Yke4 levels were 4-fold higher than Yke4 under its endogenous promoter [30]. Levels of *ZRT1* and *ZRT2* mRNAs were decreased with MCHM (Supplemental Table 2).

Discussion

The loss of the mitochondrial DNA and treatment with MCHM had pleiotropic effects on yeast. Petite yeast responded to stresses, including MCHM, differently than grande yeast. From RNA-seq analysis, iron and zinc transporters were differentially regulated in petite yeast and grande yeast in response to MCHM. In petite yeast, levels of zinc were lower while iron levels were higher. Transcriptomics pointed to metal transporters, while genetic analysis uncovered genetic variation in Yke4, an internal zinc transporter as important to MCHM. On non-fermentable carbon sources, *yke4* mutants do not grow in the presence of excess zinc [33], further highlighting that internal zinc transport in yeast differed in many other ways. To address how MCHM could affect the wide range of biochemical pathways seen, MCHM was tested and shown to be a hydrotrope *in vitro*, which could exert cell wall or membrane stress.

Petite yeast grew slower and were especially sensitive to the growth inhibition of MCHM. This may be in part due to the altered ionome of petite yeast. This includes higher levels of iron from increased expression of iron transporters and lower endogenous zinc levels. These petite

yeast were induced by loss of the mitochondrial genome, and petite yeast caused by other types of mutations also had differences in internal metals and regulation in the iron regulome [38]. There is an interplay between metal levels, as zinc transporters are also important for responding to high levels of copper [24]. Zrt1 protein levels increase in response to high levels of copper [24] and in contrast to the mRNA, Zrt1 under the control of a generic promoter and 5'UTR increased protein expressed by 66% in MCHM treatment. While genetic variation in Zrt2 contributes to copper tolerance [24], Zrt2 protein levels also increased by 26% with MCHM exposure. Supplementation with zinc alleviates copper-induced growth inhibition as well as MCHM growth inhibition.

The levels of sodium, calcium, phosphate, and magnesium also increased, suggesting drastic changes at the cell wall and membrane in response to MCHM. Although mRNAs of the iron acquisition pathway were increased in MCHM treatment, there was only a modest change in iron levels. Other stresses, such as starvation induced by rapamycin, also induce the iron regulon [39]. Deletion of *arn2* improved growth. Arn2 is localized to the ER and suggests a role in subcellular localization of iron in the yeast [30]. MCHM treated yeast are not starved for iron as *GTL1* and *GDH3* expression was not altered. These Gtl1 and Gdh3 are iron-dependent enzymes that are downregulated in iron-limiting media and upregulated in iron-replete media [40]. The addition of zinc rescued growth of yeast on MCHM at low levels of zinc (5-10 µM), while levels above 100-500 µM in combination with MCHM drastically reduced growth. Therefore, it appears that the metals have an optimum level to ameliorate the effects of MCHM. Based on RNA-seq and QTL analysis, two transporters of zinc were identified as having

an important role in MCHM response. We found no correlation between levels of internal zinc

and poor growth on MCHM, possibly because subcellular localization of zinc was critical to

growth, rather than absolute levels of the metal. There could also be a period of adjustment

that could not be captured due to the differences in timepoints between measuring metal

levels and growth. Metal levels were measured at 30 minutes of exposure while growth was

measured after two days. *YKE4* expression levels between YJM789 and S96 are not different in

unperturbed cells [41], and there are no SNPs in the 5'UTR or 3'UTR [42], suggesting that the

polymorphisms in Yke4 itself contributed to MCHM sensitivity in addition to other genetic

differences. While BY4741 *yke4* was MCHM sensitive, deletion in S288c and YJM789 had no

effect. There are hundreds of genetic differences between BY4741 and S288c and thousands in

YJM789 [43, 44]. The pleiotropic effects of MCHM on yeast combined with the many smaller,

yet not quite statistically significant peaks in the QTL point to MCHM resistance as a polygenic

trait likely spread throughout the genome. Other yeast strains have the H5Q polymorphism,

and a subset of those also have the F86L polymorphism. To separate the role of transcription

and 5'UTR dependence, the promoter and 5'UTR of *YKE4* was replaced. With MCHM treatment,

Yke4 protein levels increased 20%. 582 proteins are predicted to bind zinc with 20 proteins

binding 90% of total cellular zinc [45] and zinc sparing ensures that essential Zn binding proteins

have zinc. As YJM789 expresses less Zrt1 protein normally [24], perhaps the internal levels of

zinc and the ability of the yeast to quickly redistribute the zinc in MCHM may alter their ability

to grow.

Hydrotropes in cells prevent protein aggregation, but unlike surfactants, work at millimolar concentrations and display low cooperativity. In addition to changes in protein levels, post-translational modifications, subcellular location, and changes in protein conformation also regulate protein function. Protein aggregation is generally thought to inactivate proteins. Protein aggregation includes prions, which increase phenotypic plasticity without changing genetic diversity [46]. Intrinsically disordered regions of protein can separate proteins without being membrane-bound, which is an important step in RNA granule formation [47]. Transmembrane domain proteins such as Zrts, Yke4, and Arn transporters have multiple extracellular and intracellular loops that would be also disordered regions. MCHM slowed the aggregation of proteins in an *in vitro* assay, but with added zinc sulfate, which on its own induces some protein aggregation, MCHM appeared to combine to increase the rate of aggregation rather than preventing it. MCHM showed similar hydrotrope activity to NaXS, an industrial hydrotrope, but it was not as potent as ATP. Transcriptomics carried out after 90 minutes, approximately one generation in yeast, detected changes in mRNA encoding metal transporters. Zinc is required for both the synthesis of cell walls and phospholipids [33, 48]. Exposure of yeast to MCHM increased intracellular sodium levels yet yeast do not actively accumulate sodium [49], further supporting that MCHM alters the protein structures to either increase the bioavailability of ions or transport across cell membranes. MCHM altered the levels of ions in the cell at the earliest time points. Therefore, we conclude that this is likely through altered protein conformation at the cell membrane of many proteins because of the diverse metals that changed during MCHM exposure. Proteins and organelles are increasingly found in altered conformations to change local concentrations of proteins in (reviewed [50]).

Inside the cell, MCHM could alter how molecules interact in liquid droplet changing the function of protein functions and metabolites in the cell.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions: JEGG designed the experiments. AP carried out RNA-seq analysis. ZNS measured mCherry tagged proteins, tested protein aggregation and with RJV and MCA carried out yeast growth assays. MCA carried out QTL analysis and constructed strains. JRC analyzed results from ICP and assisted in writing.

References

1.   Maugh TI (1983) How Many Chemicals Are There? Science (80-) 220:293

2.   Guerrero PF (1994) Toxic Substances Control Act: EPA's limited progress in regulating toxic chemicals. . http://archive.gao.gov/t2pbat3/151661.pdf

3.   Lan J, Hu M, Gao C, et al (2015) Toxicity Assessment of 4-Methyl-1-cyclohexanemethanol and Its Metabolites in Response to a Recent Chemical Spill in West Virginia, USA. Environ Sci Technol

4.   Tullo AH, Kemsley J, Hogue C, Morrissey SR (2014) Obscure Chemical Taints Water Supply. Chem Eng News 92:10–15

5.   Hsu J, del Rosario MC, Thomasson E, et al (2017) Hospital Impact After a Chemical Spill That Compromised the Potable Water Supply: West Virginia, January 2014. Disaster Med Public Health Prep 11:621–624. https://doi.org/10.1017/dmp.2016.193

6.   Cui D, Mebel AM, Arroyo-Mora LE, et al (2017) Kinetic, product, and computational studies of the ultrasonic induced degradation of 4-methylcyclohexanemethanol (MCHM). Water Res 126:164–171. https://doi.org/10.1016/j.watres.2017.09.005

7.   Weidhaas JL, Dietrich AM, DeYonker NJ, et al (2016) Enabling Science Support for Better Decision-Making when Responding to Chemical Spills. J Environ Qual 45:1490. https://doi.org/10.2134/jeq2016.03.0090

8.   Gallagher JEG, Zheng W, Rong X, et al (2014) Divergence in a master variator generates distinct phenotypes and transcriptional responses. Genes Dev 28:409–421

9.   Rong-Mullins XX, Ayers MC, Summers M, Gallagher JEG (2018) Transcriptional Profiling of Saccharomyces cerevisiae Reveals the Impact of Variation of a Single Transcription Factor on Differential Gene Expression in 4NQO, Fermentable, and Nonfermentable Carbon Sources. G3 Genes, Genomes, Genet 8:. https://doi.org/10.1534/g3.117.300138

10.   Grant CM, MacIver FH, Dawes IW (1997) Mitochondrial function is required for resistance to oxidative stress in the yeast *Saccharomyces cerevisiae*. FEBS Lett 410:219–222. https://doi.org/10.1016/S0014-5793(97)00592-9

11.   Thorpe GW, Fong CS, Alic N, et al (2004) Cells have distinct mechanisms to maintain protection against different reactive oxygen species: Oxidative-stress-response genes. Proc Natl Acad Sci U S A 101:6564–6569. https://doi.org/10.1073/pnas.0305888101

12.   Liang Q, Zhou B (2007) Copper and manganese induce yeast apoptosis via different pathways. Mol Biol Cell 18:4741–9. https://doi.org/10.1091/mbc.E07-05-0431

13.   Himelblau E, Mira H, Lin SJ, et al (1998) Identification of a functional homolog of the yeast copper homeostasis gene ATX1 from Arabidopsis. Plant Physiol 117:1227–34

14.   Zhao H, Eide D (1996) The yeast ZRT1 gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation. Proc Natl Acad Sci 93:2454–2458

15.   Lyons TJ, Gasch AP, Gaither LA, et al (2000) Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. Proc Natl Acad Sci U S A 97:7957–7962. https://doi.org/97/14/7957 [pii]

16.   MacDiarmid CW, Gaither LA, Eide D (2000) Zinc transporters that regulate vacuolar zinc storage in Saccharomyces cerevisiae. EMBO J 19:2845–2855. https://doi.org/10.1093/emboj/19.12.2845 [doi]

17.   Conklin DS, McMaster JA, Culbertson MR, Kung C (1992) COT1, a gene involved in cobalt accumulation in Saccharomyces cerevisiae. Mol Cell Biol 12:3678–88. https://doi.org/10.1128/MCB.12.9.3678

18.   Lyons TJ, Villa NY, Regalla LM, et al (2004) Metalloregulation of yeast membrane steroid receptor homologs. Proc Natl Acad Sci U S A 101:5506–11. https://doi.org/10.1073/pnas.0306324101

19.   Byrne KP, Wolfe KH (2005) The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res 15:1456–61. https://doi.org/10.1101/gr.3672305

20.   Giaever G, Chu AM, Ni L, et al (2002) Functional profiling of the Saccharomyces cerevisiae genome. Nature 418:387–391

21.   Goldstein AL, McCusker JH (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast 15:1541–1553

22.   McCusker JH, Clemons K V, Stevens DA, Davis RW (1994) Genetic characterization of pathogenic Saccharomyces cerevisiae isolates. Genetics 136:1261–1269

23.   Steinmetz LM, Sinha H, Richards DR, et al (2002) Dissecting the architecture of a quantitative trait locus in yeast. Nature 416:326–330

24.   Rong-Mullins X, Winans MJ, Lee JB, et al (2017) Proteomic and genetic analysis of S. cerevisiae response to soluble copper leads to improvement of antimicrobial function of cellulosic copper nanoparticles. Metallomics 9:1304–1315. https://doi.org/10.1039/C7MT00147A

25.   Rong-Mullins X, Ravishankar A, McNeal KA, et al (2017) Genetic variation in Dip5, an amino acid permease, and Pdr5, a multiple drug transporter, regulates glyphosate resistance in S. cerevisiae. PLoS One 12:e0187522. https://doi.org/10.1371/journal.pone.0187522

26.   Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omi A J Integr Biol 16:284–287. https://doi.org/10.1089/omi.2011.0118

27.   Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

28.   Patel A, Malinovska L, Saha S, et al (2017) ATP as a biological hydrotrope. Science (80-) 356:753–756. https://doi.org/10.1126/science.aaf6846

29.   Katzmann DJ, Wendland B (2005) Analysis of Ubiquitin-Dependent Protein Sorting Within the Endocytic Pathway in Saccharomyces cerevisiae. Methods Enzymol 399:192–211. https://doi.org/10.1016/S0076-6879(05)99013-7

30.   Weill U, Yofe I, Sass E, et al (2018) Genome-wide SWAp-Tag yeast libraries for proteome exploration. Nat Methods 15:617–622. https://doi.org/10.1038/s41592-018-0044-9

31.   Cyert MS, Philpott CC (2013) Regulation of cation balance in Saccharomyces cerevisiae. Genetics 193:677–713

32.   Eide DJ, Clark S, Nair TM, et al (2005) Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in Saccharomyces cerevisiae. Genome Biol 6:R77. https://doi.org/10.1186/gb-2005-6-9-r77

33.   Kumánovics A, Poruk KE, Osborn KA, et al (2006) YKE4 (YIL023C) encodes a bidirectional zinc transporter in the endoplasmic reticulum of Saccharomyces cerevisiae. J Biol Chem 281:22566–74. https://doi.org/10.1074/jbc.M604730200

34.   Krogh A, Rn Larsson BÈ, Von Heijne G, Sonnhammer ELL (2001) Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315

35.   Komiyama T, Kimura T, Furuichi Y (2002) Round shape enlargement of the yeast spheroplast of Saccharomyces cerevisiae by HM-1 toxin. Biol Pharm Bull 25:959–65

36.   Fukudomi T, Kotani T, Miyakawa I (2011) A simple method for culture and stable maintenance of giant spheroplasts from the yeast Saccharomyces cerevisiae. J Gen Appl Microbiol 57:177–82

37.   Kelly FD, Nurse P (2011) De Novo Growth Zone Formation from Fission Yeast Spheroplasts. PLoS One 6:e27977. https://doi.org/10.1371/journal.pone.0027977

38.   Xue Y, Schmollinger S, Attar N, et al (2017) Endoplasmic reticulum-mitochondria

junction is required for iron homeostasis. J Biol Chem 292:13197–13204.

https://doi.org/10.1074/jbc.M117.784249

39.   Preiss T, Baron-Benhamou J, Ansorge W, Hentze MW (2003) Homodirectional changes in

transcriptome composition and mRNA translation induced by rapamycin and heat shock. Nat

Struct Mol Biol 10:1039–1047. https://doi.org/10.1038/nsb1015

40.   Shakoury-Elizeh M, Tiedeman J, Rashford J, et al (2004) Transcriptional Remodeling in

Response to Iron Deprivation in *Saccharomyces cerevisiae*. Mol Biol Cell 15:1233–1243.

https://doi.org/10.1091/mbc.e03-09-0642

41.   Zheng W, Zhao H, Mancera E, et al (2010) Genetic analysis of variation in transcription

factor binding in yeast. Nature 464:1187–1191

42.   Xu Z, Wei W, Gagneur J, et al (2009) Bidirectional promoters generate pervasive

transcription in yeast. Nature 457:1033–1037

43.   Song G, Dickins BJ, Demeter J, et al (2015) AGAPE (Automated Genome Analysis

PipelinE) for pan-genome analysis of Saccharomyces cerevisiae. PLoS One 10:e0120671.

https://doi.org/10.1371/journal.pone.0120671 [doi]

44.   Wei W, McCusker JH, Hyman RW, et al (2007) Genome sequencing and comparative

analysis of Saccharomyces cerevisiae strain YJM789. Proc Natl Acad Sci U S A 104:12825–12830.

https://doi.org/10.1073/pnas.0701291104

45.   Wang Y, Weisenhorn E, MacDiarmid CW, et al (2018) The cellular economy of the *Saccharomyces cerevisiae* zinc proteome. Metallomics 10:1755–1776. https://doi.org/10.1039/C8MT00269J

46.   Snell-Rood EC, Van Dyken JD, Cruickshank T, et al (2010) Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. Bioessays 32:71–81

47.   Lin Y, Protter DSW, Rosen MK, Parker R (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. Mol Cell 60:208–19. https://doi.org/10.1016/j.molcel.2015.08.018

48.   Iwanyshyn WM, Han G-S, Carman GM (2004) Regulation of phospholipid synthesis in Saccharomyces cerevisiae by zinc. J Biol Chem 279:21976–83. https://doi.org/10.1074/jbc.M402047200

49.   Haro R, Rodríguez-Navarro A (2002) Molecular analysis of the mechanism of potassium uptake through the TRK1 transporter of Saccharomyces cerevisiae. Biochim Biophys Acta - Biomembr 1564:114–122. https://doi.org/10.1016/S0005-2736(02)00408-X

50.   Boeynaems S, Alberti S, Fawzi NL, et al (2018) Protein Phase Separation: A New Phase in Cell Biology. Trends Cell Biol 28:420–435. https://doi.org/10.1016/J.TCB.2018.02.004

51.   Pupo A, Ayers MC, Sherman ZN, et al (2019) MCHM acts as a hydrotrope, altering the balance of metals in yeast. bioRxiv 606426. https://doi.org/10.1101/606426
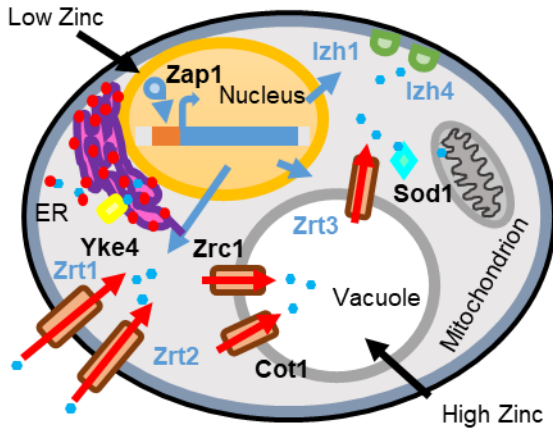
Figures:

Figure 1 Schematic of yeast cellular response to varying levels of zinc. Under low intracellular zinc levels, Zap1 induces expression of zinc uptake genes. Zrt1, the high-affinity, and Zrt2 the low-affinity transmembrane zinc transporter, are localized in the cell membrane. Izh1 and Izh4 localize to the cytoplasmic side of the cell membrane and regulate zinc homeostasis. Excess zinc (blue hexagons) is stored in the vacuole and transported in by Zrc1 and Cot1. Excess zinc is stored in the vacuole, and when needed, Zrt3 transports zinc to the cytoplasm. In addition to the vacuole, the ribosome (red circles) bind 20% of total cellular zinc [45]. Yke4 is the transmembrane transporter at the ER (pink and purple) that transports zinc in both directions. Sod1 is the copper-zinc super oxygen dismutase that is both cytoplasmic and localizes to the inner mitochondrial membrane. (Jennifer Gallagher made model figure)
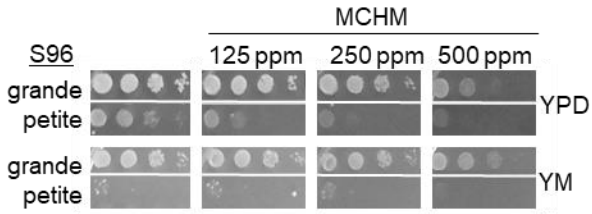
Figure 2 Serial dilution of wild-type S96 and petite yeast in increasing concentration of MCHM in rich (YPD) and minimal (YM) media. Plates were incubated at 30°C for three days and then photographed. (Jennifer Gallagher did petite/grande yeast growth assays)
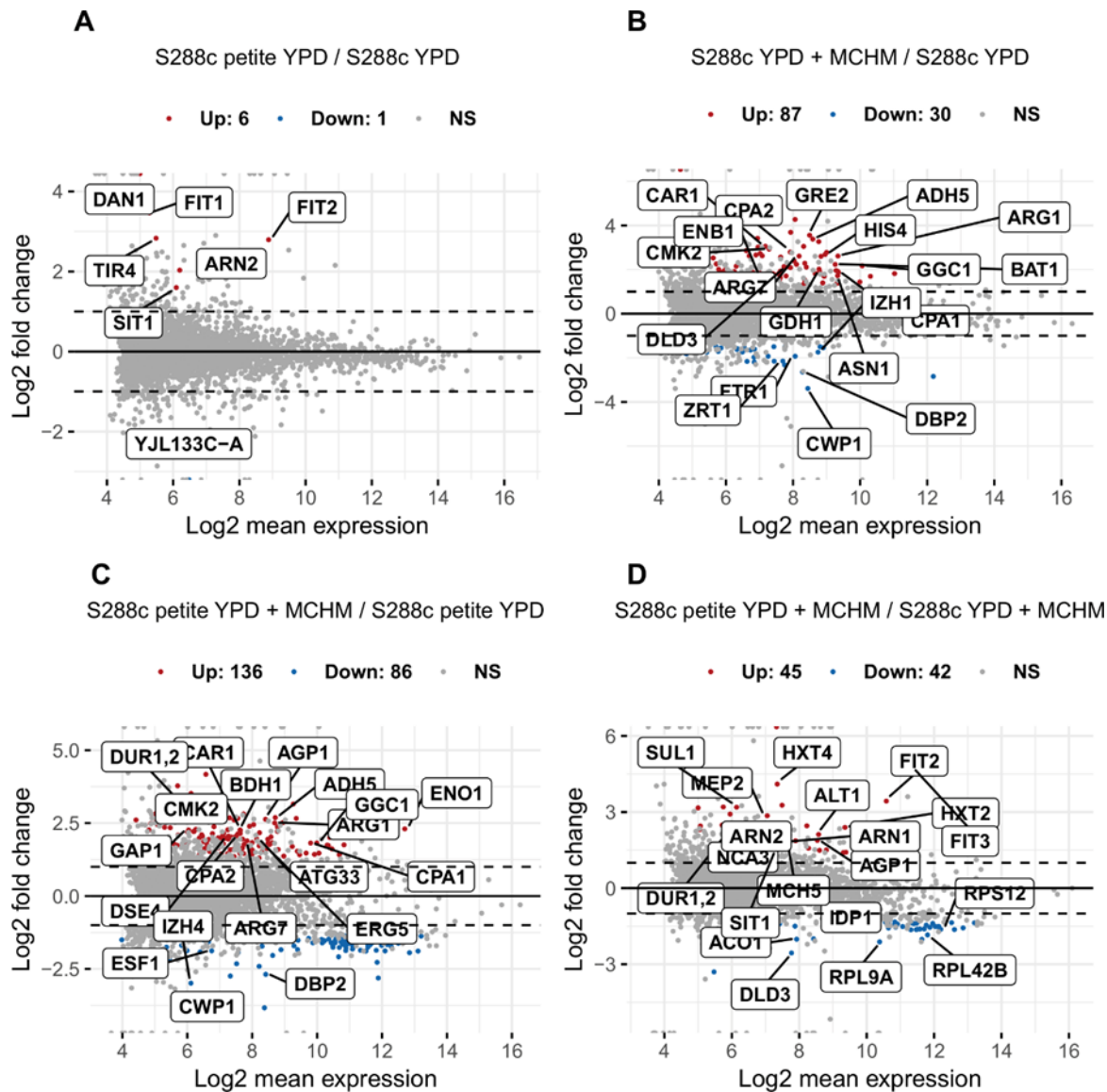
Figure 3 Changes in RNA expression between wild type and petite yeast grown in YPD treated with MCHM. The number of up- and downregulated genes is noted on the top of each panel. (A) Scatter plots of log fold 2 comparisons of RNA-seq from grande (S96) and petite (S96ρ) yeast grown in YPD. Significantly upregulated genes are labeled in red and significantly downregulated genes are labeled in blue. (B) Scatter plots of log fold 2 comparisons of RNA-seq grande yeast grown in YPD and with 550 ppm MCHM. (C) Scatter plots of log fold 2 comparisons of RNA-seq from petite yeast grown in YPD and with MCHM. (D) Scatter plots of log fold 2 comparisons of RNA-seq from S96 and S96◻ yeast grown in YPD and with MCHM. (Amaury Pupo performed differential expression analysis)
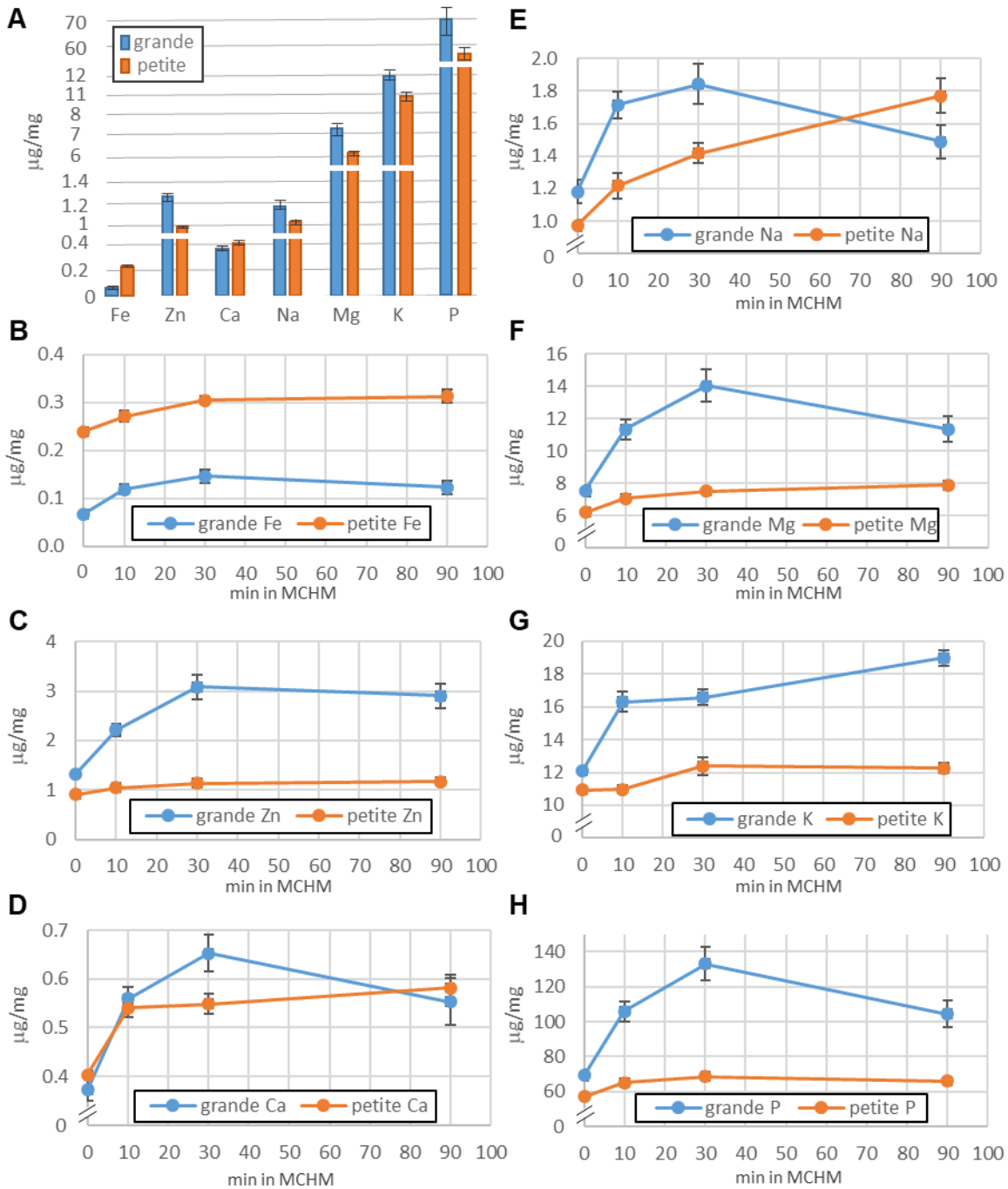
Figure 4 Measurements of metals in yeast treated with MCHM (A) Levels of Fe, Zn, Ca, Na, Mg, K and K in µg/mg from grande (S96) and petite (S96ρ) yeast grown in YPD. Levels of metals from grande (S96) and petite (S96ρ) yeast grown in YPD with 550 ppm MCHM added for the indicated time. The ions levels measured were (B) Fe, (C) Zn, (D) Ca, (E) Na, (F) Mg, (G) K and (H) F. The standard error is noted on the mean of four biological replicates. (Jennifer Gallagher and Jonathan Cumming performed ICP analysis)

110
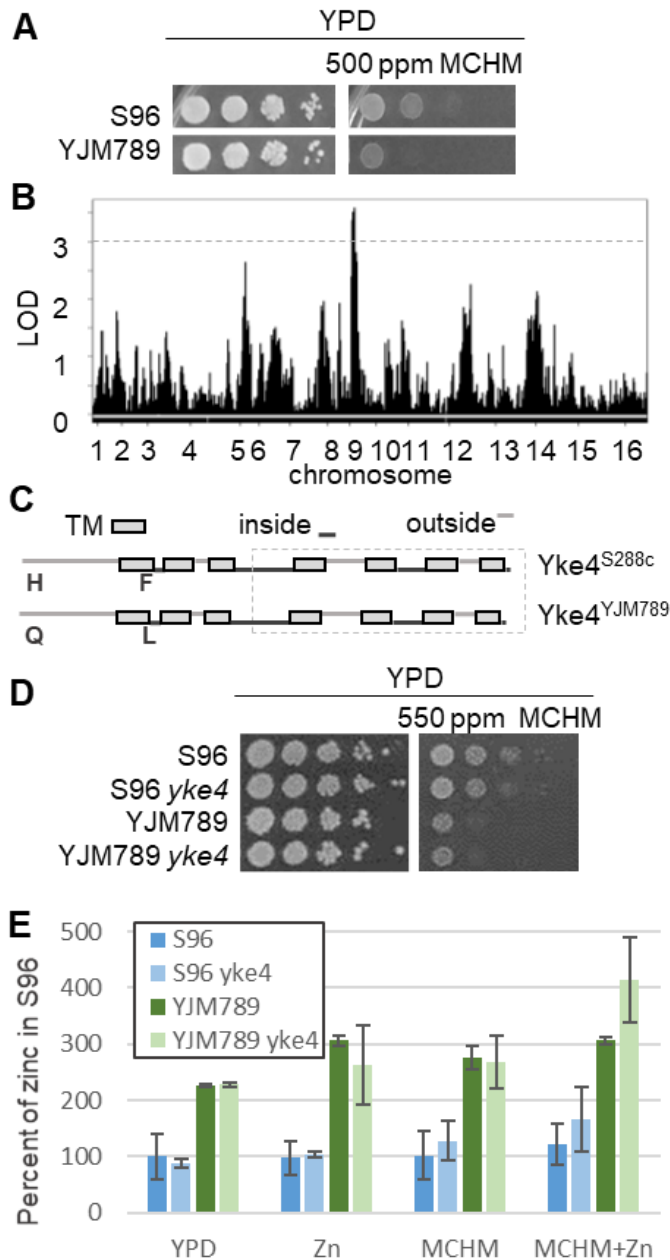
Figure 5 Genetic variation of MCHM response linked to the YKE4 locus. (A) Serial dilution of genetically diverse yeast on 550 ppm MCHM on YPD. S96 are from the S288c background and YJM789 is a clinical isolate. (B) Quantitative trait loci analysis of chromosomal regions linked to increased growth of yeast in MCHM from YJM789 and S96 segregants. (C) Diagram of Yke4 describes the transmembrane domains (TM noted with grey box) and inside or outside domains (noted with lower dark grey or upper light grey lines, respectively). Polymorphic residues are noted below in Yke4[YJM789] compared to Yke4S288c. The conserved ZIP domain is boxed in a grey

dashed line. (D) Serial dilutions of S96 and YJM789 with *yke4* mutants on MCHM. (E) Levels of total intracellular zinc from S96, YJM789 and *yke4* mutants normalized to S96 grown in YPD and total protein. Yeast were incubated with 5 μM zinc sulfate, 550 ppm MCHM and in combination for 30 minutes before metals were extracted. (Michael C. Ayers performed yeast growth assay and phenotyping and computational analyses for QTL in parts A through C. Michael C. Ayers constructed yeast knockout strains and performed growth analyses in part D. Jennifer Gallagher and Jonathan Cumming performed ICP analysis of yeast strains from part D in part E.)
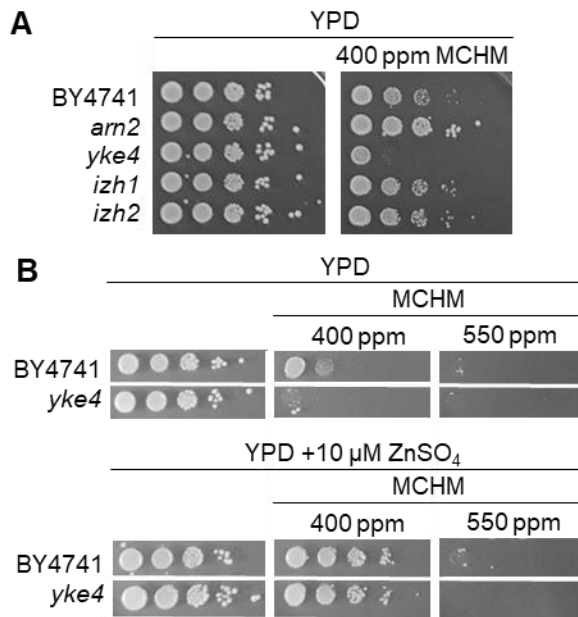
Figure 6 Impact of loss of metal transporters on growth with MCHM and zinc-containing media. (A) Serial dilution of BY4741 and yeast from the knockout collection on YPD with 400 ppm MCHM. (B) Serial dilution of wild type (BY4741) and zinc transporters knocked out yeast grown in 400 or 550 ppm MCHM on YPD with 10μM zinc sulfate. (Michael Ayers performed all experiments.)
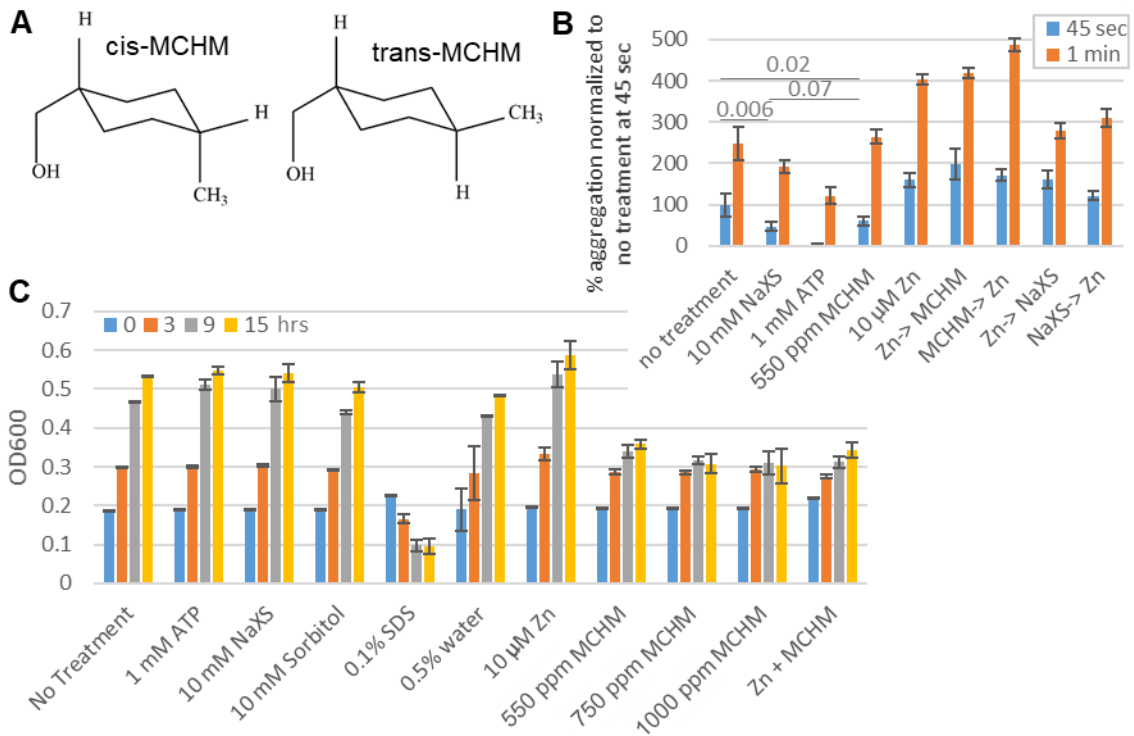
Figure 7 MCHM role in protein aggregation. (A) Chemical structure of the cis and trans conformation of MCHM. (B) Timed protein aggregation when exposed to heat. No treatment is compared to 10 mM sodium xylene sulfate, 1mM ATP, and 550ppm MCHM. The optical density of samples was measured at 450nm after incubation at 60°C. 550 ppm MCHM was added to samples then 10 ᴎM zinc sulfate was added where indicated. (C) Spheroplast yeast were incubated for between 0 and 15 hours with 0.1% SDS, 0.5% water, 10 mM sorbitol, 1 mM ATP, 10 mM NaXS, 10 µM zinc sulfate and 550-1000 ppm MCHM and samples were read at 600 nm. Biological quadruplicates were averaged and standard error. Relevant p values were calculated using the student t-test. (Zachary Sherman performed protein aggregation and Tecan growth experiments in parts B and C.)

The following Supplemental Materials are available at
for download.

Supplemental Table 1. Comparisons of RNA-seq from S96 and S96 petite yeast treated with 550 ppm MCHM for 90 minutes in YPD and YM supplemented with lysine.

Supplemental Table 2. Comparisons of zinc related genes from S96 and S96 petite yeast treated with 550 ppm MCHM for 90 minutes in YPD and YM supplemented with lysine.

Supplemental Table 3. Comparisons of iron related genes from S96 and S96 petite yeast treated with 550 ppm MCHM for 90 minutes in YPD and YM supplemented with lysine.

Supplemental Table 4. Levels of metals from S96 grande and petite yeast treated with 550 ppm MCHM grown over 90 minutes in YPD. The mean of four biological replicates is shown with standard error. Levels of metals washed with water compared to yeast washed with EDTA.

Supplemental Figure 1. Scatter plots of log fold 2 comparisons of RNA-seq from grande S96 and petite S96 yeast grown in YM supplemented with lysine. Significantly upregulated genes are labeled in red and significantly downregulated genes are labeled in blue. (A) Scatter plots of log fold 2 comparisons of RNA-seq petite and grande yeast grown in YM. (B) Scatter plots of log fold 2 comparisons of RNA-seq grande yeast grown in YM and with 550 ppm MCHM. (C) Scatter plots of log fold 2 comparisons of RNA-seq from petite yeast grown in YM and with MCHM. (D) Scatter plots of log fold 2 comparisons of RNA-seq from grande and petite yeast grown in YM and with MCHM. (Amaury Pupo performed analyses on transcriptome data.)

Supplemental Figure 2. GO term analysis of genes upregulated in all pairwise comparisons with single variable between S96 grande and S96 petite yeast grown in YM and YPD with MCHM added. The color scale of p adjust values are noted and the size of the circle notes gene ratio. The number of genes in each comparison is noted in parentheses. (Amaury Pupo performed analyses on transcriptome data.)

Supplemental Figure 3. GO term analysis of genes downregulated in all pairwise comparisons with single variable between S96 grande and S96 petite yeast grown in YM and YPD with MCHM added. The color scale of p adjust values are noted and the size of the circle notes gene ratio. The number of genes in each comparison is noted in parentheses. (Amaury Pupo performed analyses on transcriptome data.)

Supplemental Figure 4. Levels of Fe and Zn in grande and petite S96 yeast washed with water and EDTA or with only water before metal extraction. Mean of four biological replicates are shown with standard error. (Jennifer Gallagher and Jonathan Cumming performed ICP.)

Supplemental Figure 5. Role of zinc transporters in MCHM response with supplemented zinc. (A) Serial dilution of wild-type (BY4741) and zinc transporter knockout yeast were grown on YPD supplemented with increasing concentrations of zinc sulfate. (B) Serial dilution of wild-type (BY4741) and zinc transporter knockout yeast grown in 400 ppm MCHM on YPD supplemented with increasing concentrations of zinc sulfate. (Michael C. Ayers performed yeast knockout growth assays.)

Supplemental Figure 6. Levels of mCherry tagged Zrt1, Zrt2, Zrt3, and Yke4 proteins change with MCHM exposure. Yeast were exposed to 550 ppm of MCHM in YPD for 30 minutes. (A)The cell wall was stained with calcofluor white. (B) The mean fluorescence of mCherry normalized to untreated yeast for each protein. Quantification for 17-20 cells shown with standard error. (Zachary Sherman performed mCherry tagged yeast growth assays.)

CHAPTER 5: LABORATORY EVOLUTIONS LEAD TO REPRODUCIBLE MUTATIONS TO *PDR3* CONFERRING RESISTANCE TO MCHM

Michael C. Ayers and Jennifer E.G. Gallagher*

Department of Biology

*Corresponding Author

Abstract

Cell changes in response to the coal-cleaning chemical 4-methylhexanemethanol (MCHM) have previously been shown to involve nutrient and oxidative stress responses, metal homeostasis, and membrane sterol content control. MCHM acts as a hydrotrope, affecting the folding and aggregation of proteins and membrane stability. Nonetheless, the information gleaned from affected pathways has not provided a unifying mechanism sufficient to induce resistance. We produced eight strains of *S. cerevisiae* resistant to MCHM by using In-Lab evolution from a single YJM789 parent strain, and we analyzed the genomic changes that provided this new resistant phenotype. We identified thousands of SNP and indel variants per strain, which was a consistent number between strains that evolved resistance and control strains that remained sensitive. These variants did not show a pattern that would cluster resistant strains together.

The many background mutations likely masked any pattern from few large-effect loci or implicated an epistatic effect of many small mutations spread throughout the genome that was undetectable. Among coding variants in the strains which change protein sequence and thereby may alter function, only one gene showed a protein-coding mutation in every resistant strain while showing no variants at all in the control strains. This gene, *PDR3*, controls transcription for the pleiotropic drug response and is likely the most significant driver of adaptive MCHM resistance in yeast. While many of the evolved alleles of *PDR3* would likely produce functional proteins, a knockout in the parent YJM789 strain was sufficient to produce resistance to MCHM. This transcription factor is the first gene where mutations are known to increase resistance to this novel hydrotropic chemical.

Introduction

The chemical 4-methylcyclohexanemethanol (MCHM) came to relevance when a crude formulation used in coal-cleaning processes spilled into the Elk River outside Charleston, WV in January 2014. The spill affected the drinking water of approximately 300,000 residents, so research became widespread to characterize everything from the volatility and adsorption properties of the chemical in water systems to the toxicological effects on various model organisms (Han et al., 2017; Horzmann et al., 2017; Phetxumphou et al., 2016; Sain et al., 2015; Weidhaas et al., 2016; *West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). The cellular biological effects of the chemical have been explored in more detail in yeast than in other organisms.

More recent work has produced genomic datasets to determine the cellular pathways involved

in MCHM response. MCHM is a hydrotrope which affects protein folding and has been

implicated to have a role in zinc ion concentration changes in the cell to alleviate this stress

(Pupo, Ayers, et al., 2019). Lipid biosynthetic changes induced by MCHM may also be a source

of stress on the plasma membrane (Pupo, Ku, et al., 2019). MCHM also causes amino acid

accumulation according to metabolic data, while exhibiting a nutrient starvation signal

activating the environmental stress response (Ayers et al., 2020; Pupo, Ku, et al., 2019). One

early study using stress gene reporters showed oxidative stress and DNA damage response

activation in response to MCHM and its potential metabolites (Lan et al., 2015). The more

recent work using yeast genomic datasets confirmed MCHM causes the production of reactive

oxygen species (ROS) in yeast cells, as well as DNA damage, making yeast strains sensitive to

ROS, such as petite yeast, unable to grow (Ayers et al., 2020; Lan et al., 2015; Pupo, Ayers, et

al., 2019). This data all points to widespread cellular pathways important for MCHM tolerance.

In-Lab evolution (ILE) of yeast is a tool that allows for the genetic analysis of many

phenotypes and processes of evolution. The short generation time of yeast allows for the quick

production of large populations that accumulate mutations that may respond to selection or

drift in the lab. This has been used to experimentally observe evolution in industrial

environments with yeast hybrid strains to determine if they evolve similarly in terms of

aneuploidy and copy number to historical lager brewing yeast hybrids (Gorter de Vries et al.,

2019) and with wine strains where ploidy events control significant changes in fermentation

efficiency (Mangado et al., 2018). Experiments also explore processes such as yeast and

bacterial populations in competition and yeast populations that converge towards phenotypes

with similar fitness due to epistatic interactions of beneficial mutations (Kryazhimskiy et al., 2014; Zhou et al., 2018). To understand stress resistance, ILE has been used to produce yeast adapted to a multitude of stressors. Adaptation to copper has been produced in *S. cerevisiae* and *Candida humilis*, revealing mechanisms such as overexpression of superoxide dismutases and catalases to detoxify ROS and overexpression of proteins that bind excess copper (Adamo et al., 2012). Thermotolerance was also evolved in yeast stains to explore mechanisms that could improve industrial biomass to ethanol conversion processes. Cellular changes involving sterol components of the membrane and concentrations of glycerol also provided increased tolerance to stresses from osmolarity and excess glucose and ethanol (Caspeta & Nielsen, 2015). The evolution of resistance to a glyphosate-based herbicide in the lab was shown to involve copy number changes in proteins affecting cell wall stability in response to the presumably inert non-glyphosate additives in the herbicide (Ravishankar et al., 2020). ILE experiments are an unbiased and insightful method for understanding specific mechanisms and processes of adaptation to environments.

Yeast cells use the pleiotropic drug response to detoxify and remove a multitude of general xenobiotics and chemical stressors. Much of the work of this response is performed by a class of proteins called ATP-binding cassette (ABC) transporters that translocate chemicals out of the cell and are a conserved group of proteins throughout the tree of life (reviewed by Jungwirth & Kuchler, 2006). Some important stress-responsive yeast ABC transporters include Pdr5, Snq2, Pdr15, and Yor1, all of which have roles in exporting drugs or various chemically unrelated toxic chemicals out of the cell in response to stress (Balzi et al., 1994; Bissinger & Kuchler, 1994; Jungwirth & Kuchler, 2006; Katzmann et al., 1999; Kolaczkowski et al., 1998;

120

Servos et al., 1993; Wolfger et al., 1997; Wolfger et al., 2004). The transcriptional control of the genes encoding these proteins involves the interplay of the transcription factors Pdr1 and Pdr3, two paralogs sharing 36% amino acid sequence similarity along the entire length of the proteins (Delaveau et al., 1994). These proteins form heterodimers and homodimers that compete to occupy *cis*-regulatory elements termed PDREs upstream of pleiotropic drug response genes (Katzmann et al., 1996; Mamnun et al., 2002). Their interaction in different combinations and on PDREs of different genes likely plays a role in the relative inhibition and activation of these same genes (Mamnun et al., 2002). The importance of the ABC transporters in a cell's response to chemical stressors is important for evolved stress adaptation. In particular, induced mutagenesis experiments on *PDR3* have produced activated gain-of-function alleles where Pdr3 increases the expression of *PDR5* and *SNQ2* and increases resistance to a multitude of chemicals (Nourani et al., 1997; Simonics et al., 2000). The role of the ABC transporters in the response to chemical stress makes them and the proteins that control their expression targets for adaptation to chemicals such as MCHM.

In this study, we utilized In-Lab evolutions to produce strains adapted to the spill chemical MCHM. The goals of the study were to identify mechanisms as yet uncovered in the cellular response to MCHM as well as explore the evolutionary process under novel stress that yeast were unlikely to have encountered in their environmental niche. Eight strains evolved in MCHM became resistant to high dosages of the chemical, while the eight control strains evolved under similar conditions without MCHM remained sensitive. All 16 strains accumulated similar numbers of mutations overall. There were no clear patterns in mutations that led to resistance, except for variants appearing in the gene *PDR3*. The mutations in this gene were the

121

major reproducible drivers of resistance to MCHM, but other mutations spread across the genome likely contributed to the convergent phenotype through epistatic interactions of small effects.

Materials and Methods

In-Lab evolutions:

The YJM789 strain (MATalpha, *lys2Δ*) (McCusker et al., 1994b, 1994a; Tawfik et al., 1989; Wei et al., 2007) was grown in biological quadruplicate in 2mL liquid cultures of YPD (1% yeast extract, 2% peptone, 2% dextrose) media with and without 700ppm crude MCHM at 30°C. Cultures were grown for two days before 1% of the culture was used to inoculate fresh tubes containing 2mL of YPD or YPD+700ppm MCHM media. This passaging process was performed for six passages before plating serial dilutions onto YPD plates containing ranges of 0-1000ppm MCHM. Two colonies were isolated from each final evolved population. The isolates were grown for 2-3 passages in liquid YPD without MCHM to allow for removal of epigenetic memory of resistance that may impact growth on MCHM media. The resistance of the 16 isolated strains, eight control evolutions, and eight MCHM evolutions, was evaluated by growing on 1000ppm MCHM solid media (YPD as above, plus 2% agar). Following the epigenetic check consisting of passaging isolates in the absence of MCHM then rechecking for continued resistance, genomic DNA was then isolated from the strains and submitted for sequencing at the WVU Genomics Core as previously (Ravishankar et al., 2020).

Sequencing and analysis:

Strains were sequenced on an Illumina Miseq producing 151bp paired-end reads of between 1,638,426 and 3,445,498 reads per strain. Raw sequence reads and vcf files produced by the mapping and variant calling analysis are available on figshare at the following DOI link: 10.6084/m9.figshare.13028222. Reads were aligned to the S288c reference genome R64.2.1 release obtained from yeastgenome.org (Engel et al., 2013) using BWA version 0.7.17-r1188 (Li & Durbin, 2009), giving coverage of approximately 20-30x across the genome for each strain. SAM and BAM files were created using samtools version 1.7 (Li, 2011; Li et al., 2009). Variants were called using the HaplotypeCaller function of GATK 4.1.6.0 (Van der Auwera et al., 2013) on Java Runtime Environment 11.0.6. Variants were filtered by removing all variants that did not pass GATK quality measures. As YJM789 has approximately 60,000 SNPs (Wei et al., 2007) when compared with the S288c reference, these variants were removed by filtering all variants that were shared amongst all 16 evolved strains (control and MCHM) and using an existing SNP list for the strains. Any variants shared by all 16 strains were assumed to be existing variation between S288c and YJM789, as opposed to random mutations that happened in every sample. The 6874 remaining variants in the 16 strains were analyzed for presence in coding regions and the intersection between strains as below. The numbers of variants per strain in the eight control strains and eight MCHM strains were compared with a two-tailed Student's T-test to determine if there was any difference in total number of variants between resistant phenotypes.

Variants were analyzed further using R version 3.6.3 and the following packages: VariantAnnotation version 1.32.0 (Obenchain et al., 2014), pcadapt 4.3.1 (Luu et al., 2020), BSgenome.Scerevesiae.UCSC.sacCer3_1.4.0 (Pagès, 2019; Team, 2014),

TxDb.Scerevisiae.UCSC.sacCer3.sgdGene_3.2.2 (Carlson & Maintainer, 2015), and all dependencies are available in Supplemental Table 1. The entire R analysis script and supplemental files and tables are available at https://github.com/wvumayers/MCHM.git. In short, the 6874 variants that remained following filtering were analyzed for PCA clustering analysis using the pcadapt package with a k=3 based on screeplot. Variants were also plotted for their density across the genome by first creating a table of approximately 40000 variants by including every instance of the 6874 unique variants in all 16 strains, as many variants appeared in multiple strains. A second plot was made with variants shared by less than 10 strains, reducing the table to approximately 15000 variants across all strains from 4968 unique variants. The plots were produced using 2.5kb windows in the karyoploteR package, version 1.14.1, in R on the sacCer3 genome (Gel & Serra, 2017). Variants then were filtered to remove all those found in any of the control strains, regardless of presence in the MCHM strains. This left 1171 variants existing only in at least one of the eight MCHM evolved strains. These variants were analyzed for their effect on coding sequences using the predictCoding command of the VariantAnnotation package and the UCSC sacCer3 genome and transcript packages mentioned above. There are three protein/ORF annotation differences between the sacCer3 version of the genome and the S288c R64.2.1 version of the genome used to map the original sequences and call variants that did not affect the coding predictions of this dataset. The SIFT 4G variant annotator software was also used to determine the predicted tolerance score for each variant from the original ILE vcf files (Vaser et al., 2016). The *Saccharomyces cerevisiae* R64-1-1.23 database included with the SIFT4G jar was used to annotate the variants.

GO terms and intersection of strain variants:

GO terms for the 107 genes that contained nonsynonymous and synonymous variants in the MCHM evolved strains were determined using the DAVID bioinformatic database at https://david.ncifcrf.gov (Huang et al., 2009a, 2009b). Default options were used for the analysis. The intersection analysis of strains containing the same genes was performed using the shiny app tool version of UpSetR at https://gehlenborglab.shinyapps.io/upsetr/(Lex et al., 2014).

PDR3 knockouts and allele analysis:

The YJM789 *PDR3*::*NAT* knockout was produced by transforming PCR construct containing the NAT^R gene amplified from the pAG25 plasmid (Goldstein & McCusker, 1999) using the following primers: forward 5'-GCATCAGCAGTTTTATTAATTTTTTCTTATTGCGTGACCGCAACGGATCCCCGGGGTTAAT-3' and reverse 5'-TACTATGGTTATGCTCTGCTTCCCTATTTTCTTTGCGTTTGAATTCGAGCTCGTTTAAAC-3'. All transformations were performed using lithium acetate, as described previously (Gietz & Schiestl, 2007). Colony PCR confirmation of the knockout cassette integration was performed using the following primers: forward 5'- TCCAAAGCAACCATAGGTC-3' and reverse 5'-ATGTACGGGCGACAGTCACATCAT-3'. Serial dilution growth assays for resistance phenotyping of the knockout were performed on solid MCHM YPD media as above and compared to growth of the BY4742 knockout collection knockout of *PDR3* as well (Brachmann et al., 1998; Giaever et al., 2002).

Results

In-Lab evolution of YJM789 in MCHM:

Previous studies of MCHM's effects on yeast cells used genetic knockouts,

transcriptomics, metabolomics, and biochemical assays to determine that MCHM resistance is

dependent on a complex interaction of many cellular and genetic networks (Ayers et al., 2020;

Pupo, Ayers, et al., 2019; Pupo, Ku, et al., 2019). To expand upon this knowledge, resistant

strains were developed through In-Lab evolution (ILE) experiments to help elucidate genetic

changes that might produce a new resistant genotype. There is considerable genetic variation

across different yeast strains that may affect resistance to chemicals. Previously, a mapping

population between S96 (an S288c laboratory strain) and YJM789 (a clinical isolate) uncovered

standing genetic variation that is linked to MCHM response (Pupo, Ayers, et al., 2019). S96 is

more MCHM resistant than YJM789, and we used these strains to select for increased MCHM

resistance. Four replicates of each strain were evolved in rich liquid YPD media with and

without 700ppm MCHM (Figure 1A). The lineages of the original eight samples (four control,

four treated) were passaged every two days for six passages. Control and treated lineages were

plated after the sixth passage, and two single colonies were isolated from each lineage to

produce eight total control strains and eight total treated strains. After six passages, the MCHM

treated strains grew faster than the earlier passaged. However, this was only seen in the

YJM789 cultures, and increased MCHM resistance in S96 strains did not develop after multiple

attempts (not shown). S96 was already more MCHM resistant than YJM789 and further

resistance could not be achieved in the ILEs. We continued with YJM789 strains. These YJM789

strains were named S1-16, with S1-8 being the control strains (and S1 and S2 being from the

same original lineage), and S9-16 being the treated strains (and S9 and S10 being from the

same original lineage). The 16 strains were tested for their MCHM resistance at 1000ppm on

126

YPD media. The eight control strains showed wildtype levels of resistance, while all eight

treated strains showed increased resistance as compared to wildtype (Figure 1B). There was

plate to plate variation in MCHM at 1000ppm, likely due to nearing the solubility limit of the

chemical (Dietrich et al., 2015; Phetxumphou et al., 2016; Sain et al., 2015). The wildtype strain

was plated on each plate to control for plate to plate variation (Figure 1B).

Principal component analysis of mutations:

After successfully producing resistant strains of evolved YJM789 yeast, the 16 samples

were sequenced at approximately 25x coverage using Illumina paired-end reads. The reads of

all 16 strains were mapped to the S288c reference genome and variants were called using GATK

HaplotypeCaller. After removing the variants shared between all 16 strains, considered to be

the existing variants between the reference genome and the YJM789 genetic background,

variants were clustered on three principal components (Figure 2A-C). To analyze whether there

may be a general pattern of evolved resistance in the remaining variants, a PCA analysis was

performed. The strains did not cluster according to resistance on any of these components. As

an aside, they also did not cluster by original lineage, so none of the pairs (individual isolates

from the same population), such as S1 and S2 or S9 and S10, clustered. This indicates the

colonies isolated from the same cultures were genetically distinct. A large number of remaining

background variants, approximately 6900 (Table 1), likely prevented clustering by resistance

phenotype because there are too many variants in the analysis that are non-causative for

resistance. A few high effect loci would be masked by large numbers of non-causative variants.

Similarly, resistance spread among many lower effect loci spread throughout the genome

would also be masked if resistance in each strain was due to different combinations of these

small effects. In total, looking at the sum of variants produced by the In-Lab evolution

conditions did not produce clear patterns of resistance. The mutation patterns of the MCHM

and control strains showed there was no significant difference in the number of variants in

control vs. MCHM ILEs (p = 0.15) and all 16 strains had between 2326 and 2681 variants per

strain (Table 1). The total number of mutations does not seem to be from any effect of the

chemical.

Table 1: Summary of mutated genes in ILE strains. This table contains the number of overall
mutations and mutated genes in each ILE strain. Ty retrotransposon genes, such as gag/pol
domains, were excluded from this analysis.

Strain	Total mutations (6874 total)	MCHM specific mutations only (1171 total)	Mutations in coding regions – Excluding Ty	Genes with mutations (107)	Nonsynonymous mutations only	Genes with Nonsynonymous mutations (75)
S1	2550					
S2	2543					
S3	2681					
S4	2516					
S5	2436					
S6	2657					
S7	2521					
S8	2471					
S9	2579	212	54	25	33	16
S10	2496	187	63	28	31	18
S11	2493	180	43	24	20	16
S12	2415	187	53	25	28	16
S13	2512	189	59	33	35	23
S14	2475	170	64	33	42	20

| S15 | 2572 | 180 | 53 | 24 | 34 | 18 |
| S16 | 2326 | 170 | 45 | 25 | 25 | 15 |

Filtering MCHM evolved variants to those involved in coding sequence changes:

To find patterns of mutations that contributed to MCHM resistance, we considered variants shared by both the control ILE and the MCHM resistant strains were likely to be passenger mutations and not causative of the MCHM resistance phenotype. These mutations in the control ILE strains were filtered from the variants in the eight resistant ILEs. This left 1172 variants between the eight treated ILE strains (Table 1). Many of these variants were in non-coding regions of the genome, and while they could contribute to resistance to MCHM through changes in expression levels of genes, mutations in intergenic regions were filtered out, leaving 485 variants in genes between the eight strains (Table 1). The final filter used to narrow the focus to those that may be causative of resistance was to remove variants in Ty retrotransposon Gag/Pol proteins. The YJM789 strain does not contain the same Ty complement as the S288c reference strain (Wei et al., 2007). Therefore, the detection of these variants is likely due to the existing variation between the parent strain of the ILEs and the reference. The reason they remained after using earlier filters that removed all variants shared between the 16 ILEs, and a filter containing known SNPs in YJM789, is due to issues of 25x coverage with short reads not able to detect repetitive regions and large deletions. These final filters were supported by analysis of variant density across the genome (Supplemental Figure 1). This analysis plotted the location of variants for all 16 strains based on chromosomal location. With many of the 6800 variants being shared amongst multiple strains, sometimes up

129

to 15 strains, approximately 40000 variants were plotted initially (Supplemental Figure 1A).

When filtering out variants shared by 10 or more strains, 4900 unique variants remained

instead of 6800, resulting in plotting 15000 variants total across the 16 strains instead of 40000

(Supplemental Figure 1B). The differences in these plots include a region on chromosome VIII

around 90kb that corresponds to a Ty region in S288c that is missing in YJM789. These parental

variants were picked up in most strains, but not all. Filtering down to those variants shared by

less than 10 strains minimized the effect of missed parental variants on the analysis. The

removal of all variants existing in control strains, leaving the 1172 variants analyzed for coding

effects in MCHM strains, removed these variants. After removing Gag/Pol Ty retrotransposon

detected variants, there were between 45 and 64 variants per strain in the eight strains,

contained in 107 total genes (Table 1 and Supplemental Table 2). Finally, limiting our analysis to

nonsynonymous coding changes left between 20 and 42 variants per strain, contained in only

75 genes (Table 1 and Supplemental Table 3).

In order to analyze the intersection of these 75 genes with nonsynonymous changes

between strains, UpSetR software was utilized (Lex et al., 2014). Each strain contained

nonsynonymous variants in three to seven genes that were exclusive to that individual strain

(Figure 3A). This accounted for 42 of the genes containing variants in the MCHM evolved

strains, so more than half of genes with variants were found to have variants in multiple strains.

An additional 18 genes contain variants in two strains. The remaining 15 genes contained

variants within at least three strains, but only one gene was found to have a nonsynonymous

variant in all eight strains (Figure 3A). This gene was *PDR3* (Figure 3A), a transcription factor

that controls the expression of ABC transporters involved in the pleiotropic drug response

130

which exports chemicals from the cell that are involved in general stress (Delaveau et al., 1994).

When the intersection of all 107 genes containing nonsynonymous or synonymous variants was

analyzed, *PDR3* remained the only gene with a variant in all eight MCHM ILE strains

(Supplemental Figure 2, Table 2).

Table 2: Summary of genes containing unique variants in at least four MCHM ILE strains. This table shows a list of genes that appear in at least four MCHM ILE strains with an MCHM-unique variant. Uncharacterized ORFs have no annotated function in yeast but may produce functional proteins. Dubious ORFs are unlikely to produce functional proteins. The last two columns show how many MCHM and control ILE strains contain any variant at all in the gene, including variants that appear in both MCHM and control strains. These columns are included to show to what extent this gene seems to accumulate variants overall. Total number of variants is in parentheses in the last two columns, as some strains contain multiple variants in that gene.

Gene Name	MCHM ILEs Containing Unique Variant(s) in this Gene	ORF Character-ization	Chr	Start	End	MCHM ILEs - any Variant [genes (variants)]	Control ILEs - any Variant [genes (variants)]
AGP3	4	Verified	chrVI	17004	18680	4 (13)	3 (10)
COS4	5	Verified	chrVI	6426	7565	7 (36)	6 (22)
FLO1	4	Verified	chrI	203403	208016	8 (187)	8 (280)
HPF1	4	Verified	chrXV	28703	31606	8 (96)	8 (111)
HXT13	4	Verified	chrV	21537	23231	8 (61)	8 (60)
HXT17	5	Verified	chrXIV	772657	774351	6 (16)	4 (5)
HXT9	5	Verified	chrX	19497	21200	7 (13)	5 (19)
PAU24	4	Verified	chrII	809057	809419	7 (24)	7 (10)
PDR3	8	Verified	chrII	217470	220400	8 (8)	0 (0)
SAM3	4	Verified	chrXVI	22938	24701	7 (20)	8 (29)
TPO2	5	Verified	chrVII	763762	765606	5 (6)	0 (0)
YAL069W	7	Dubious	chrI	335	649	8 (91)	8 (97)

						8 (36)	8 (26)
YCR108C	6	Uncharacterized	chrIII	315997	316188		
YEL077C	4	Uncharacterized	chrV	264	4097	8 (44)	8 (63)
YEL077W-A	4	Dubious	chrV	630	1112	5 (16)	3 (7)
YER188C-A	4	Uncharacterized	chrV	569608	569907	8 (68)	8 (56)
YNL018C	4	Uncharacterized	chrXIV	599936	601774	8 (67)	8 (64)
YNL337W	4	Dubious	chrXIV	7165	7419	7 (18)	6 (12)

To determine if mutations in any cellular process might be important for evolved MCHM resistance, the 107 genes containing at least one variant in at least one MCHM ILE strain, including synonymous variants, were analyzed via GO Term analysis (Figure 3B). Most GO terms pointed to mutations in plasma membrane (FDR = 0.0096%) or cell wall proteins (FDR = 5.85 x 10^{-5}%), including those important for sugar transport or flocculation (Figure 3B). In agreement with previous genomic work with MCHM, the vacuole (FDR = 0.21%) and response to stress (FDR = 0.19%) also contained mutations that may impact resistance (Ayers et al., 2020).

To find only mutations that directly contributed to MCHM resistance, we focused on genes that contained mutations in multiple isolates. There were 18 genes with at least one synonymous or nonsynonymous variant not found in control strains within their coding regions in four or more MCHM ILEs (Table 2). However, many of these genes also seem to accumulate a high number of mutations in the rich liquid media culture environment. This is exemplified by *FLO1*, where all 16 strains contained variants, 467 in total. Some genes in the list accumulate more variants, possibly through their position near the ends of chromosomes where mutation rates are known

to be higher (Agier & Fischer, 2012; Ivanova et al., 2020; Lang & Murray, 2011) and/or reduced

selective pressure on their function in laboratory growth media. Only two genes showed no

mutations in any of the eight control ILE strains while also showing mutations in at least half of

the MCHM ILEs, *PDR3,* and *TPO2*.

The list included seven uncharacterized or dubious ORFs, mostly found at the ends of

their respective chromosomes (Table 2). The unknown or nonexistent function of these genes

made them unlikely candidates for adaptive resistance mechanisms. The remaining genes'

functions varied, but most were membrane or cell wall proteins (*FLO1, HPF1, PAU24*), often

involved in sugar or nutrient transport (*AGP3, HXT9, HXT13, HXT17, TPO2*). These cell wall

genes, as well as stress response genes such as *PDR3,* may be situated as targets of adaptive

changes for MCHM resistance due to their changes that may decrease the internal cellular

concentration of the chemical.

Prediction analysis for the tolerance of protein-coding changes from variants in MCHM ILEs:

Because there were thousands of variants per strain, including dozens of variants in

coding regions of each MCHM ILE, all mutants unique to the MCHM resistant strains were

scored as deleterious or well tolerated for protein function. SIFT analysis predicts whether

amino acid substitutions from single nucleotide polymorphisms will be tolerable for protein

function using multiple sequence alignment of homologous protein sequences (Vaser et al.,

2016). The variants in our dataset included both SNPs and indels, but indel variants could not

be scored because SIFT exclusively analyzes SNPs. We also analyzed our genome for CNVs but

none were detected in the evolved strains. Most of the variants were predicted to be tolerated

133

(Table 3 and Supplemental File 1). There are two reasons this may be the case. First, many of these variants are synonymous with coding changes, so SIFT analysis will align proteins and predict no deleterious effect based on conservation. That does not eliminate the possibility that the proteins' expression levels change due to codon bias changes (Bennetzen & Hall, 1982; Letzring et al., 2010; Presnyak et al., 2015), but these synonymous changes are less likely to be the major drivers of resistance change than protein sequence changes. The second reason most of these variants would be tolerated is selective pressures to keep conserved functions of proteins with nonsynonymous changes. The nonsynonymous changes would be more likely to replace similar amino acids that do not disrupt protein function, or in regions of the protein less likely to disrupt function. Variants in this dataset that were predicted to be tolerated were deemed less likely to be causative of the evolved resistance to MCHM.

The deleterious and NA scored variants all include variants that change the protein sequences. A variant given a score of NA, or not applicable, was usually an indel variant that could not be scored, but occasionally was a SNP variant that did not appear in the SIFT database. Many of the deleterious and NA scored variants appear in proteins that are uncharacterized or dubious ORFs (Supplemental File 1). The variants in *PDR3* in all eight MCHM ILEs received a score of NA. The only pattern of genes with these scores amongst the ILEs was the *PDR3* mutations. Other deleterious and NA scored mutations may provide unique effects on MCHM resistance in individual strains but were not a consistent pathway to resistance in different strains.

Table 3: Summary of predicted tolerance for coding variants found only in MCHM ILE strains. The eight MCHM ILE strains are represented from left to right in columns. Each strain is summarized for whether its coding variants are predicted to be tolerated, deleterious, or NA

based on SIFT4G analysis. NA scores mean a classification as either tolerated or deleterious was not possible to determine, or not applicable (NA). SIFT4G scores approaching 0 are expected to alter protein function to be deleterious based on amino acid substitutions. The NA results are usually due to the inability of SIFT to predict the effects of indels or occasionally SNPs that do not appear in the SIFT database. Coding variants in this table exclude those in Ty regions.

MCHM ILE Strain	S9	S10	S11	S12	S13	S14	S15	S16
Total Variants	54	63	43	53	59	64	53	45
Tolerated	38	41	30	35	42	30	31	23
Deleterious	7	2	5	2	8	12	6	4
NA	9	20	8	16	9	22	16	18

Comparison of MCHM evolved genes with genes appearing in other genomic datasets from MCHM studies:

Previous work has been done to produce genomic datasets that implicate genes involved in MCHM resistance mechanisms, so it hypothesized that some variants may target these same genes to adapt to MCHM treatment. The list of 107 genes containing at least one variant in an MCHM ILE strain was compared to a genetic screen of the BY4742 knockout collection and transcriptomic dataset (Figure 4) (Ayers et al., 2020). There was little overlap with the genetic screen (Figure 4A). This is not surprising, as the knockout screen selected for mutants with reduced MCHM resistance, revealing genes that are required for resistance. Therefore, in the 329 genes found from the previous knockout screen, any variants in evolved strains would have to avoid reducing the function of the proteins produced by those ORFs, as their existing function is known to be required for MCHM resistance. Adaptive changes to these

genes would most likely need to improve their current function in MCHM responses, which is less likely than a mutation neutral or deleterious to protein function.

Yeast were treated with MCHM and the changes in gene expression were quantitated. The transcriptomic dataset included a list of 592 upregulated genes and a list of 576 downregulated genes. There was more overlap with these lists and the 107 ILE genes, including 14 upregulated (Figure 4B) and eight downregulated genes (Figure 4C). The overlapping genes in these cases include several of the genes with variants in many ILE strains. Some genes of special note are the hexose transporters and carbon metabolism genes *HXT11*, *MAL13*, and *HXT17* from the upregulated list, and *HXT7* from the downregulated list. *HXT17* contains variants unique to MCHM ILEs in five strains and did not accumulate more mutations in some of the other ORFs containing dozens of mutations. Another gene involved in glucose metabolism with the hexose transporters above is *DOG1*, which encodes a deoxyglucose phosphatase involved in resistance to 2-deoxyglucose (Randez-Gil et al., 1995; Sanz et al., 1994; Soncini et al., 2020). While *DOG1* is responsible for alleviating stress from toxic glucose analogs in yeast, *GEX2* encodes a glutathione transporter important for oxidative stress resistance, a known source of stress from MCHM treatment (Ayers et al., 2020; Dhaoui et al., 2011). Carbon metabolism and stress-responsive genes are not the only types of genes in the overlapping datasets. Several cell wall proteins appear as well, including *PIR3, TIR1, SCW10,* and *FLO1*, each with varying functions from cell wall stability (*PIR3*) to flocculation (*FLO1*) (Cappellaro et al., 1998; Doolin et al., 2001; Hodgson et al., 1985; Kitagaki et al., 1997; Kowalski et al., 1995; Rossouw et al., 2015; Teunissen & Steensma, 1995; Toh-E et al., 1993; Yun et al., 1997). These datasets continue to point to the cell wall and sugar metabolism as functions that are important

136

ways to adapt resistance to MCHM. However, most of these genes also seem to accumulate mutations in control strains as well (Supplemental File 2). This could be from the passages themselves. The control yeast were also grown to saturation before being diluted into fresh media. Any mutation that decreases the lag before growth would be advantageous in these conditions, and these variants would begin to dominate the culture. These variants may have occurred in the variant dataset only by chance or from other selective pressures and not due to functional changes that contribute to MCHM resistance. Mutations in *PDR3* and *TPO2* were detected in many MCHM ILEs, contain variants in no control ILEs, and are differentially regulated at the transcriptome level in MCHM treated cells (Supplemental File 2 and Figure 4). *TPO2* belongs to a family of transporter proteins that are part of the multidrug resistance pathway. The protein localizes to the plasma membrane and has been characterized as an exporter of polyamines (Tomitori et al., 2001), which could be related to the amino acid biosynthesis pathways implicated in MCHM resistance in previous work (Ayers et al., 2020). However, the mutations in *TPO2* are all synonymous (Supplemental File 1). The codon changes could potentially affect protein levels if they affected translation rates.

Analysis of evolved alleles in *PDR3*:

The *PDR3* variants in the MCHM ILEs showed a pattern of evolution that pointed to a reproducible pathway to resistance. We decided to look more closely at the individual mutations. Previous research has been done with *PDR3* mutagenesis to produce gain-of-function alleles that improve resistance to different chemicals (Nourani et al., 1997). For instance, amino acid mutations in the region from approximately residues 220-280 created alleles that increased the expression of multiple ABC transporters that pump chemicals out of

cells, specifically *SNQ2* and *PDR5* (Nourani et al., 1997). The mutations in *PDR3* in the S12, S13, and S15 ILEs mutated single amino acids at residues 288, 209, and 229 respectively (Figure 5A). It is possible that these mutations are creating gain-of-function alleles that increase the expression of ABC transporters, thereby conveying resistance to MCHM. The gene *SNQ2* is required for resistance to MCHM (Ayers et al., 2020), but it has yet to be shown if overexpression would be sufficient to produce resistance.

The remaining five mutations resulted in changes to the C-terminal portion of the protein, where an activation domain homologous to Gal4-like transcriptional activators is located (Delaveau et al., 1994). The S9 mutation was a M842L single amino acid change (Figure 5A). One study produced six different gain-of-function mutations in this region with single nucleotide changes that increased expression of *PDR3*, *PDR5*, and *SNQ2* and conferred resistance to several chemicals (Simonics et al., 2000). None of the mutations were the same as found in the ILE strains, but they do implicate this region of the protein for possible gain-of-function resistance mutations. The S16 mutation was an insertion and frameshift occurring at the amino acid 972 that altered and extended the remaining four amino acids into an extra nine amino acids before the new stop codon (Figure 5A).

The S10 and S14 mutations truncated the protein by 415 and 406 amino acids, respectively (Figure 5A). The S11 mutation also truncated the protein, but by 82 amino acids (Figure 5A). One hypothesis is that deletions of this much of the protein sequence would produce nonfunctional products, effectively acting as a knockout. Alternatively, these truncations could still produce proteins with some function, considering the DNA binding domain is at the N-terminus, like other Gal4-like transcription factors (Delaveau et al., 1994).

138

Furthermore, the motifs implicated in the homo- and heterodimer interactions of Pdr3 and Pdr1 are also in the N-terminal 400 amino acids. With dimerization and DNA interacting regions of the protein products intact, the truncated alleles of S10, S11, and S14 could be functional proteins.

To test the more fundamental hypothesis that a knockout-like truncation may confer resistance to MCHM, we knocked the *PDR3* gene out of the wildtype parent of the ILEs, YJM789. The *pdr3Δ* strain produced by knocking out the gene with a NAT[R] marker showed similar resistance as the S11 resistant MCHM strain (Figure 5B). S11 corresponds to the strain containing the shortest truncation of the Pdr3 protein, only 82 amino acids. If the mutations to *PDR3* in the evolved strains are mimicking a knockout by making Pdr3 nonfunctional, that would be sufficient to produce the resistant phenotype. The *pdr3Δ* in the BY4742 strain background was also tested for MCHM resistance to see if the relationship between this genotype and phenotype was specific to the YJM789 background. At 2 days of growth, BY4742 *pdr3Δ* grew faster than the BY4742 wildtype strain at 550ppm MCHM (Figure 5B). However, by three days, the wildtype strain reached similar growth to the knockout (not shown), and the knockout did not show a resistant phenotype above 550ppm MCHM. The knockout of this gene produced some increased resistance, indicating its effects may be important outside of the YJM789 genetic background. The effect in YJM789 appears greater, producing a resistant phenotype at concentrations much higher than the wildtype strain's tolerance.

Discussion

The In-Lab evolution of the YJM789 strain of *S. cerevisiae* produced thousands of mutations in each strain. The mutations were similar in number from strain to strain, including both control media and MCHM-treated conditions. MCHM is known to induce DNA damage (Ayers et al., 2020), which could possibly be a mechanism for mutation, but as the control strains accumulated similar numbers of mutations, it likely played little role other than a selective pressure on mutations occurring normally in the growth conditions. There was no clear pattern of mutations resulting in MCHM resistance across strains, though the sheer number of mutations could mask any pattern with background variants. When filtered for mutations unique to MCHM ILEs that were only found in coding sequences, patterns such as cell wall and responses to stress did begin to emerge. These patterns are consistent with previous knowledge of MCHM effects on cells, including oxidative stress activation (Ayers et al., 2020). MCHM is also known to act as a hydrotrope (Pupo, Ayers, et al., 2019), so variants that may stabilize the plasma membrane or cell wall could be likely resistance mechanisms. Despite consistency with previous work, these patterns also have issues that reduced confidence that they were causative variants. In particular, the variants included synonymous mutations and genes that were accumulating mutations in control strains as well. Synonymous mutations do not change protein sequence and are unlikely to significantly affect the function of that protein unless altering expression levels. The plentiful mutations in the same genes in control strains may indicate that the genes are hotspots for mutation accumulation, reducing the likelihood any particular mutation was found due to selection for a competitive advantage in MCHM. However, based on the QTL study previously done in MCHM, associative loci that failed to reach

140

LOD score significance were spread relatively evenly throughout the genome (Pupo, Ayers, et al., 2019). This indicates that the presence of small effect loci contributing to a combined pleiotropic effect on resistance phenotype is likely. These variants may be contributing small phenotypic effects even if they are not sufficient to produce resistance on their own.

There was only one gene that seemed to mutate exclusively in MCHM strains: *PDR3*. The role of this gene in activating the stress response to chemicals, combined with it being the only gene to mutate in every MCHM ILE but no control ILE, made it the most likely candidate for reproducible evolved resistance to MCHM. Follow-up genetic work by knocking the gene out of the original parent strain revealed that inactivating the protein was sufficient to induce resistance at the level of the evolved strains. This gene is the first to be identified in yeast as a target to induce resistance to MCHM. Previous work, such as the genetic screen, focused only on finding genes required for tolerance. The screen was not designed to detect an increased resistance phenotype in any of the mutants tested. The model for this resistance involves the importance of Pdr3 in controlling the activation of multiple ABC transporters that pump stress-inducing chemicals out of cells. Pdr3 and its paralog Pdr1 form homo- and heterodimers and then bind to transcriptional response regions termed PDREs, where they can inhibit or activate transcription of genes such as *PDR5*, *SNQ2*, *YOR1*, *PDR15*, *PDR10*, and other transporters (Delaveau et al., 1994; Jungwirth & Kuchler, 2006; Katzmann et al., 1996; Kolaczkowski et al., 1998; Mahé et al., 1996; Mamnun et al., 2002; Servos et al., 1993) (Figure 5C). Changing the levels of functional Pdr3 or altering the ability of Pdr3 to dimerize could affect expression of individual or multiple ABC transporters and change the amounts of ones more important for MCHM resistance. We tested via knockout in the YJM789 parent strain if inactivating mutations

could be sufficient to produce MCHM resistance. (Ayers et al., 2020; Pupo et al., 2019). The mutations in *PDR3* could also be activating mutations that make Pdr3 increase transcription of all or some transporters, such as *SNQ2* or *PDR5*.

The *SNQ2* gene is required for MCHM resistance, so it is a likely candidate for this increased expression (Ayers et al., 2020). There is a 637 amino acid N-terminal truncation in the Snq2 protein in YJM789 as compared to the reference strain, or the resistant QTL parent strain S96. The YJM789 allele of *PDR5* is also divergent from the reference and S96 yeast strains, with a 5.3% amino acid difference compared to the reference strain (Guan et al., 2010; Wei et al., 2007). The YJM789 allele of *PDR5* has been shown to alter the strain's resistance to antifungals (Guan et al., 2010), increasing or decreasing resistance depending on the chemical. *PDR5* is not required for resistance to MCHM like *SNQ2* according to previous work, but this work was done in the BY4742 strain, which has the same allele as the reference strain (Ayers et al., 2020). It is possible that the YJM789 *PDR5* allele can provide resistance to MCHM if the expression is increased. *PDR5* is significantly upregulated by gain-of-function mutations in *PDR3* in the same regions as the MCHM ILE variants (Nourani et al., 1997; Simonics et al., 2000), but changes in expression of transporters in strains with mutated Pdr3 have not been measured. The divergent allele complements of ABC transporters between YJM789 and S96 may also point to a possible explanation for the lack of evolved resistance we observed in the S96 strain. While changes to the function of Pdr3 may be a good target for evolved resistance in YJM789, the alleles of ABC transporters in S96 may be less amenable to the *PDR3* genotype-phenotype relationship. This may also explain why the BY4742 strain seemed to receive a less extreme increase in resistance from its *pdr3Δ* deletion, as BY4742 and S96 are closely related strains.

This study was able to produce MCHM resistant evolved strains of YJM789. The patterns from the evolved strains point most clearly to a model involving resistance via mutation of the gene encoding the transcription factor Pdr3. Background mutations in the ILEs, such as less consistently mutated genes or mutated *cis*-regulatory elements, may also be contributing to resistance in each strain. In this case, resistance to MCHM might involve various pathways that allow for a mix of small effect changes throughout the genome. This would be consistent with work that has shown hundreds of required genes, hundreds of differentially regulated genes, and a QTL showing widespread low effect loci throughout the genome (Ayers et al., 2020; Pupo, Ayers, et al., 2019). Analysis of these types of small combinations of alleles will be difficult. Therefore, future work should explore the hypotheses from the model that the *PDR3* alleles in the ILEs are sufficient to provide MCHM resistance by altering the expression of ABC transporters. Cloning each allele into the YJM789 *PDR3* knockout strain will create strains that can test directly for MCHM resistance and test the altered expression of the ABC transporters in the same genetic background, isolated from the background mutations of the ILE strains. Quantitative PCR for all relevant genes, especially *SNQ2* and *PDR5,* and western blots of the protein products will answer whether these alleles change the expression of these genes. The results of these experiments should reveal whether MCHM resistance can be evolved reproducibly through one large effect gene, or if it is mainly acquired through many small changes. The cloned alleles of *PDR3* may also be relevant for studies of the transcription factor's role in other stress resistance mechanisms. Other future directions for this dataset could explore hypotheses about *TPO2* codon usage bias to determine if the evolved synonymous allele changes contribute significantly to MCHM resistance. The dataset may also

be informative of patterns of evolution in stress treatments when combined with ILE analysis in other environments. Adaptive changes to MCHM provide insight into this specific novel chemical stressor of yeast, but also inform studies about stress responses and evolution in the field.

Acknowledgments:

Author Contributions:

MCA designed and carried out yeast experiments, all analysis, and wrote the paper. JEGG designed the study and supervised the project.

Competing Financial Interests

The authors declare no competing interests.

References

Adamo, G. M., Brocca, S., Passolunghi, S., Salvato, B., & Lotti, M. (2012). Laboratory evolution of copper tolerant yeast strains. *Microbial Cell Factories*, *11*(1), 1. https://doi.org/10.1186/1475-2859-11-1

Agier, N., & Fischer, G. (2012). The Mutational Profile of the Yeast Genome Is Shaped by Replication. *Molecular Biology and Evolution*, *29*(3), 905–913. https://doi.org/10.1093/molbev/msr280

Ayers, M. C., Sherman, Z. N., & Gallagher, J. E. G. (2020). Oxidative Stress Responses and Nutrient Starvation in MCHM Treated Saccharomyces cerevisiae. *BioRxiv*, 2020.08.17.253799. https://doi.org/10.1101/2020.08.17.253799

Balzi, E., Wang, M., Leterme, S., Van Dyck, L., & Goffeau, A. (1994). PDR5, a novel yeast multidrug resistance conferring transporter controlled by the transcription regulator PDR1. *The Journal of Biological Chemistry*, *269*(3), 2206–2214.

Bennetzen, J. L., & Hall, B. D. (1982). Codon selection in yeast. *The Journal of Biological Chemistry*, *257*(6), 3026–3031.

Bissinger, P. H., & Kuchler, K. (1994). Molecular cloning and expression of the Saccharomyces cerevisiae STS1 gene product. A yeast ABC transporter conferring mycotoxin resistance. *The Journal of Biological Chemistry*, *269*(6), 4180–4186.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains

and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, *14*(2), 115–132.

https://doi.org/10.1002/(SICI)1097-0061(19980130)14:2<115::AID-YEA204>3.0.CO;2-2

Cappellaro, C., Mrsa, V., & Tanner, W. (1998). New Potential Cell Wall Glucanases of

Saccharomyces cerevisiae and Their Involvement in Mating. *Journal of Bacteriology*, *180*(19),

5030–5037.

Carlson, M., & Maintainer, B. P. (2015). *TxDb.Scerevisiae.UCSC.sacCer3.sgdGene: Annotation*

package for TxDb object(s).

Caspeta, L., & Nielsen, J. (2015). Thermotolerant Yeast Strains Adapted by Laboratory Evolution

Show Trade-Off at Ancestral Temperatures and Preadaptation to Other Stresses. *MBio*, *6*(4).

https://doi.org/10.1128/mBio.00431-15

Delaveau, T., Delahodde, A., Carvajal, E., Subik, J., & Jacq, C. (1994). PDR3, a new yeast

regulatory gene, is homologous toPDR1 and controls the multidrug resistance phenomenon.

Molecular and General Genetics MGG, *244*(5), 501–511. https://doi.org/10.1007/BF00583901

Dhaoui, M., Auchère, F., Blaiseau, P.-L., Lesuisse, E., Landoulsi, A., Camadro, J.-M., Haguenauer-

Tsapis, R., & Belgareh-Touzé, N. (2011). Gex1 is a yeast glutathione exchanger that interferes

with pH and redox homeostasis. *Molecular Biology of the Cell*, *22*(12), 2054–2067.

https://doi.org/10.1091/mbc.E10-11-0906

Dietrich, A. M., Thomas, A., Zhao, Y., Smiley, E., Shanaiah, N., Ahart, M., Charbonnet, K. A.,

DeYonker, N. J., Alexander, W. A., & Gallagher, D. L. (2015). Partitioning, Aqueous Solubility,

and Dipole Moment Data for cis- and trans-(4-Methylcyclohexyl)methanol, Principal

Contaminants of the West Virginia Chemical Spill. *Environmental Science & Technology Letters*, *2*(4), 123–127. https://doi.org/10.1021/acs.estlett.5b00061

Doolin, M.-T., Johnson, A. L., Johnston, L. H., & Butler, G. (2001). Overlapping and distinct roles of the duplicated yeast transcription factors Ace2p and Swi5p. *Molecular Microbiology*, *40*(2), 422–432. https://doi.org/10.1046/j.1365-2958.2001.02388.x

Engel, S. R., Dietrich, F. S., Fisk, D. G., Binkley, G., Balakrishnan, R., Costanzo, M. C., Dwight, S. S., Hitz, B. C., Karra, K., Nash, R. S., Weng, S., Wong, E. D., Lloyd, P., Skrzypek, M. S., Miyasato, S. R., Simison, M., & Cherry, J. M. (2013). The Reference Genome Sequence of Saccharomyces cerevisiae: Then and Now. *G3: Genes|Genomes|Genetics*, *4*(3), 389–398. https://doi.org/10.1534/g3.113.008995

Gel, B., & Serra, E. (2017). karyoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics*, *33*(19), 3088–3090. https://doi.org/10.1093/bioinformatics/btx346

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., … Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, *418*(6896), 387–391. https://doi.org/10.1038/nature00935

Gietz, R. D., & Schiestl, R. H. (2007). Quick and easy yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, *2*(1), 35–37. https://doi.org/10.1038/nprot.2007.14

Goldstein, A. L., & McCusker, J. H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast (Chichester, England)*, *15*(14), 1541–1553. https://doi.org/10.1002/(SICI)1097-0061(199910)15:14<1541::AID-YEA476>3.0.CO;2-K

Gorter de Vries, A. R., Voskamp, M. A., van Aalst, A. C. A., Kristensen, L. H., Jansen, L., van den Broek, M., Salazar, A. N., Brouwers, N., Abeel, T., Pronk, J. T., & Daran, J.-M. G. (2019). Laboratory Evolution of a Saccharomyces cerevisiae × S. eubayanus Hybrid Under Simulated Lager-Brewing Conditions. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.00242

Guan, W., Jiang, H., Guo, X., Mancera, E., Xu, L., Li, Y., Steinmetz, L., Li, Y., & Gu, Z. (2010). Antagonistic changes in sensitivity to antifungal drugs by mutations of an important ABC transporter gene in a fungal pathogen. *PLoS ONE*, *5*(6), 1–10. https://doi.org/10.1371/journal.pone.0011309

Han, A. A., Fabyanic, E. B., Miller, J. V., Prediger, M. S., Prince, N., Mouch, J. A., & Boyd, J. (2017). In vitro cytotoxicity assessment of a West Virginia chemical spill mixture involving 4-methylcyclohexanemethanol and propylene glycol phenyl ether. *Environmental Monitoring and Assessment*, *189*(4), 190. https://doi.org/10.1007/s10661-017-5895-5

Hodgson, J. A., Berry, D. R., & Johnston, J. R. (1985). Discrimination by Heat and Proteinase Treatments between Flocculent Phenotypes Conferred on Saccharomyces cerevisiae by the Genes FLO1 and FLO5. *Microbiology, 131*(12), 3219–3227. https://doi.org/10.1099/00221287-131-12-3219

Horzmann, K. A., de Perre, C., Lee, L. S., Whelton, A. J., & Freeman, J. L. (2017). Comparative analytical and toxicological assessment of methylcyclohexanemethanol (MCHM) mixtures

associated with the Elk River chemical spill. *Chemosphere*, *188*, 599–607.

https://doi.org/10.1016/j.chemosphere.2017.09.026

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009a). Systematic and integrative analysis of

large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57.

https://doi.org/10.1038/nprot.2008.211

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009b). Bioinformatics enrichment tools: Paths

toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1),

1–13. https://doi.org/10.1093/nar/gkn923

Ivanova, T., Maier, M., Missarova, A., Ziegler-Birling, C., Dam, M., Gomar-Alba, M., Carey, L. B.,

& Mendoza, M. (2020). Budding yeast complete DNA synthesis after chromosome segregation

begins. *Nature Communications*, *11*(1), 2267. https://doi.org/10.1038/s41467-020-16100-3

Jungwirth, H., & Kuchler, K. (2006). Yeast ABC transporters—A tale of sex, stress, drugs and

aging. *FEBS Letters*, *580*(4), 1131–1138. https://doi.org/10.1016/j.febslet.2005.12.050

Katzmann, D. J., Epping, E. A., & Moye-Rowley, W. S. (1999). Mutational disruption of plasma

membrane trafficking of Saccharomyces cerevisiae Yor1p, a homologue of mammalian

multidrug resistance protein. *Molecular and Cellular Biology*, *19*(4), 2998–3009.

https://doi.org/10.1128/mcb.19.4.2998

Katzmann, D. J., Hallstrom, T. C., Mahé, Y., & Moye-Rowley, W. S. (1996). Multiple Pdr1p/Pdr3p

binding sites are essential for normal expression of the ATP binding cassette transporter

protein-encoding gene PDR5. *The Journal of Biological Chemistry*, *271*(38), 23049–23054. https://doi.org/10.1074/jbc.271.38.23049

Kitagaki, H., Shimoi, H., & Itoh, K. (1997). Identification and Analysis of a Static Culture-Specific Cell Wall Protein, Tir1p/Srp1p in Saccharomyces Cerevisiae. *European Journal of Biochemistry*, *249*(1), 343–349. https://doi.org/10.1111/j.1432-1033.1997.t01-1-00343.x

Kolaczkowski, M., Kolaczowska, A., Luczynski, J., Witek, S., & Goffeau, A. (1998). In vivo characterization of the drug resistance profile of the major ABC transporters and other components of the yeast pleiotropic drug resistance network. *Microbial Drug Resistance (Larchmont, N.Y.)*, *4*(3), 143–158. https://doi.org/10.1089/mdr.1998.4.143

Kowalski, L. R. Z., Kondo, K., & Inouye, M. (1995). Cold-shock induction of a family of TIP1-related proteins associated with the membrane in Saccharomyces cerevisiae. *Molecular Microbiology*, *15*(2), 341–353. https://doi.org/10.1111/j.1365-2958.1995.tb02248.x

Kryazhimskiy, S., Rice, D. P., Jerison, E. R., & Desai, M. M. (2014). Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science*, *344*(6191), 1519–1522. https://doi.org/10.1126/science.1250939

Lan, J., Hu, M., Gao, C., Alshawabkeh, A., & Gu, A. Z. (2015). Toxicity assessment of 4-Methyl-1-cyclohexanemethanol and its metabolites in response to a recent chemical spill in west virginia, USA. *Environmental Science and Technology*, *49*(10), 6284–6293. https://doi.org/10.1021/acs.est.5b00371

Lang, G. I., & Murray, A. W. (2011). Mutation Rates across Budding Yeast Chromosome VI Are Correlated with Replication Timing. *Genome Biology and Evolution*, *3*, 799–811. https://doi.org/10.1093/gbe/evr054

Letzring, D. P., Dean, K. M., & Grayhack, E. J. (2010). Control of translation efficiency in yeast by codon–anticodon interactions. *RNA*, *16*(12), 2516–2528. https://doi.org/10.1261/rna.2411710

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., & Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics (Oxford, England)*, *27*(21), 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Luu, K., Blum, M., & Privé, F. (2020). *pcadapt: Fast Principal Component Analysis for Outlier Detection*. https://CRAN.R-project.org/package=pcadapt

151

Mahé, Y., Parle-McDermott, A., Nourani, A., Delahodde, A., Lamprecht, A., & Kuchler, K. (1996). The ATP-binding cassette multidrug transporter Snq2 of Saccharomyces cerevisiae: A novel target for the transcription factors Pdr1 and Pdr3. *Molecular Microbiology*, *20*(1), 109–117. https://doi.org/10.1111/j.1365-2958.1996.tb02493.x

Mamnun, Y. M., Pandjaitan, R., Mahé, Y., Delahodde, A., & Kuchler, K. (2002). The yeast zinc finger regulators Pdr1p and Pdr3p control pleiotropic drug resistance (PDR) as homo- and heterodimers in vivo. *Molecular Microbiology*, *46*(5), 1429–1440. https://doi.org/10.1046/j.1365-2958.2002.03262.x

Mangado, A., Morales, P., Gonzalez, R., & Tronchoni, J. (2018). Evolution of a Yeast With Industrial Background Under Winemaking Conditions Leads to Diploidization and Chromosomal Copy Number Variation. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.01816

McCusker, J. H., Clemons, K. V., Stevens, D. A., & Davis, R. W. (1994a). Genetic characterization of pathogenic Saccharomyces cerevisiae isolates. *Genetics*, *136*(4), 1261–1269. https://doi.org/10.1111/j.1567-1364.2008.00464.x

McCusker, J. H., Clemons, K. V., Stevens, D. A., & Davis, R. W. (1994b). Saccharomyces cerevisiae virulence phenotype as determined with CD-1 mice is associated with the ability to grow at 42??C and form pseudohyphae. *Infection and Immunity*, *62*(12), 5447–5455.

Nourani, A., Papajova, D., Delahodde, A., Jacq, C., & Subik, J. (1997). Clustered amino acid substitutions in the yeast transcription regulator Pdr3p increase pleiotropic drug resistance and

identify a new central regulatory domain. *Molecular and General Genetics MGG*, *256*(4), 397–405. https://doi.org/10.1007/s004380050583

Obenchain, V., Lawrence, M., Carey, V., Gogarten, S., Shannon, P., & Morgan, M. (2014). VariantAnnotation: A Bioconductor package for exploration and annotation of genetic variants. *Bioinformatics*, *30*(14), 2076–2078. https://doi.org/10.1093/bioinformatics/btu168

Pagès, H. (2019). *BSgenome: Software infrastructure for efficient representation of full genomes and their SNPs*.

Phetxumphou, K., Dietrich, A. M., Shanaiah, N., Smiley, E., & Gallagher, D. L. (2016). Subtleties of human exposure and response to chemical mixtures from spills. *Environmental Pollution*, *214*, 618–626. https://doi.org/10.1016/j.envpol.2016.04.056

Presnyak, V., Alhusaini, N., Chen, Y.-H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K. E., Graveley, B. R., & Coller, J. (2015). Codon Optimality Is a Major Determinant of mRNA Stability. *Cell*, *160*(6), 1111–1124. https://doi.org/10.1016/j.cell.2015.02.029

Pupo, A., Ayers, M. C., Sherman, Z. N., Vance, R. J., Cumming, J. R., & Gallagher, J. E. G. (2019). MCHM Acts as a Hydrotrope, Altering the Balance of Metals in Yeast. *Biological Trace Element Research*. https://doi.org/10.1007/s12011-019-01850-z

Pupo, A., Ku, K. M., & Gallagher, J. E. G. (2019). Effects of MCHM on yeast metabolism. *PLOS ONE*, *14*(10), e0223909. https://doi.org/10.1371/journal.pone.0223909

Randez-Gil, F., Blasco, A., Prieto, J. A., & Sanz, P. (1995). DOGR1 and DOGR2: Two genes from Saccharomyces cerevisiae that confer 2-deoxyglucose resistance when overexpressed. *Yeast*, *11*(13), 1233–1240. https://doi.org/10.1002/yea.320111303

Ravishankar, A., Pupo, A., & Gallagher, J. E. G. (2020). Resistance Mechanisms of Saccharomyces cerevisiae to Commercial Formulations of Glyphosate Involve DNA Damage Repair, the Cell Cycle, and the Cell Wall Structure. *G3: Genes, Genomes, Genetics*, *10*(6), 2043–2056. https://doi.org/10.1534/g3.120.401183

Rossouw, D., Bagheri, B., Setati, M. E., & Bauer, F. F. (2015). Co-Flocculation of Yeast Species, a New Mechanism to Govern Population Dynamics in Microbial Ecosystems. *PLoS ONE*, *10*(8). https://doi.org/10.1371/journal.pone.0136249

Sain, A. E., Dietrich, A. M., Smiley, E., & Gallagher, D. L. (2015). Assessing human exposure and odor detection during showering with crude 4-(methylcyclohexyl)methanol (MCHM) contaminated drinking water. *Science of The Total Environment*, *538*, 298–305. https://doi.org/10.1016/j.scitotenv.2015.08.050

Sanz, P., Randez-Gil, F., & Prieto, J. A. (1994). Molecular characterization of a gene that confers 2-deoxyglucose resistance in yeast. *Yeast*, *10*(9), 1195–1202. https://doi.org/10.1002/yea.320100907

Servos, J., Haase, E., & Brendel, M. (1993). Gene SNQ2 of Saccharomyces cerevislae, which confers resistance to 4-nitroquinoline-N-oxide and other chemicals, encodes a 169 kDa protein homologous to ATP-dependent permeases. *Molecular and General Genetics MGG*, *236*(2), 214–218. https://doi.org/10.1007/BF00277115

Simonics, T., Kozovska, Z., Michalkova-Papajova, D., Delahodde, A., Jacq, C., & Subik, J. (2000).

Isolation and molecular characterization of the carboxy-terminal pdr3 mutants in

Saccharomyces cerevisiae. *Current Genetics*, *38*(5), 248–255.

https://doi.org/10.1007/s002940000164

Soncini, S. R., Chandrashekarappa, D. G., Augustine, D. A., Callahan, K. P., O'Donnell, A. F., &

Schmidt, M. C. (2020). Spontaneous mutations that confer resistance to 2-deoxyglucose act

through Hxk2 and Snf1 pathways to regulate gene expression and HXT endocytosis. *PLOS*

Genetics, *16*(7), e1008484. https://doi.org/10.1371/journal.pgen.1008484

Tawfik, O. W., Papasian, C. J., Dixon, A. Y., & Potter, L. M. (1989). Saccharomyces cerevisiae

pneumonia in a patient with acquired immune deficiency syndrome. *J Clin Microbiol*, *27*(7),

1689–1691.

Team, T. B. D. (2014). *BSgenome.Scerevisiae.UCSC.sacCer3: Saccharomyces cerevisiae (Yeast)*

full genome (UCSC version sacCer3).

Teunissen, A. W. R. H., & Steensma, H. Y. (1995). The dominant flocculation genes of

Saccharomyces cerevisiae constitute a new subtelomeric gene family. *Yeast*, *11*(11), 1001–

1013. https://doi.org/10.1002/yea.320111102

Toh-E, A., Oguchi, T., Matsui, Y., Yasunaga, S., Nisogi, H., & Tanaka, K. (1993). Three yeast genes,

PIR1, PIR2 and PIR3, containing internal tandem repeats, are related to each other, and PIR1

and PIR2 are required for tolerance to heat shock. *Yeast*, *9*(5), 481–494.

https://doi.org/10.1002/yea.320090504

Tomitori, H., Kashiwagi, K., Asakawa, T., Kakinuma, Y., Michael, A. J., & Igarashi, K. (2001). Multiple polyamine transport systems on the vacuolar membrane in yeast. *Biochemical Journal*, *353*(Pt 3), 681–688.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ data to high confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Current Protocols in Bioinformatics*, *43*, 11.10.1-11.10.33. https://doi.org/10.1002/0471250953.bi1110s43

Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., & Ng, P. C. (2016). SIFT missense predictions for genomes. *Nature Protocols*, *11*(1), 1–9. https://doi.org/10.1038/nprot.2015.123

Wei, W., McCusker, J. H., Hyman, R. W., Jones, T., Ning, Y., Cao, Z., Gu, Z., Bruno, D., Miranda, M., Nguyen, M., Wilhelmy, J., Komp, C., Tamse, R., Wang, X., Jia, P., Luedi, P., Oefner, P. J., David, L., Dietrich, F. S., … Steinmetz, L. M. (2007). Genome sequencing and comparative analysis of Saccharomyces cerevisiae strain YJM789. *Proceedings of the National Academy of Sciences*, *104*(31), 12825–12830. https://doi.org/10.1073/pnas.0701291104

Weidhaas, J. L., Dietrich, A. M., DeYonker, N. J., Ryan Dupont, R., Foreman, W. T., Gallagher, D., Gallagher, J. E. G., Whelton, A. J., & Alexander, W. A. (2016). Enabling Science Support for Better Decision-Making when Responding to Chemical Spills. *Journal of Environment Quality*, *45*(5), 1490. https://doi.org/10.2134/jeq2016.03.0090

West Virginia Chemical Spill: Collective NTP Findings and Supporting Files. (n.d.). National

Toxicology Program U.S. Department of Health and Human Services. Retrieved March 8, 2017,

from https://ntp.niehs.nih.gov/results/areas/wvspill/collective.html

Wolfger, H., Mahé, Y., Parle-McDermott, A., Delahodde, A., & Kuchler, K. (1997). The yeast ATP

binding cassette (ABC) protein genes PDR10 and PDR15 are novel targets for the Pdr1 and Pdr3

transcriptional regulators. *FEBS Letters*, *418*(3), 269–274. https://doi.org/10.1016/s0014-

5793(97)01382-3

Wolfger, Hubert, Mamnun, Y. M., & Kuchler, K. (2004). The yeast Pdr15p ATP-binding cassette

(ABC) protein is a general stress response factor implicated in cellular detoxification. *The

Journal of Biological Chemistry*, *279*(12), 11593–11599.

https://doi.org/10.1074/jbc.M311282200

Yun, D.-J., Zhao, Y., Pardo, J. M., Narasimhan, M. L., Damsz, B., Lee, H., Abad, L. R., D'Urzo, M.

P., Hasegawa, P. M., & Bressan, R. A. (1997). Stress proteins on the yeast cell surface determine

resistance to osmotin, a plant antifungal protein. *Proceedings of the National Academy of

Sciences of the United States of America*, *94*(13), 7082–7087.

Zhou, N., Katz, M., Knecht, W., Compagno, C., & Piškur, J. (2018). Genome dynamics and

evolution in yeasts: A long-term yeast-bacteria competition experiment. *PLOS ONE*, *13*(4),

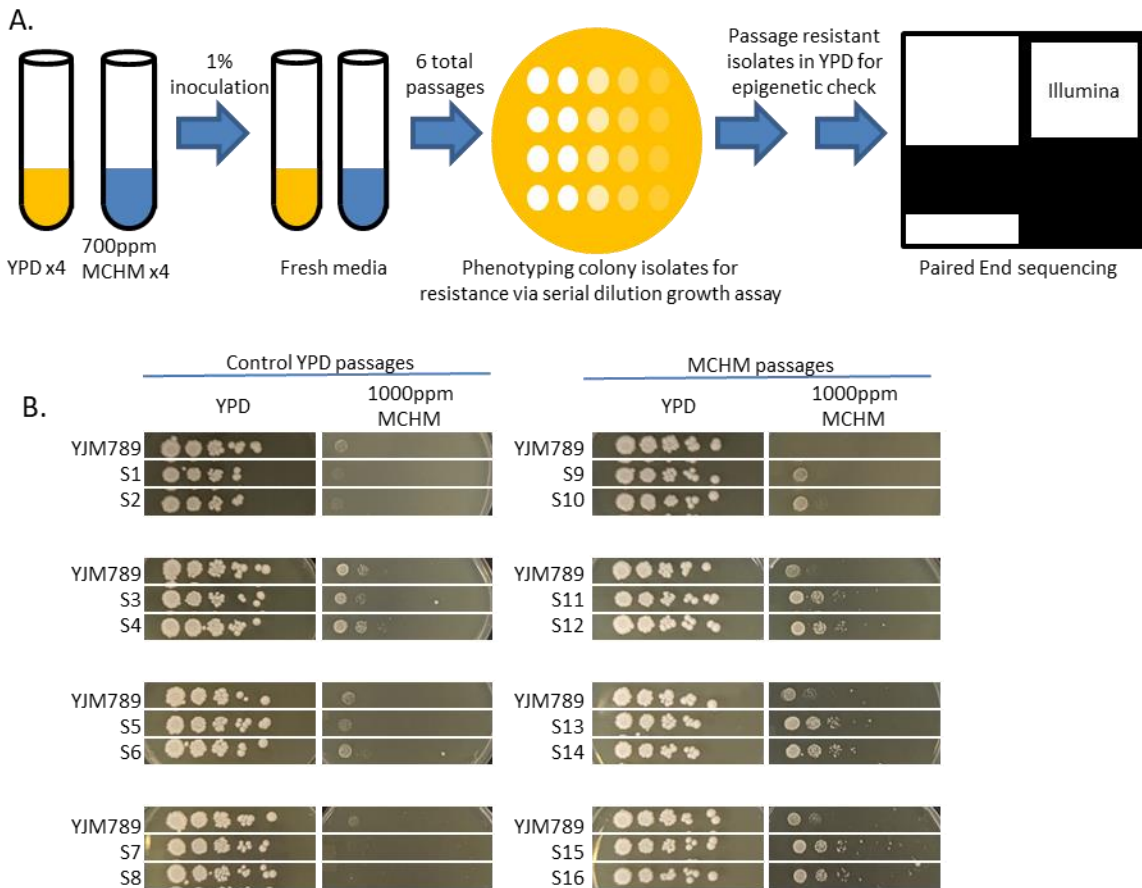e0194911. https://doi.org/10.1371/journal.pone.0194911

Figures:

Figure 1: In-Lab evolutions produce MCHM resistant yeast strains. A. Four biological replicates of YJM789 yeast were grown in liquid rich media (YPD) with and without 700ppm MCHM. Each replicate was grown for two days, then passaged by inoculating 1% of the culture into fresh media. Six total passages were done, testing growth by serial diluting the cultures onto solid media containing 1000ppm MCHM to determine that resistance was achieved. Resistant isolates were passaged on solid rich media plates without MCHM to check for loss of epigenetically inherited resistance mechanism. Two single colony isolates from each of the four control and four MCHM replicates were selected for analysis. Genomic DNA was extracted and sequenced with Illumina. **B.** Resistance of isolates from the In-Lab evolutions were examined by serial dilution assay on solid YPD with and without 1000ppm MCHM. The eight control strains are labeled S1-S8. Pairs of strains (S1 and S2, S3 and S4, etc.) are isolates from the same original biological replicate. Pairs were plated on the same plates with a control YJM789 parent strain for growth assays. The YJM789 parent appears on each plate to control for plate to plate variation in MCHM dosage due to volatility and the high concentration of MCHM approaching its limits. The eight MCHM evolved strains are labeled S9-S16, and pairs such as S9 and S10 are isolates from the same biological replicate.
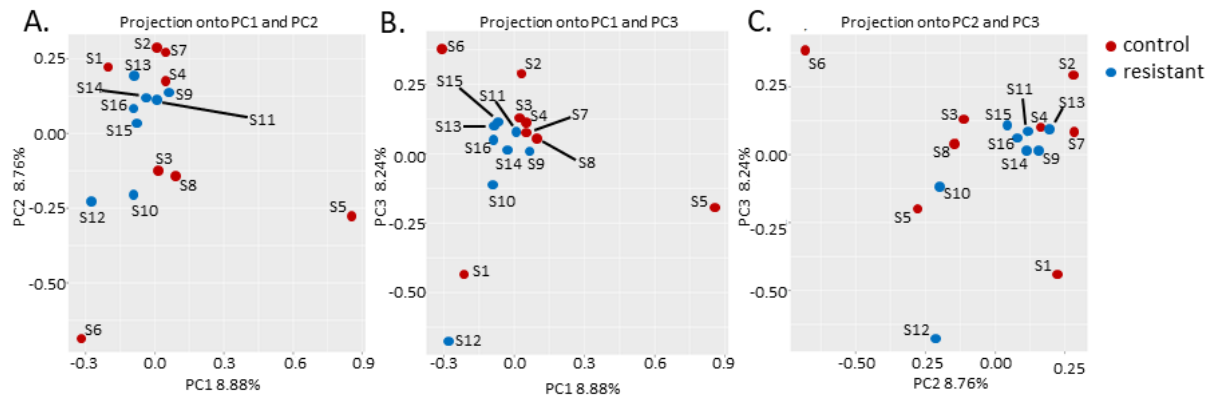
Figure 2: PCA analysis of evolved variants in In-Lab evolution strains according to resistance phenotype. The 6874 variants found in the control and MCHM strains after sequencing, aligning to the reference yeast genome, and filtering of existing variants in the YJM789 parent genome were considered to constitute the entirety of variants produced through the In-Lab evolution passaging experiment. There were approximate 2500 variants per strain, and many variants appeared in more than one strain. The 16 strains were clustered by PCA analysis according to their variant content. Strains are colored by their resistance phenotype, with the eight control strains in red and the eight MCHM resistant strains in blue. Samples are also labeled S1-S16 corresponding to strain names from Figure 1. The first three principal components are plotted against each other as follows: **A.** principal component 1 vs principal component 2, **B.** principal component 1 vs. principal component 3, and **C.** principal component 2 vs. principal component 3.
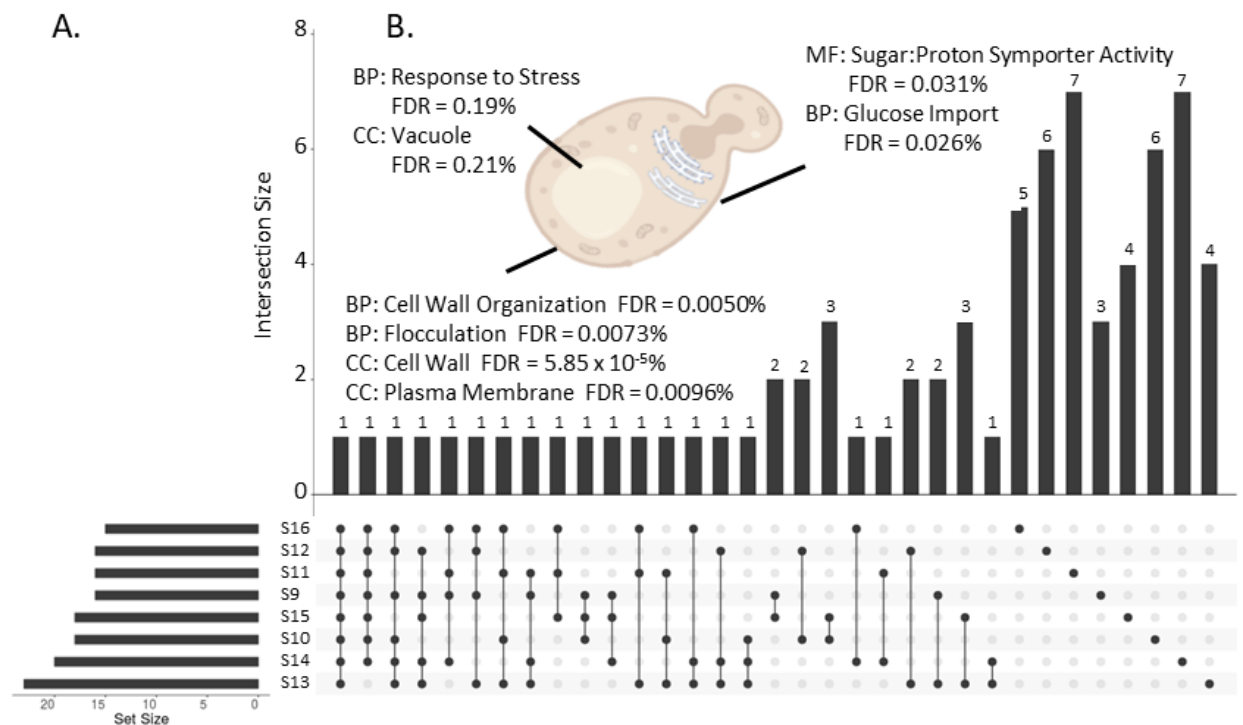
Figure 3: Intersection and functional analysis of genes containing variants in the MCHM evolved strains. A. Variants were filtered for their presence in coding sequences and for their presence exclusively in MCHM evolved strains. The variants from this list that resulted in nonsynonymous changes were found in 75 genes. Each of the eight MCHM strains contained nonsynonymous variants in between 15 and 23 genes, so there was some overlap in genes between strains. The intersection of the genes between the strains was analyzed by UpSetR. The bar graph vertical axis shows the intersection size, or the number of genes present in each intersection. The definition of each intersection appears in the filled and connected ovals below each bar. The eight strains are organized from smallest to largest (top to bottom) number of genes (set size) on the left. **B.** A similar set of genes as in part A, but containing both synonymous and nonsynonymous variants, consisting of 107 genes containing coding variants for the MCHM strains. These genes were analyzed for function, process, and cellular component terms. The GO Term and the False Discovery Rate (FDR) for each term is shown on the yeast cell diagram. FDR is represented as percentages, not ratios.
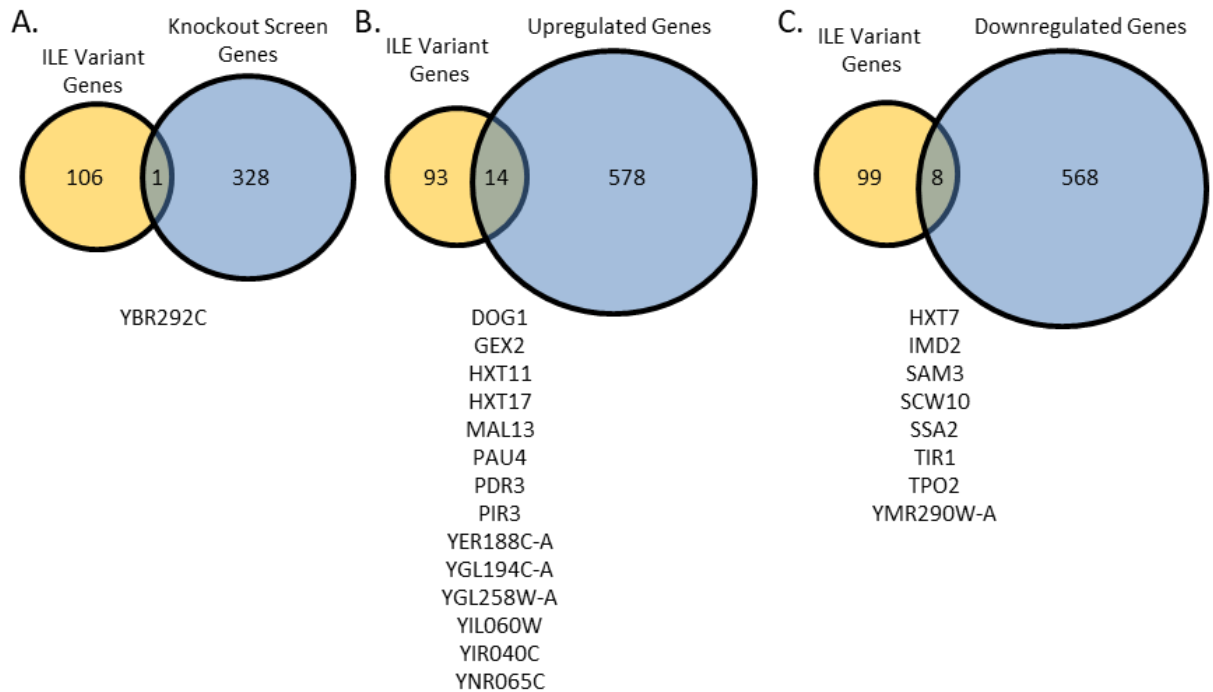
160

Figure 4: Venn diagrams show overlap between the genes with variants in MCHM evolved strains with genes in other MCHM genomic datasets. A. This Venn diagram shows the overlap of the gene lists from the MCHM evolution strains (ILE Variant Genes) and the genetic screen of the BY4742 strain knockout collection showing genes required for MCHM resistance (knockout screen genes) (Ayers et al., 2020). The number of genes found only in each list and in the overlap are shown in the diagram. The one gene in the overlap of the lists is written below the diagram. **B.** This Venn Diagram shows the overlap of the MCHM evolution strains genes and the upregulated genes from a transcriptome of MCHM treated BY4741 strain (Ayers et al., 2020). The 14 genes in the overlap of the datasets are listed below the diagram. **C.** This Venn Diagram shows the overlap of the MCHM evolution strains genes and the downregulated genes from the same transcriptomic analysis as B. The eight genes in the overlap of the datasets are listed below the diagram.
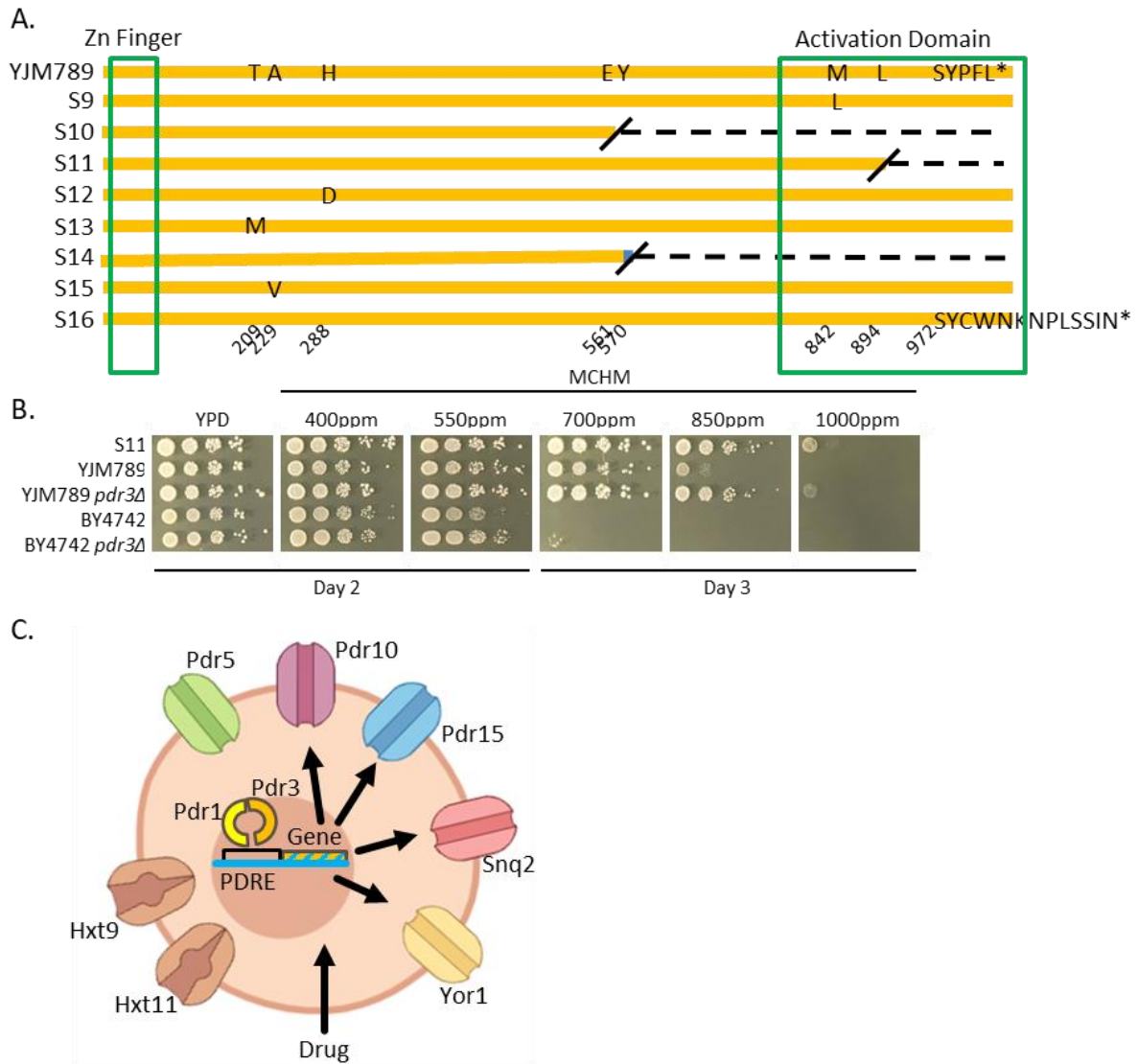
161

Figure 5: Pdr3 alleles in the MCHM evolved strains and how the pleiotropic drug response transporters are genetically controlled by Pdr1 and Pdr3. A. The eight MCHM evolved strains each contained exactly one mutation in the *PDR3* gene, all of which created protein sequence changes. The alleles are diagrammed alongside the wildtype YJM789 allele. The names of each strain are the left of each allele. The Zn Finger DNA-binding domain and Gal4-like Activation Domains are highlighted with green boxes in the N-terminal and C-terminal regions, respectively. The residue number for the start of each of the eight mutations is at the bottom of the diagram. Single letter amino acid abbreviations are used to indicate the identity of the residues in the parental allele. Each of the evolved strains (S9-S16) have exactly one mutation, so S9, S12, S13, and S15 show their single residue changes from the parent. The S10, S11, and S14 strains have early stop codons shown by a slash and dashed line where the missing portion of the protein coding sequence will not be translated. The final strain, S16, contains an insertion that alters the last five amino acids of the protein to a new 13 amino acid sequence. **B.** Shown

162

are serial dilution growth assays of one of the evolved resistant strains (S11), the YJM789 parent strain, the YJM789 *PDR3* knockout strain, the BY4742 wildtype strain, and BY4742 *PDR3* knockout. The growth assays are on increasing concentrations of MCHM from 0ppm (YPD) to 1000ppm. The day pictures were taken is shown underneath assays, day two or three shown for lower and higher concentrations of MCHM respectively to exemplify growth differences. **C.** This diagram models how Pdr1 and Pdr3 transcription factors control the expression of ABC transporters involved in the export of chemicals in response to stress. As cells are exposed to drugs, the transcription factor dimerizes, which results in binding to *cis*-regulatory elements called PDREs upstream of genes encoding various ABC transporters. The major ABC and hexose transporters (Pdr5, Snq2, etc.) contain PDREs bound by Pdr1 and Pdr3 are shown in the plasma membrane and labeled.
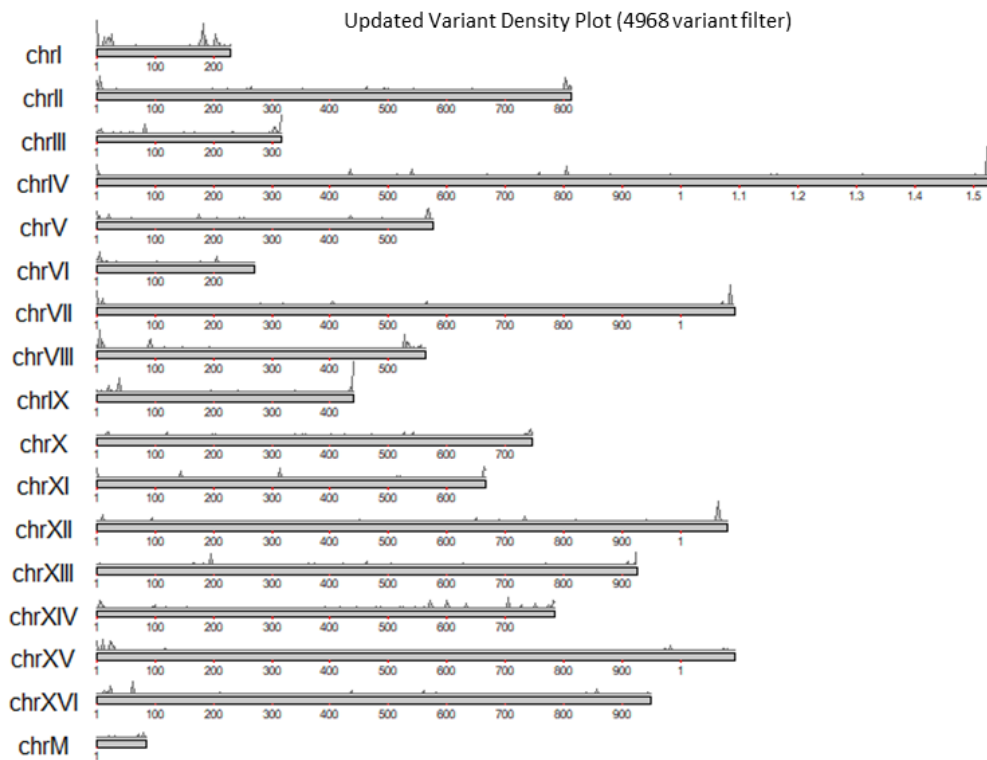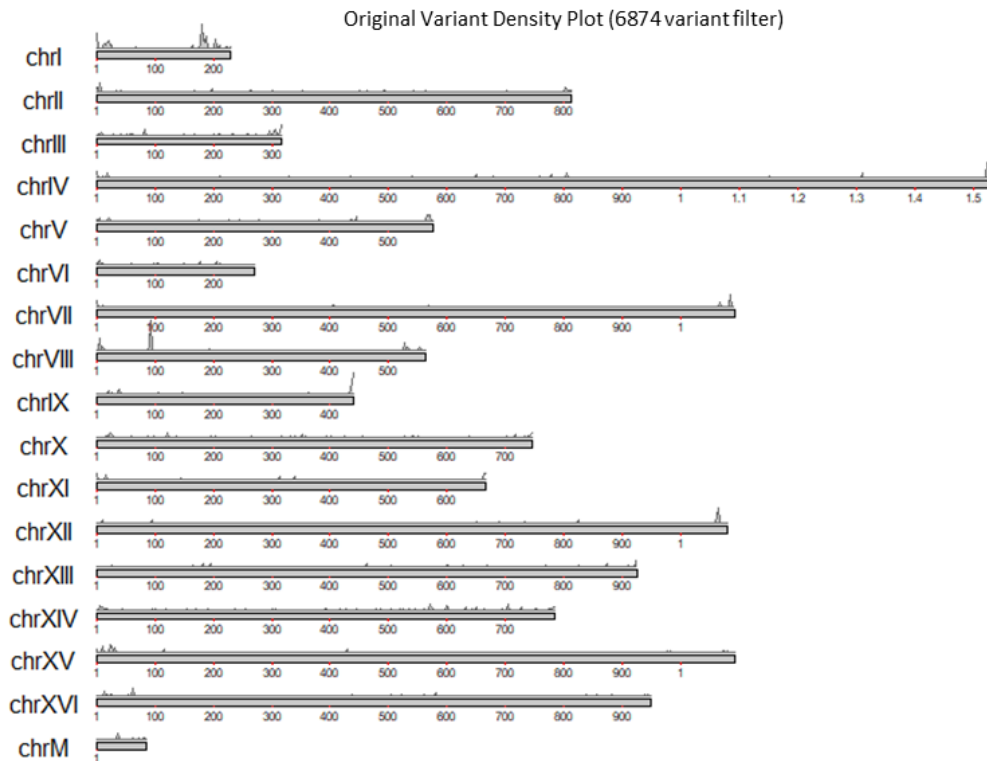
Supplemental Figure 1: Variant density plot of 6874 variants across the 16 evolved strains. A. The 6874 variants remaining after removing variants shared by all 16 strains and known YJM789

SNPs were plotted on the yeast genome. All instances of each variant in each strain were plotted, so approximately 40000 variants were plotted in total, as many remaining variants were shared by upwards of 10 strains each. Each chromosome is represented as a bar, with base pair locations labeled every 100kb. The variants are plotted as a density curve using 2.5kb windows along the top of each chromosome. **B.** The variants were filtered to only those shared among less than 10 strains, leaving 4968 individual variants. This translated to approximately 15000 instances of these variants across the 16 strains, which were then plotted on the yeast genome. Each chromosome is represented as a bar, with base pair locations labeled every 100kb. The variants are plotted as a density curve using 2.5kb windows along the top of each chromosome.
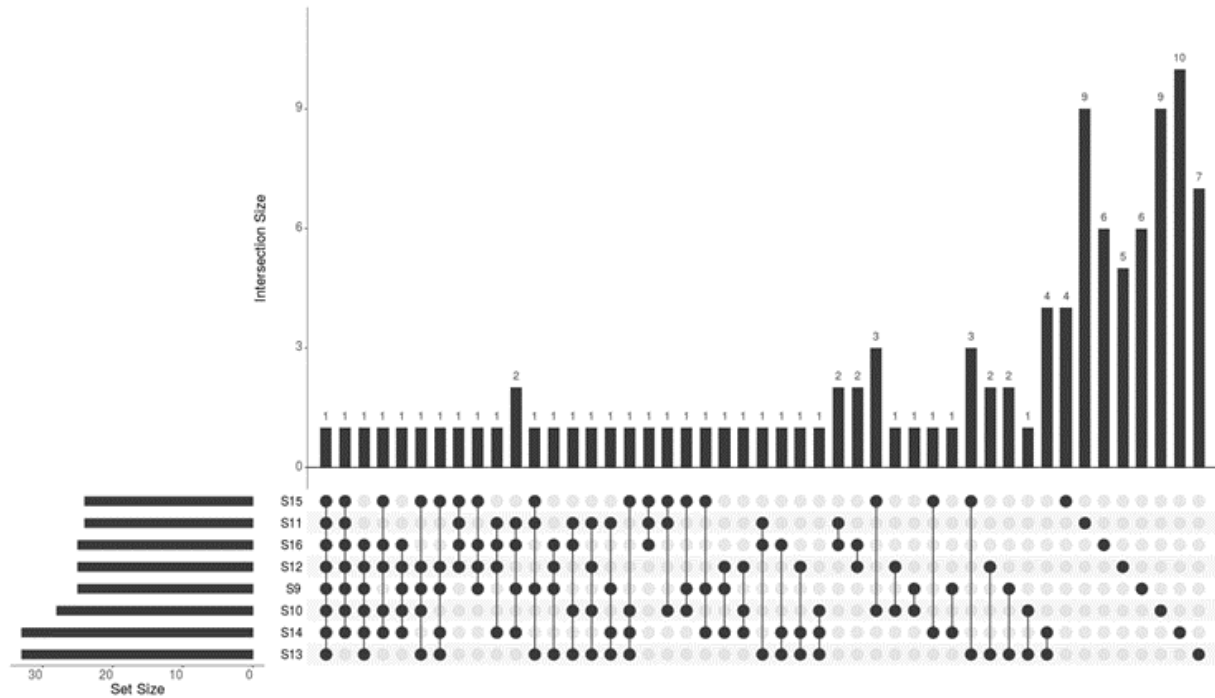
Supplemental Figure 2: Intersection of genes containing synonymous and nonsynonymous variants in the MCHM evolved strains. Variants were filtered for their presence in coding sequences and for their presence exclusively in MCHM evolved strains, as in Figure 2. The variants from this list that resulted in synonymous and nonsynonymous changes were found in 107 genes. The intersection of the genes between the strains was analyzed by UpSetR. The bar graph vertical axis shows the intersection size, or the number of genes present in each intersection. The definition of each intersection appears in the filled and connected ovals below each bar. The eight strains are organized from smallest to largest (top to bottom) number of genes (set size) on the left.

The following Supplemental Material is available at
https://github.com/wvumayers/MCHM.git.

Supplemental File 1: Tables of coding variants in each evolved MCHM strain. This file contains eight tabs, each with one of the eight MCHM evolved strains labeled S9 through S16. Each tab contains all the variants found in the coding regions of each strain, and its accompanying information. This includes the location of the variant, identity of the gene containing the variant, the nucleotide allele change from the variant, the coding consequence change from the variant, and a prediction using SIFT as to whether the mutation is tolerable or deleterious to protein function.

Supplemental File 2: Matrix of all mutations in 107 genes identified as containing unique variants in MCHM strains. The 107 genes that were identified as having unique variants in MCHM evolved strains did not necessarily have no variants at all in control strains. Therefore, a matrix was created using coding analysis in all 16 strains for these 107 genes. All synonymous and nonsynonymous changes are included in the analysis. The number of variants in each gene in each strain is shown, with summary columns to the right. These columns identify how many control and MCHM strains have at least one variant and the total number of variants contained in control and MCHM strains. Those genes with variants present in between one and three control strains are colored yellow, while those with zero variants in control strains are labeled red in the summary columns. Those genes with variants in greater than three MCHM evolved strains are labeled in green. The intersection of these two summaries (genes with variants in very few control strains but in many MCHM strains) helped identify *PDR3* as the only gene with mutations in all MCHM strains and no control strains.

Supplemental Table 1: R packages used to perform variant analysis. This table contains the identity of all R packages that were used, including dependencies, in the analysis of variants for annotation and coding prediction. The package is listed with its version used. The citation for each package is also included in the third and successive columns.

Supplemental Table 2: List of 107 genes containing unique variants in MCHM evolved strains. This list includes all genes that contain a unique variant that results in either synonymous or nonsynonymous changes to the protein. This list was used for the GO term analysis in Figure 3, the Venn diagrams in Figure 4, and the intersection graph in Supplemental Figure 2.

Supplemental Table 3: List of 75 genes containing unique variants causing nonsynonymous changes in MCHM evolved strains. This list includes all genes that contain a unique variant that results in only nonsynonymous changes to the protein. This list was used for the intersection graph in Figure 3.

Supplemental Table 4: Matrix table of all 6874 variants and their presence or absence in all 16 strains. This table uses a matrix containing binary values of 1 meaning presence of a variant in a strain and 0 meaning absence. Variant identity is described explicitly in the first column with *S. cerevisiae* chromosome reference id, bp location, and allele change shown. Columns 2 through 4 summarize the location of the variant using chromosome, start bp location, and stop bp location.

Chapter 6: CONCLUSION

The study of the chemical 4-methylcyclohexanemethanol following it spilling into the Elk River in January 2014 initially focused on its physical and toxicological properties. As summarized in Chapter 2, this chemical was relatively uncharacterized for any biological effects, so immediate work focused on determining any health outcome effects. The National Toxicology Program ran a battery of studies and came to the conclusion that, while MCHM and components of the spill could have developmental impacts such as reduced birth weight in rodents and skeletal changes in fish, and while it could cause neurological effects based on fish photomotor responses, neither of these outcomes occurred in concentrations simulating the levels found in the spill (*West Virginia Chemical Spill: Collective NTP Findings and Supporting Files*, n.d.). Follow-up analyses of real birth weight outcomes in the impacted region in the aftermath of the spill also showed no noticeable changes (Benson et al., 2018). There remained open questions about how MCHM impacted cell biology, which this dissertation sought to address. This dissertation explored the relationship between genotype and resistance phenotype for MCHM in yeast.

The data shown in Chapter 3 reveals that the way MCHM initially impacts yeast cells is by very quickly causing cell cycle arrest in G1. The chemical stress does not kill the cells because they successfully respond with a pre-programmed cellular network called the environmental stress response. This is confirmed by the transcriptome data that identified downregulation of ribosome biogenesis genes and upregulation of stress genes involved in the pleiotropic drug response, oxidative stress response, and metal and ion homeostasis. The source of the MCHM-induced stress that activates the ESR in this way is likely to be nutrient deprivation. The genetic

screen of the BY4742 knockout collection showed the requirement of aromatic amino acid

biosynthetic genes *ARO1, ARO3, TRP2, TRP3, TRP4,* and *TRP5*, all of which were also

upregulated in the wildtype transcriptome data. While supplementation of tryptophan,

tyrosine, and phenylalanine was not sufficient to rescue growth phenotypes in MCHM, it is

possible these end products cannot prevent the arrest that impacts growth on MCHM and are

more important to resume growth after cells adapt to the stress. Other metabolites in the

aromatic amino acid and *trp* pathways, such as chorismate and indoles, may also be important

instead of or in combination with the amino acid products. Finally, the proteins themselves may

have unannotated functions other than amino acid biosynthesis. Knockouts of *TRP2* in

particular have been shown to grow poorly even with tryptophan supplementation, which

rescues other *trp* mutants (Krömer et al., 2013).

The nutrient deprivation signal is a likely cause of the early activation of cell cycle arrest

and ESR programming in yeast, but MCHM also induces reactive oxygen species over longer

time courses. After 6 and 12hrs, MCHM produces ROS, as detected by flow cytometry in

Chapter 3. A consequence of the production of ROS in the cell can be DNA damage, which was

suggested as a consequence of MCHM + S9 treatment based on expression of DNA damage

response proteins like Rad51 in the earliest yeast study on MCHM (Lan et al., 2015). The comet

assay in the Chapter 3 shows that DNA damage is a result of MCHM treatment of yeast. The

ROS and resulting DNA damage stresses are key consequences of MCHM exposure in yeast,

though they are unlikely to cause the initial cell cycle arrest and environmental stress response

programming. Instead, they may be a result of the cell's metabolism of MCHM over time. We

propose that once the cell has metabolized and overcome any concomitant stress induced by the metabolites, it resumes growth and the cell cycle.

The ability for the cell to overcome the ROS and related stresses from MCHM is dependent on the cell's ability to respond to ion, metal, and ROS homeostasis. The transcriptome of Chapter 3 points to mitochondrial genes and the vacuolar ATPase, which are important to these pathways, but Chapter 4 provides the evidence that confirms their roles. Petite yeast, missing functional mitochondria, are sensitive to ROS and metal stressors (Grant et al., 1997), so their sensitivity to MCHM is as predicted. The QTL experiment of MCHM growth in the YJM789xS96 yeast segregant collection pointed to loci spread broadly throughout the genome with only one peak able to pass statistical significance. This peak was above a gene for the protein Yke4, a zinc transporter in the ER. The levels of zinc increase with MCHM in resistant grande yeast and the BY4741 strain knockout for *YKE4* is sensitive, but growth is rescued with supplementation of low amounts of zinc to the growth media. The simplest explanation points to increased levels of zinc for resistance, but the sensitive parent in the QTL has significantly higher intrinsic levels of zinc. As higher levels of zinc supplementation fail to rescue MCHM growth, it seems that it is a more complicated role in the tight control of zinc levels in the cell that affect resistance. MCHM acts as a hydrotrope, affecting the solubility of proteins *in vitro* and reducing aggregation to 60% of no treatment in the aggregation assay of Chapter 4. Zinc treatment increases aggregation in the same assay. The intracellular and extracellular loop domains of transmembrane transporters such as Yke4 have disordered regions that may be affected by this solubility effect. In this way, MCHM can affect the ability of

the cell to control intracellular and compartmental levels of nutrients. The tight control of zinc

levels may work *in vivo* to allow cells to adapt to hydrotrope effects of MCHM.

The study of genotypes related to MCHM resistance in the genetic screen of Chapter 3

and the QTL of Chapter 4 was continued with an In-Lab evolution experiment in Chapter 5. The

QTL resistant parent strain S96 was unable to produce an evolved strain with improved

resistant via ILE over several attempts. The evolution of the sensitive QTL parent strain YJM789

produced eight resistant strains. Variant analysis of these strains and eight control strains

evolved similarly, but without MCHM, showed thousands of mutations spread broadly

throughout the genome in all strains. There was no clear pattern according to clustering

analysis of these mutations for resistance phenotype. In agreement with the peaks of the QTL

in Chapter 4 that did not reach significance, the variants spread broadly throughout the

genome point to resistance to MCHM involving small contributions of many cellular pathways

and networks to produce the single complex phenotype. However, the ILEs did point to one

gene in particular that may be a large determinant of MCHM resistance. The gene *PDR3* had a

mutation in every strain that became resistant to MCHM and none in any of the control strains.

This was the only gene to have this profile. Every other gene that had mutations in the resistant

strains also contained mutations in the control strains. The *PDR3* mutations resulted in protein

sequence changes, including four with single amino acid changes, three with C-terminal

truncations, and one with an insertion changing the five C-terminal amino acids to a new 13

amino acid terminus. The effects of these eight new alleles on resistance mechanisms are still

under investigation, but the gene knockout in the YJM789 parent strain is sufficient to produce

resistance similar to the evolved strains. The *PDR3* gene is a nominal target for gain-of-function

171

resistant genotypes due to its role in controlling the expression of ABC transporters that pump chemical stressors out of the cell. In retrospect the importance of *PDR3* mutations in the induction of new resistance phenotypes may also point to why the resistant parent strain S96 was unable to easily produce evolved resistance. While the wildtype transcription factor shows few mutations between strains, several ABC transporter genes whose expression is controlled by Pdr3 are polymorphic between the strains. There is a large 637 amino acid truncation in the N-terminus of Snq2 in the YJM789 strain and a 5.3% amino acid divergence spread throughout the protein of Pdr5 between the two strains (Wei et al., 2007). Differences in ABC transporter complements may have prevented the large-scale changes that induced resistance in YJM789 from being replicated. Indeed, while a knockout of *PDR3* in the BY4742 strain, which is closely related to S96, did produce some increased resistance, it was not sufficient to produce resistance beyond 550ppm exposure.

The role of genotypes in creating stress response phenotypes was explored in this dissertation using the toxicologically-relevant spill chemical MCHM. We found that wildtype yeast strains exhibit MCHM tolerance that requires aromatic amino acid biosynthetic genes, redox stress responsive genes, mitochondrial genes, and metal and ion homeostatic genes to mitigate nutrient, ROS, and protein folding stress from the hydrotrope nature of the chemical. These many sources of stress rely on pathways and networks within the cell that are impacted by loci spread broadly throughout the genome, as seen through QTL, screen, and ILE analyses. This study was able to identify only one locus where the genotype could produce an increased resistant phenotype beyond wildtype levels, the transcription factor controlling expression of multiple drug transporters. Stress responsive phenotypes are controlled by many important

small effect loci, but some individual loci may provide outsize effect on chemical sources of

stress.

References

Benson, S. M., Ruestow, P., Keeton, K. A., Novick, R. M., Marsh, G. M., & Paustenbach, D. J. (2018). The 2014 crude 4-methylcyclohexanemethanol chemical release and birth outcomes in West Virginia. *Archives of Environmental & Occupational Health*, *73*(5), 292–301. https://doi.org/10.1080/19338244.2017.1350132

Grant, C. M., MacIver, F. H., & Dawes, I. W. (1997). Glutathione synthetase is dispensable for growth under both normal and oxidative stress conditions in the yeast Saccharomyces cerevisiae due to an accumulation of the dipeptide gamma-glutamylcysteine. *Molecular Biology of the Cell*.

Krömer, J. O., Nunez-Bernal, D., Averesch, N. J. H., Hampe, J., Varela, J., & Varela, C. (2013). Production of aromatics in Saccharomyces cerevisiae—A feasibility study. *Journal of Biotechnology*, *163*(2), 184–193. https://doi.org/10.1016/j.jbiotec.2012.04.014

Lan, J., Hu, M., Gao, C., Alshawabkeh, A., & Gu, A. Z. (2015). Toxicity assessment of 4-Methyl-1-cyclohexanemethanol and its metabolites in response to a recent chemical spill in west virginia, USA. *Environmental Science and Technology*, *49*(10), 6284–6293. https://doi.org/10.1021/acs.est.5b00371

Wei, W., McCusker, J. H., Hyman, R. W., Jones, T., Ning, Y., Cao, Z., Gu, Z., Bruno, D., Miranda, M., Nguyen, M., Wilhelmy, J., Komp, C., Tamse, R., Wang, X., Jia, P., Luedi, P., Oefner, P. J., David, L., Dietrich, F. S., … Steinmetz, L. M. (2007). Genome sequencing and comparative analysis of Saccharomyces cerevisiae strain YJM789. *Proceedings of the National Academy of Sciences*, *104*(31), 12825–12830. https://doi.org/10.1073/pnas.0701291104

West Virginia Chemical Spill: Collective NTP Findings and Supporting Files. (n.d.). National

Toxicology Program U.S. Department of Health and Human Services. Retrieved March 8, 2017,

from https://ntp.niehs.nih.gov/results/areas/wvspill/collective.html