Title
Interventions to Encourage and Facilitate Greener Industrial Chemicals Selection

Author
Faulkner, David

Publication Date
2017

Peer reviewed|Thesis/dissertation

Interventions to Encourage and Facilitate Greener Industrial Chemicals Selection

by

David Michael Faulkner

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular Toxicology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Christopher D. Vulpe, Co-Chair
Associate Professor Jen-Chywan Wang, Co-Chair
Associate Professor Daniel K. Nomura
Professor John Arnold

Fall 2017

Abstract

Interventions to Encourage and Facilitate Greener Industrial Chemicals Selection

by

David Michael Faulkner

Doctor of Philosophy in Molecular Toxicology

University of California, Berkeley

Professor Christopher D. Vulpe, Co-Chair

Associate Professor Jen-Chywan Wang, Co-Chair

Despite their ubiquity in modern life, industrial chemicals are poorly regulated in the United States. Statutory law defines industrial chemicals as chemicals that are not foods, drugs, cosmetics, nor pesticides, but may be used in consumer products, and this distinction places them under the purview of the Toxic Substances Control Act (TSCA), which received a substantial update when the US congress passed a revision of the act in 2016. The revised law, the Frank R. Lautenberg Chemical Safety for the 21st Century Act addresses many but not all of TSCA's failings, and rightfully emphasizes the development and adoption of high throughput screens, *in vitro*, and alternative assays to improve the process for registering new chemicals and to address the tens of thousands of untested chemicals currently in the TSCA inventory. As the discipline of toxicology gradually shifts from its' history as a reactive science (responding to problems after they've occurred) to a proactive science (attempting to predict and circumvent dangers to human and environmental health), two things become clear: 1.) traditional low throughput toxicological testing methodologies are inadequate to address both the volume of chemicals of interest and the pace of research; and 2.) the modern industrial chemical ecosystem is complex and no single testing solution will be appropriate for all the actors that populate that ecosystem.

To address these challenges, three interventions are proposed, each of which targets a different population within the industrial chemical ecosystem. The first intervention is a suite of computational toxicology methods targeted towards chemists in the initial phases of chemical design and development. The second intervention is an alternative assay, the yeast functional toxicogenomic assay, permits industrial or government labs to rapidly investigate differences in cytotoxic mechanism between different chemicals – even if they are structurally very similar. The third intervention explored in this work is a method for enhancing the metabolic capacity of cell lines currently used by regulators for high throughput cytotoxicity testing. These interventions individually are not necessarily appropriate for all actors across the US industrial chemicals ecosystem, but as bespoke solutions they may be quite useful, and, it is hoped, support a larger exploration of where and how similar efforts may be spent most effectively to reduce industrial chemical hazard.

Dedication page

I dedicate this dissertation to my mother and sister, whom I love most dearly.

Table of Contents

List of Figures and Tables

Chapter 2: Tools for Green Molecular Design to Reduce Toxicological Risk

Chapter 3: Functional Toxicogenomics and Combinatorial Chemistry For Greener Design

Chapter 4
Using CRISPR To Enhance The Metabolic Capacity Of Cell Lines Commonly Used In High Throughput Screening

Acknowledgements

I cannot properly express my gratitude to all of the patient, generous, and loving people in my life who made this dissertation possible, to do so would necessitate at least the space of another one hundred and eighty or so pages, and even then I'm sure I'd be forgetting someone. Nevertheless:

Thank you Chris, first and foremost, for all of your support over these last four and a half years. Thank you for indulging my intellectual flights of fancy and for giving me so many opportunities to explore and grow. Your patient generosity and brilliance have been tremendous gifts during my time in the Vulpe lab, and I am profoundly indebted to you. I'll even forgive you for moving to Florida.

Thank you Mani, for all of your help and training in lab. It is no exaggeration to say that there's no way that any of this would have gotten done without your teaching and encouragement. I'll always treasure our lengthy and winding conversations about language, culture, history, and, of course, potatoes.

Thank you Vanessa, for being my friend and teaching me about the wide wonderful world of yeast work. My only regret is that we now live on opposite sides of the country!

Thank you Marty and Tom, for your support and for making the BCGC my adoptive home here on campus. I'm thrilled to be joining the Center in January.

Thank you Heather for your friendship, and for pulling me into so many new and exciting green chemistry projects. Your support and kindness over the last couple years have meant so much to me.

Thank you Faith for all of your help in lab, and for making Florida feel like home, at least for a little while. Us Michiganders have to stick together.

Thank you Brianna for all your help in getting the comptox project off the ground and for being a friend (and NST gossip buddy).

Thank you Amanda, Alexis, Devrah, Justine, Cody, Candice, and everyone else at CEHT for adopting me during my time in Florida.

Thank you Jon and Allison for all the bowling nights, the questionable skiing adventure, and Bay to Breakers.

Thank you to my family, for all of the love and encouragement.

And lastly, thank you to my second family: Joe, Nick Jarosz, Randy, Miyoko, Amy, Brett, Nick Beier, and Zach, for all the emotional and spiritual support over the years. I love you all.

Chapter 1
Introduction to Green and Industrial Chemistry

A Brief History Of The Green Chemicals Industry In Relation To Consumer Products

It is estimated that the global chemical manufacturing industry produces between 70,000 and 100,000 chemicals each year, and the US EPA reports that it receives roughly 1,000 new chemical pre-manufacture notices annually[1,2]. Unfortunately, only a few thousand of the chemicals currently in trade in the United States have been thoroughly evaluated for chemical hazard due to the material and temporal cost of traditional toxicological methods. Increasing consumer demand and recent changes in chemical regulation in the US and Europe have spurred a renaissance in the field of hazard assessment, leading companies, regulators and researchers to develop new methods for rapid chemical hazard assessment and to further refine existing methods to improve cost, speed, and accuracy.

Consumer demand has been a powerful accelerant for the adoption of more comprehensive chemical testing, and most large producers of consumer products consider sustainability metrics when developing marketing materials[3]. Just as the modern environmental movement began with Rachel Carson's 1962 book "Silent Spring," and lead to the passing of the Clean Water Act, The Toxic Substances Control Act (TSCA), and the founding of the US EPA, increasing evidence of consumer exposure to, and absorption of, chemicals found in consumer products[4] has spurred a transition from concerns about environmental restoration and protection to a more preemptive model of pollution prevention[5]. Recent studies have identified common antimicrobial compounds such as triclosan as well as brominated flame retardants and plasticizing compounds in body fluid samples from a representative sample of the American population[6]. Alarmingly, it is clear that exposure to brominated flame retardants – which have been linked to a variety of health effects related to endocrine disruption – commonly occurs through household dust generated in part by consumer products such as furniture or electrical appliances[7]. The furor that these findings have generated, as well as massive public backlash to chemicals commonly found in consumer products such as Bisphenol A[8], have contributed to the groundswell of concern about industrial chemicals that has been fomenting for decades.

Part of the regulatory response to public concern about chemicals in consumer products is the Lautenberg amendment to the TSCA, which expands the EPAs regulatory authority over industrial chemicals, requiring that new chemicals must demonstrate safety before they can be brought to market and enabling the EPA to require additional safety testing for existing chemicals if there is insufficient data to evaluate chemical risk[9]. This brings the US chemical policy closer in line with the European Union's REACH regulation, which requires that companies register substances that they are using and to provide environmental and human health data for each substance[10]. In both regulatory schema, alternative – i.e. non-animal – testing strategies are encouraged, as both a means of reducing testing costs and as part of a larger effort to streamline the data collection process.

Of course, neither producers of consumer products nor the chemical industry are blind to these societal and regulatory changes and have over the past few decades bent in the direction of environmental sustainability through the embrace of "Green Chemistry[11]." The

field of "Green Chemistry" has among its tenants the principle that chemicals should be "benign by design" – that is, they ought to possess the minimum necessary hazard to human and environmental health as can be achieved while preserving their function. Of the 12 principles of green chemistry articulated by Anastas and Warner[12], "benign by design" has proven to be the most difficult to achieve for obvious reasons: unlike improving atom economy or reducing solvent use, the process of developing less hazardous chemicals is spectacularly complex and fraught with all the peril that is endemic to the field of biology. Chemical reactions are complicated, but living organisms are worlds unto themselves. The challenge of reducing hazard while maintaining performance requires not only the faculties of a savvy toxicologist, but the investment of effort by chemists, engineers, and designers – hardly a trivial problem[11].

Despite the costs, the result of this combination of consumer demand, corporate strategy and shareholder fiat[13,14] is that most of the worlds largest producers of consumer products have environmental sustainability and safer chemicals policies in various stages of development[15]. Microsoft freely distributes its list of substances restricted from use in its products[16], while Amazon offers guarantees for responsible materials sourcing and conflict minerals disclosures[17]. Wal-Mart provides a thorough treatment of their chemicals policy which directly references the text and principles of Green Chemistry as formulated by Anastas and Werner[12] alongside announcements of joining The Chemical Footprint Project[18,19] in 2017, and requirements for its suppliers to provide materials disclosures for priority chemicals in their products[20]. These efforts pale in comparison to those of Apple, which has published Environment reports that detailing metrics for the company's energy, water, and raw materials usage since 2008, and as of this writing, all of these reports are still publically available on the company website[21]. While some of the early reports provide limited, basic information about energy and water usage, in recent years, the environmental safety reports have become incredibly thorough documents with surprising granularity of data disclosure, and more specialized reports are available for individual products. In 2016, the company announced its full materials disclosure (FMD) program, which sought to work up their supply chains and identify every chemical substance composing every part in every product. As of 2017, more than 20,000 chemicals have been identified, comprising nearly the entire repertoire of Apple products. This is a rare accomplishment among producers of consumer products, allowing Apple to join SeaGate as one of the few electronics companies to have completed an FMD program[19]. Such programs are by far the exception, rather than the rule, owing to the cost and difficulty of such projects. The result of such investments is that FMD provides companies the ability to improve their product chemistries as never before, because safer products cannot be developed if the product composition is not known[19].

Once the constituent chemicals of a product have been identified, it becomes possible to perform alternatives assessments and determine which, if any chemicals might provide the same functionality with lower human and environmental health costs. An ecosystem of consulting firms, databases, and methodologies has developed to provide scientific support and certifications for companies with an interest in safer consumer chemistry. The Healthy Building Network, a 501 (c)(3) non-profit, developed Pharos, a database of chemicals commonly found in building materials, as well as the available human and environmental

health data for those chemicals[22]. Interested parties can use Pharos to compare the hazard endpoints of one chemical versus another, and use that information to guide their decision-making. Pharos incorporates the GreenScreen method of hazard assessment, a framework developed by Clean Production Action, another 501 (c)(3) organization[23], to provide simplified summaries of hazard endpoints for users without substantial toxicology expertise to ease the process of comparing chemicals.

The practical challenges of producing safe and sustainable products has led to "greenwashing" behaviors wherein companies make bold claims about a product's safety and environmental friendliness, but do not follow through on the important work of ensuring that these claims are accurate. As a result, eco-conscious consumers not only become disenchanted[15], but may misuse a product or engage in unsafe product use practices due to inaccurate perceptions of its safety profile or health benefits[24]. While there is room for optimism in some the recent market shifts towards environmental ethics in business practices[25,26], if the full environmental and public health potential of these changes are to be realized, more rigor and thoughtfulness must be applied to the practical challenges of developing new and alternative chemical solutions that are indeed "greener" but which also satisfy market and regulatory demands.

While the growth and development of the green chemistry industry is encouraging, and there are many signs of progress as producers of consumer products gradually emphasize safer product chemistries, significant challenges remain, particularly in the field of chemical hazard assessment. Data gaps continue to be a significant challenge for alternatives assessments, despite the proliferation of chemical hazard databases such as Pharos, the Chemical Footprint Project[18], EPA's ToxCast[27] and Toxnet[28], and ECHA's Registered Substances Database[29]. Though useful in narrow circumstances, these databases only house a fraction of the thousands of compounds used in the production of consumer products, and in many cases entries are limited to only a few hazard endpoints. To address this problem, significant developments in the field of chemicals testing are necessary, as is a willingness to implement them at crucial points in the development pipeline for consumer products.

The Necessity Of Bespoke Solutions

The modern industrial chemical system is sprawling and complex, resistant to any single palliative that will induce all synthetic chemists to generate benign compounds and also enable regulators to work through the backlog of chemicals awaiting investigation. To borrow a theoretical tool from the field of Public Health, one must meet the subject of an intervention where they are, and one must tailor the intervention to suit the needs of the subject, otherwise, no improvement may be expected[30,31]. Therefore, this dissertation explores three different interventions, each intended to improve product chemical safety at a different part of the industrial chemical apparatus: The first, aimed at the earliest phase of synthetic chemistry, is an exploration of the computational tools which may enable chemical designers to incorporate reduced hazard into their lead compounds from the very beginning. Second, the utility of an alternative testing strategy, functional toxicogenomics, is presented to indicate that it may be used by industrial or regulatory actors to assess compounds for safety, and potentially improve the chemical design process. Finally, the

nascent CRISPR technology is examined to determine how it may be used to retrofit existing high throughput screens to improve predictive accuracy with minimal effort.

Computational Toxicological Tools For Greener Molecular Design

The most effective means of reducing chemical hazard is to avoid producing hazardous chemicals in the first place. This notion is not purely aphoristic, but rather, a statement of intent with regards to the first intervention strategy which is discussed: the use of computational toxicology programs to guide chemists in the design of new molecules. This chapter demonstrates that although very few computational toxicology programs are intended for designing industrial chemicals, some programs intended for pharmaceutical development and medicinal chemistry may be used to this end. A survey of computational toxicology platforms or models is presented to identify those tools which are most accessible to the average synthetic chemist who possesses minimal toxicological expertise, but is interested in designing safer molecules. Broadly speaking, the "green" chemist wishes to develop compounds without any of the properties which are prized by pharmacologists, and seeks to minimize a compounds' uptake, stability, and biological activity. Building on the work of Voutchkova et al[32,33], physicochemical properties that facilitate uptake, stability and biological activity of molecules are considered using principles derived from medicinal chemistry: size, number of hydrogen bonding moieties, polar surface area, pKa, and logP. Then, a set of three different workflows is developed to guide green chemists through the process of green molecular design[11], which accounts for their level of toxicological expertise and available funding. A version of this chapter has been published as a chapter in a textbook on computational pharmacology and toxicology[34].

Functional Toxicogenomics and Combinatorial Chemistry Streamline Hazard Evaluation

Once a chemical has been developed, it must be tested for potential hazards, a process which may take months or years using traditional models for hazard assessment. To remedy this, regulators have encouraged the development of alternative assays to streamline the process of screening chemicals, one of which is the yeast functional toxicogenomic assay. The yeast functional toxicogenomic assay has been used to identify mechanisms of cytotoxicity for thousands of chemicals and to dissect biochemical pathways in baker's yeast (*Saccharomyces cerevisiae*), improving both our understanding of chemical toxicity and of basic cellular biology[35–38]. Here, this assay to test the biofuel candidate compound 2,5-dimethylfuran as well as a series of related compounds to determine if minor modifications to the core furan ring would result in different mechanisms of cellular toxicity.

Because 2,5-dimethylfuran (2,5-DMF) has such promise as a biofuel, it is prudent to thoroughly evaluate it for toxicological hazard before adopting it as a replacement for petroleum, especially since there is some evidence that it may induce double-strand breaks in genomic DNA[39]. However, merely testing 2,5-DMF does not help chemists working in the chemical space of simple furans, and there is ample evidence throughout the history of toxicology indicating that minor changes in structure may result in substantial changes in toxicity. Therefore, the related compounds 2-methylfuran, 2-ethylfuran, and 2,3-

dimethylfuran were also screened to see if changes to the functionalization of the furan ring ameliorated the genetic toxicity endpoint. Bar-Seq analysis of the data from the yeast functional toxicogenomic assay suggests that under chronic exposure conditions, 2,5-DMF does indeed result in DNA damage, including double strand breaks, likely as a result of DNA-protein cross-links, supporting earlier findings for 2,5-DMF's genotoxic properties. However, Bar-Seq data also indicate that these DNA damaging properties may be unique to 2,5-DMF, and that the other furan compounds tested here mediate their cytotoxic effects through nonspecific oxidative protein damage and interference with intracellular transport pathways. The implications of these findings are twofold: 1. the chapter demonstrates the validity of GMD as a means to develop functionally similar molecules with different toxicological properties, and 2. the work adds to the corpus of evidence supporting the use of the yeast functional toxicogenomic assay for rapid and scalable assessment of industrial compounds.

Use Of CRISPRa To Improve Metabolic Competence Of In Vitro Cultures For High-Throughput Screens

Regulators and industrial firms alike make use of mammalian cell culture and high-throughput screening (HTS) assays to provide baseline activity data and hazard assessment data for chemicals, often using immortalized human liver or kidney cell lines.[40] Unfortunately, it has been demonstrated that the metabolic capabilities of two cell lines commonly used for HTS, HEK293-T and HepG2, are far inferior to their primary cell equivalents – kidney and liver cells, respectively. This metabolic deficit means that bioactivated compounds may appear less toxic in HTS while detoxified compounds may appear more toxic. Several methods exist for restoring metabolic competence to cultured cells, including the use of DMSO[41], mRNA, and (other techniques), but these techniques are invasive and often necessitate the use of adjuvants to work, which complicates and may confound testing done with them. As a drop-in solution, the CRISPRa method described by Konermann et al[42]. is employed to develop a simple transfection that will result in up-regulated metabolism-related genes which are otherwise suppressed in HepG2 and HEK293T cell cultures. The development and optimization of this technique is presented, along with early evidence suggesting the potential for this technology to improve metabolic competence in HTS.

References

1. Clomburg, J. M., Crumbley, A. M. & Gonzalez, R. Industrial biomanufacturing: The future of chemical production. *Science* **355,** aag0804 (2017).
2. Card, M. L. *et al.* History of EPI Suite™ and future perspectives on chemical property estimation in US Toxic Substances Control Act new chemical risk assessments. *Env. Sci Process. Impacts* **19,** 203–212 (2017).
3. Hofenk, D., Birgelen, M. van, Bloemer, J. & Semeijn, J. How and When Retailers' Sustainability Efforts Translate into Positive Consumer Responses: The Interplay Between Personal and Social Factors. *J. Bus. Ethics* 1–20 (2017). doi:10.1007/s10551-017-3616-1
4. Wambaugh, J. F. *et al.* High Throughput Heuristics for Prioritizing Human Exposure to Environmental Chemicals. *Environ. Sci. Technol.* **48,** 12760–12767 (2014).
5. History of Green Chemistry. Available at: https://www.acs.org/content/acs/en/greenchemistry/what-is-green-chemistry/history-of-green-chemistry.html.
6. Crinnion, W. J. The CDC fourth national report on human exposure to environmental chemicals: what it tells us about our toxic burden and how it assist environmental medicine physicians. *Altern. Med. Rev. J. Clin. Ther.* **15,** 101–109 (2010).
7. Fromme, H., Becher, G., Hilger, B. & Völkel, W. Brominated flame retardants – Exposure and risk assessment for the general population. *Int. J. Hyg. Environ. Health* **219,** 1–23 (2016).
8. Resnik, D. B. & Elliott, K. C. Bisphenol A and Risk Management Ethics. *Bioethics* **29,** 182–189 (2015).
9. USEPA. Highlights of Key Provisions in the Frank R. Lautenberg Chemical Safety for the 21st Century Act. *Assessing and Managing Chemicals Under TSCA* Available at: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/highlights-key-provisions-frank-r-lautenberg-chemical.
10. European Agency For Safety At Work. REACH - Regulation for Registration, Evaluation, Authorisation and Restriction of Chemicals. (2017). Available at: https://osha.europa.eu/en/themes/dangerous-substances/reach.
11. Coish, P. *et al.* Current Status and Future Challenges in Molecular Design for Reduced Hazard. *ACS Sustain. Chem. Eng.* **4,** 5900–5906 (2016).
12. Anastas, P. & Warner, J. *Green Chemistry: Theory and Practice, Oxford University Press: New York, 1998, p.30. By permission of Oxford University Press.* (Oxford University Press, 1998).
13. Scruggs, C. E. & Van Buren, H. J. Why Leading Consumer Product Companies Develop Proactive Chemical Management Strategies. *Bus. Soc.* **55,** 635–675 (2016).
14. IEHN — Shareholder Resolutions. Available at: http://iehn.org/resolutions.shareholder.php. (Accessed: 11th December 2017)
15. Nyilasy, G., Gangadharbatla, H. & Paladino, A. Perceived Greenwashing: The Interactive Effects of Green Advertising and Corporate Environmental Performance on Consumer Reactions. *J. Bus. Ethics* **125,** 693–707 (2014).
16. Sustainable Production Process | Microsoft Environment. *Microsoft* Available at: https://www.microsoft.com/en-us/environment/product. (Accessed: 11th December 2017)

17. Sustainability-Responsible Sourcing. Available at: https://www.amazon.com/p/feature/uknj5z35m3ev8as. (Accessed: 11th December 2017)
18. The Chemical Footprint Project. Available at: http://www.chemicalfootprint.org/. (Accessed: 11th December 2017)
19. Konkel, L. Chemical Footprinting: Identifying Hidden Liabilities in Manufacturing Consumer Products. *Environ. Health Perspect.* **123,** A130–A133 (2015).
20. Walmart Commitment to Sustainable Chemistry - Walmart Sustainability. Available at: https://www.walmartsustainabilityhub.com/sustainable-chemistry/sustainable-chemistry-policy. (Accessed: 11th December 2017)
21. Environment - Reports. *Apple* Available at: https://www.apple.com/environment/reports/. (Accessed: 11th December 2017)
22. Pharos Project. Available at: https://www.pharosproject.net/. (Accessed: 11th December 2017)
23. Chemicals, G. F. S. GreenScreen® For Safer Chemicals | An open, transparent, and publicly accessible…. Available at: https://www.greenscreenchemicals.org. (Accessed: 11th December 2017)
24. Royne, M. B., Levy, M. & Martinez, J. The Public Health Implications of Consumers' Environmental Concern and Their Willingness to Pay for an Eco-Friendly Product. *J. Consum. Aff.* **45,** 329–343 (2011).
25. The Sustainability Consortium – Sustainable Products, Sustainable Planet. *The Sustainability Consortium* Available at: https://www.sustainabilityconsortium.org/. (Accessed: 10th December 2017)
26. Leonidou, C. N., Skarmeas, D. & Saridakis, C. Ethics, Sustainability, and Culture: A Review and Directions for Research. in *Advances in Global Marketing* 471–517 (Springer, Cham, 2018). doi:10.1007/978-3-319-61385-7_19
27. US EPA, O. Toxicity ForeCaster (ToxCast™) Data. *US EPA* (2015). Available at: https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data. (Accessed: 11th December 2017)
28. Toxnet. *Toxnet: Toxicology Data Network* Available at: http://toxnet.nlm.nih.gov/.
29. Registered substances - ECHA. Available at: https://echa.europa.eu/information-on-chemicals/registered-substances. (Accessed: 11th December 2017)
30. Craig, P. *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* **337,** a1655 (2008).
31. Moore, G. *et al.* Process evaluation in complex public health intervention studies: the need for guidance. *J Epidemiol Community Health* **68,** 101–102 (2014).
32. Voutchkova, A. M., Ferris, L. A., Zimmerman, J. B. & Anastas, P. T. Toward molecular design for hazard reduction—fundamental relationships between chemical properties and toxicity. *Tetrahedron* **66,** 1031–1039 (2010).
33. Voutchkova, A. M., Osimitz, T. G. & Anastas, P. T. Toward a Comprehensive Molecular Design Framework for Reduced Hazard. *Chem. Rev.* **110,** 5845–5882 (2010).
34. Faulkner, D. *et al.* CHAPTER 3:Tools for Green Molecular Design to Reduce Toxicological Risk. in *Computational Systems Pharmacology and Toxicology* 36–59 (2017). doi:10.1039/9781782623731-00036
35. Hoepfner, D. *et al.* High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiol. Res.* **169,** 107–120 (2014).

36. Costanzo, M. *et al.* The Genetic Landscape of a Cell. *Science* **327,** 425–431 (2010).
37. Giaever, G. & Nislow, C. The Yeast Deletion Collection: A Decade of Functional Genomics. *Genetics* **197,** 451–465 (2014).
38. Chen, M., Zhang, M., Borlak, J. & Tong, W. A Decade of Toxicogenomic Research and Its Contribution to Toxicological Science. *Toxicol. Sci.* **130,** 217–228 (2012).
39. Fromowitz, M. *et al.* Bone marrow genotoxicity of 2,5-dimethylfuran, a green biofuel candidate. *Environ. Mol. Mutagen.* **53,** 488–491 (2012).
40. Krewski, D. *et al.* Toxicity Testing in the 21st Century: A Vision and a Strategy. *J. Toxicol. Environ. Health Part B* **13,** 51–138 (2010).
41. Nishimura, M., Ueda, N. & Naito, S. Effects of Dimethyl Sulfoxide on the Gene Induction of Cytochrome P450 Isoforms, UGT-Dependent Glucuronosyl Transferase Isoforms, and ABCB1 in Primary Culture of Human Hepatocytes. *Biol. Pharm. Bull.* **26,** 1052–1056 (2003).
42. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588 (2014).

Chapter 2
Tools for Green Molecular Design to Reduce Toxicological Risk

Introduction

The green chemistry design philosophy has proven useful in both academic and business settings[1,2], but some aspects have yet to be fully exploited during the earliest stages of chemical design. The fourth principle of green chemistry, which states chemicals should be "benign by design," is possibly the most difficult to achieve, because it relies on the ability to predict the behavior of a given compound in a biological system, including the degree and rate at which a chemical substance enters the systemic circulation and/or is present at the site of physiological or biochemical activity.

The aim of this chapter is to illustrate in practical terms how some of these tools might be used by bench chemists working in academic or government research settings to design molecules that are less bioavailable and less toxic. For those already comfortable with the fields of molecular biology or toxicology, more comprehensive resources are available,[3] but for those without such backgrounds, this chapter provides an introduction to toxicological hazard reduction via rational molecular design. The chapter begins with a survey of currently available tools and a broad discussion of their attributes. Next, a case study using a panel of test molecules to is presented to demonstrate the utility and applicability of these tools. Several practical workflows are presented to help guide chemists in the early stages of molecule design, and then the chapter concludes with a meditation on the nature of the ideal green molecular design tool.

An Introduction to Green Molecular Design

Voutchkova and colleagues[4] have discussed the three fundamental requirements for chemical toxicity in a living system: (1) there must be exposure of the chemical to the living system, (2) the chemical substance must be bioavailable (the degree and rate at which a chemical substance enters the systemic circulation and/or is present at the site of physiological or biochemical activity) by the route of exposure, and (3) the chemical and/or metabolites must be capable of directly or indirectly causing an alteration (initiating event) which leads to an adverse outcome. The physicochemical properties of the compound(s) play a large role in these factors. As an example, reducing or blocking potential bioavailability through chemical design has been shown to be a successful process in GMD[4].

Fortunately for chemists, the field of computational toxicology has grown tremendously in recent years in response to several initiatives to anticipate and reduce toxic hazard, improve and accelerate toxicity testing of new chemicals, and reduce the use of animals in the testing process.[5,6] While the field of computational toxicology owes its genesis to the fields of medicinal chemistry and drug development, both new and existing platforms can be harnessed to evaluate industrial chemicals as well. Importantly, in June 2016, the Frank R. Lautenberg Chemical Safety for the 21st Century Act was signed into law, which overhauls the United States' 40-year-old statute governing chemicals, The Toxic Substances Control Act. In summary, the law requires the U.S. EPA to ensure that no chemical in the U.S. commerce poses an unreasonable risk to human health or the

environment. The law also aims to reduce the use of laboratory animals in toxicology studies, and to develop, validate, and use reliable alternatives to animal studies. The alternatives, already in process, include computational toxicology and high-throughput cell-based assays. In addition, new technologies such as organ-on-a-chip assays will also be validated and used, which in several cases allows human endpoints to be incorporated into both research and regulatory efforts. These efforts will eventually create a large transparent information source (frequently called "Big Data") that can be harnessed into the GMD process. Data transparency is the key to these efforts, and it is anticipated that information will be available to create new and/or update existing platforms expanding past the chemical space of pharmaceutical compounds. In an effort to propose a new innovative tool for GMD, the authors have analyzed several on-line and subscription based tools involving chemical design and human health and environmental toxicological endpoints. Several of these tools are discussed in this chapter and a brief case study is presented using three tools individually. Based on these assessments, a new proposed GMD tool is outlined.

Physiochemical, Genotoxicity, and Blood Brain Barrier Passage Properties of Chemicals

The traditional model used by medicinal chemists, pharmacologists, and toxicologists to understand how xenobiotics (compounds not naturally produced by the body) enter and leave the body considers the Absorption, Metabolism, Distribution, Excretion, and Toxicity (ADMET) of the compound. Although remarkable progress has been made in recent years regarding the development of computational tools for predicting toxicological endpoints[7], accurately anticipating toxicological hazard remains a significant challenge.[8] Currently the most effective means to proactively reduce toxicological risk of environmental or industrial chemicals is to design molecules that are not readily absorbed by a biological system.[3,4] An additional factor is to redesign structural motifs to avoid metabolism of compounds into more toxic intermediates.[9,10]

A review by Raunio describes *in silico* procedures for predicting metabolic processes.[11] In drug research, medicinal chemists have developed models to predict which physicochemical properties make a molecule more "drug-like": that is, likely to be absorbed into the body, to be distributed within the body to certain targets, and to produce an effect.[12] Those same principles can provide guidance to design molecules that are less likely to interact with biological targets. These guiding principles are primarily concerned with the likelihood that a compound will be absorbed if it is administered orally, but they overlap significantly with the rules of dermal and respiratory absorption as well. Generally speaking, the rules that govern the likelihood of a compound crossing the cellular membrane are consistent across routes of exposure. *Lipinski's rules* are a commonly used set of five properties that are used to predict whether a compound is likely to be readily absorbed through oral ingestion. The rules are as follows: chemicals with (1) more than five hydrogen bond donors, (2) more than 10 hydrogen bond acceptors, (3) a molecular weight greater than 500 Da, or (4) a logP value (sometimes called $\log K_{ow}$) greater than 5 are unlikely to be well absorbed, unless the compound is (5) a substrate for a biological transporter, in which case can be an exception to the previous rules.[13] A comparable analysis by Veber and colleagues similarly found that compounds with more than 10 rotatable bonds, more than 12 total hydrogen bond acceptors and

donors, or a polar surface area greater than 140 Å were unlikely to be orally bioavailable in rats.[14]

While numerous properties are instrumental in the absorption of exogenous compounds, lipophilicity, charge, similarity to endogenous substances, blood-to gas partition molecular weight, and polar surface area appear to be of the greatest value when designing molecules.[3,4] Therefore, molecules that are large, hydrophilic, charged at neutral pH, and that possess a large polar surface area are not readily absorbed in the GI tract. If the molecule of interest is not a substrate for a one of the body's many biological transport proteins, it may be actively transported into cells, and possibly distributed to the rest of the body. More complex and sophisticated models have been developed for predicting intestinal absorption based on *in vitro* data collected in the MCDK or Caco-2 cell lines, which are used as models for intestinal absorption.[15,16] MCDK and Caco-2 permeability predictions are not included in many computational platforms but are notable when they appear, as they provide some direct suggestions about bioavailability. Given the importance of genetic toxicity in the evaluation of commercial compounds, computational tools that predict whether a chemical has mutagenic potential are common.[10,17] The Ames test is an *in vitro* test which uses the bacterium *S. typhimurium* to determine the potential mutagenicity of chemicals.[18–20] A positive Ames result should prompt reconsideration of structural features (such as reactive nucleophiles) likely to elicit mutagenic activity, as well as additional chemical biotransformations that can change mutagenic potential. Currently, assessing the potential mutagenicity of metabolites is limited to specific computational platforms and in most cases, potential metabolite structures have to be assessed separately. Knowing if a chemical can cross the blood-brain barrier (BBB) is a significant aspect of drug design. The highly selective permeability of the BBB can make it challenging to deliver some drugs to the brain as intended, but it can also allow other compounds through that are not meant to affect the central nervous system (CNS), resulting in undesirable effects.[21–23] Predictive models for BBB permeability use lipophilicity, polar surface area, and whether the compound is a substrate for specific transporters to identify compounds that are likely to cross the BBB and potentially interact with neurological pathways. However, computational predictions of BBB permeability must be interpreted cautiously, because the BBB is a complex membrane that is difficult to model.[24,25] The key knowledge required for a chemist wishing to use such predictive toxicology tools is to understand which molecular interactions and characteristics that are most benign or worrisome for a physiological system - concepts which are central to the design of medicinal compounds.

Tools for Green Molecular Design

Currently, several types of tools exist to predict the physicohemical properties and potential toxic effects of compounds; they may be broadly grouped into three categories: Expert Systems, Decision Trees, and (Quantitative) Structure-Activity Relationships (QSARs).

Expert Systems

Expert Systems use known relationships between chemical structures and toxicological outcomes to build rule-based predictive systems based on Structure-Activity Relationships (SARs).[26,27] Assembly of an expert system requires development and curation of a large database of "toxicophores" – specific functional groups or fragments of molecular structures known to cause toxicity – and implementing a set of rules that connects known toxins and toxicophores to appropriate toxic endpoints. If a user searches for a chemical that isn't in the database, expert systems may employ "read-across," a technique which predicts hazard based on comparison to known toxins or toxicophores with similar structures.[28] It is anticipated that as new data become available, expert systems will be continually updated and may provide a more rapid path to validation of computational toxicology tools. Derek Nexus is an extensively used example of an expert system[29] and a detailed analysis of the system follows.

Derek Nexus is a proprietary reasoning-based expert system for the prediction of toxicity. Knowledge-based expert systems can be broadly placed into two categories: rule-based and reasoning based, with Derek Nexus being the latter. A rule-based system relies upon a series of rules that use "IF" and "THEN" statements to present the user with a prediction and the justification for it. However, there are limitations regarding the complexity of rules than can be encoded and the extent to which uncertainty can be handled thus limiting moderation of the outcome from a model. Reasoning based systems are capable of handling uncertainty to make predictions that take account of the interaction between rules which may themselves be imprecise.[1] The interactions between rules which agree or disagree result in a prediction being strengthened, weakened, overturned or contradicted,[2] and users are given information about confidence in the prediction, see Figure 1.

Derek Nexus makes predictions by using alerts which have been developed using appropriate data sources. The alerts are described by "patterns", or "Markush structures", which define the scope of the alert and in combination with reasoning rules allow Derek Nexus to make a prediction for or against toxicity for a given query compound and to advise on the level of confidence in it. They form a knowledge base covering over 50 toxicological endpoints including mutagenicity, carcinogenicity, teratogenicity and skin irritation. Each alert covers a specific area of chemical space by describing a chemical substructure, often referred to as a toxicophore, which is believed to be responsible for inducing a specific toxicological outcome. An alert often contains mitigating features or exclusions to ensure it is triggered only by the relevant toxicophore in the right structural environment.

Alerts are written by studying multiple data sources. They include toxicity data from relevant assays and chemical data such as reactivity as well as information about the biological mechanism of action and any mechanistic data. Mechanistic information is important since this is what drives toxicity and the information is included in the alert comments, which provide an expert summary of the molecular initiating event (MIE) and key events (KEs) linking it to a toxicological endpoint, thus describing in outline the adverse outcome pathway (AOP) for the alert. Metabolism data are taken into account, if

available, since this enables prediction for a compound which is not directly toxic but is metabolized to an active species, Figure 2.

In addition, rules are written about the species for which the alert is relevant and factors which influence the manifestation of toxicity such as physico-chemical properties (e.g. fat-water partition coefficients). These rules allow the inference engine to modulate the prediction for a query containing an alert and the result can range from a high expectation of activity all the way down to an explicit prediction of inactivity (as distinct from merely the absence of a prediction of activity). The use of human expertise is the key difference between a knowledge based system and a machine-learnt statistical system such as Sarah Nexus.[8] Like a human expert, a knowledge based system can assess the level of confidence in predictions from an alert for a given type of query structure on the basis of a wide range of relevant information such as chemical reactivity, characteristics of an assay or an understanding of the likely mechanism, whereas a machine-learnt system bases its overall call only on a statistical analysis of the training data and the descriptors employed. The derivation of each alert in Derek Nexus is described in the alert commentary along with supporting references and example compounds to provide a transparent prediction. Historically, Derek Nexus alerts have been restricted to qualitative prediction of the hazards posed by a chemical. Advances in both human and machine-learnt knowledge mean that it is now possible for alerts to make at least semi-quantitative predictions. Derek Nexus can predict the skin sensitization potency class for a query compound in the local lymph node assay (LLNA), EC3, based on experimental data observed for the nearest neighbors, weighted by the Tanimoto similarity score, and selected from a set of compounds that exclusively fire the same alert as the query compound. The use of structural alerts for the prediction of toxicity is widely understood[3,9] and has found application in many settings.[10–14] Recent regulatory acceptance of *in silico* toxicity predictions is a significant milestone in the use of (Q)SAR for risk assessment. The ICH M7 guidelines allow the acceptance of negative computer predictions for the genotoxic risk assessment of low-level pharmaceutical impurities, provided that they come from two complementary (Q)SAR methodologies; one expert rule-based and one statistically based, for example Derek Nexus can be used in conjunction with Sarah Nexus. This dual approach reduces the risk that mutagens will be missed. In many cases, the development of new alerts and rules is supported by data sharing initiatives in which organizations pool relevant data. These consortia share knowledge and data with the aim of advancing scientific knowledge as opposed to a specific product. ICH M7 has prompted new data and knowledge sharing initiatives for which Derek Nexus has been utilized as a tool.[12] Derek Nexus has also been used as part of an integrated testing strategy (ITS) to predict for compounds outside the applicability domain of *in chemico/in vitro* assays[13] and as a tool to develop novel methods for making negative predictions for mutagenicity.[14] Derek Nexus is an example of computer prediction as a fast and green method with which to assess the toxicity of chemicals. The advent of stricter regulations on animal testing will increase the importance of computer models as predictive tools. As the models continue to improve and with increased sharing of data between companies, their predictions can be expected to become increasingly accurate and their use more commonplace.

Decision Trees

Decision trees use a series of Yes or No questions to classify and prioritize compounds based on their structural properties. The data for developing decision trees comes from mining the available literature and data of existing compounds and subsequently categorizing chemicals based on structural features and known toxic outcomes. An example is the Cramer Classification scheme, a common decision tree for ranking chemicals in terms of their expected oral toxicity. It consists of 33 questions that place chemicals in one of three classes: Class I - low oral toxicity, Class 2 - moderate oral toxicity, and Class 3 - high oral toxicity. The original decision tree was developed by Cramer and Ford[30] in the late seventies, but modifications were proposed in 2002 to improve the accuracy of classifications. Although some consider the Cramer scheme in need of additional revision,[30–33] this decision tree has proven useful enough to be included in many computational tools such as *Toxtree* and the *OECD QSAR Toolbox*. Verhaar's scheme is yet another decision tree, developed to predict the likelihood of a chemical causing environmental toxicity.[34]

QSAR Tools

QSARs (sometimes called quantitative structure-property relationships (QSPR)) model the relationship between a chemical structure and a specific biological endpoint.[26] QSAR processes are discussed in detail in another chapter in this book. There are many ways to create a QSAR, but the basic strategy is as follows: starting with a "training set" of molecules (compounds with known positive and negative values for an endpoint of interest), a variety of mathematical techniques are used. These include: multiple linear regression, partial least squares, neural networks, logistic regression, or linear discriminant analysis. Yee and Wei provide an excellent overview of various methods in a chapter on Statistical Modeling of Molecular Descriptors in QSAR/QSPR[35] to explain the variation in molecules in accordance with their values for that endpoint. QSARs are then validated with molecules that have known values for the endpoint, but which are not in the training set. The best computational tools provide information about the training sets used to build their QSARs, so that it is possible to determine how accurate their predictions are for a given molecule. If a molecule is very dissimilar from the compounds used in the training set, the QSAR will not generate a statistically meaningful prediction for that molecule. The chemical space for which a QSAR can provide statistically meaningful output is referred to as the "applicability domain." Examples of tools that use QSARs are programs like *ACD Percepta, ADMET Predictor, Derek, Medchem Designer, The OECD QSAR Toolbox, Mobyle@RPBS, and QikProp*.

Representative Tools

To assess the utility and accessibility of computational tools currently available for molecular designers, a representative list of available computational tools was generated using click2drug.org, an online "directory of computer-aided drug design tools" maintained by the Swiss Institute of Bioinformatics.[36] The list included single-function QSARs and collections of ADMET prediction models, with a mix of downloadable software, databases, and web-based platforms. The list includes the tools that the authors considered most useful and accommodating to general trends for the overlapping fields of

computational toxicology, computational chemistry, and computational medicinal chemistry.

ACD Percepta[37]
This medicinal chemistry software is distributed by Advanced Chemistry Development, Inc.[37] The program accepts single and batch .sdf or .mol files and also allows users to draw structures directly in the software. The outputs include basic physicochemical parameters (logP, H-donors/acceptors, rotational bonds, rings, Lipinski's Rules violations, solubility) as well as ADME (Caco-2 permeability, Plasma Protein Binding, CNS Penetration), and some drug safety information (some CYP inhibition tests, Ames, hERG). The program provides color-coded interpretations of those outputs and indicates on sliding scales how drug-like a molecule is. Navigation is easy, and different compounds can be compared in convenient tables or tabular format. Saving and sharing data is simple through export into either .pdf or .csv format, and the user interface is aesthetically pleasing and easy to work with. Of particular utility is the *Structure Design Engine* module, as it allows the chemist to edit a molecule in a drawing window and observe how structural modifications affect predicted physicochemical and ADME properties in real time: as one modifies the molecule, sliders move and alerts appear based on structural changes. Additionally, this module is capable of proposing a set of analogs for a given molecule based on a set of desired physicochemical parameters – a feature which greatly simplifies the "iteration" phase of the workflows. It is worth noting that this (like most of the programs discussed herein) is intended for the development of pharmaceuticals, not industrial chemicals, so appropriate care must be taken when interpreting the results generated by *ACD Percepta,* as the molecule of interest may fall outside the training sets used to develop the predictive algorithms. For chemists looking to improve the ADMET profile for a compound, this software offers most of what is needed.

ADMET Predictor[38]
A medicinal chemistry program distributed by Simulations Plus, this program contains a large number of QSARs capable of predicting ADME values, a plethora of toxicity endpoints, and numerous metabolism parameters including CYP metabolism kinetics.[45] Chemicals may be input as SMILES, .sdf, .rdf, or mdl files, and may be entered in single or batch format. Predictions are returned in the form of a table, and hovering over a prediction produces an explanatory tooltip to ease interpretation. ADMET Predictor includes several "summary" toxicity prediction models, which assign rankings and codes indicating the specific toxicity concern(s) to chemicals based on the data outputs from certain physicochemical or QSAR models. These indicate if, for example, a compound is predicted to cause acute toxicity or carcinogenicity in rodents, to cause hepatotoxicity in humans, or to be Ames positive. Aside from the physicochemical property predictions, these summary predictions are likely to be of the most interest to chemists because they greatly simplify data analysis. Prediction data is easily shared in the form of .tsv files. Like ACD Percepta, ADMET Predictor is intended for use in the development of pharmaceuticals, so predictions for the properties of industrial chemicals must be considered thoughtfully. The .tsv output files include "applicability of domain" statistics that indicate if the submitted molecule falls within the scope of the predictive model (based on the chemical structures used in the model training set). This feature provides a

degree of confidence of the prediction, and incorporates a degree of transparency into the software.

Medchem Designer[39]
A chemical drawing program coupled with ADMET Predictor or available as a free standalone program, Medchem Designer, features basic predictions of drug-likeness and bioavailability of molecules.[39] Structures may be drawn or uploaded in SMILES, .mol, and .sdf formats. Predictions appear in table format below the workspace and are very easily exported to Excel. The interface is simple, clean, and can be learned quickly, making it ideal for the novice user. The program uses the same codes to indicate predicted toxicity as are used in ADMET Predictor. Although Medchem Designer has far fewer QSARs than ADMET Predictor, it is sufficient for collecting enough physicochemical and ADMET parameters to make iterative molecular design adjustments. One of the useful features of this program is the Optical Structure Recognition tool, which allows the user to draw a box around an onscreen chemical structure and import that structure into the program. This can be helpful when the .mol or .sdf file of a compound is not immediately available for import.

Lhasa Derek and Meteor Suites[40,41]
Lhasa Limited distributes the *Nexus* software suite through which the Derek and Meteor products are licensed. The Derek platform is detailed earlier in 3.3.1.

As mentioned earlier, Derek provides expert, knowledge-based toxicity predictions using the Lhasa Knowledge Base, a curated database derived from literature and proprietary sources.[30,47] Structures may be drawn in the workspace, imported singly, or imported in batch format from .mol files, SMILES strings, or any delimited structure file. Derek is capable of providing predictions for charged or metal-containing compounds – a feat few medicinal chemistry tools are capable of. Outputs are tabular and the program provides a list of species and endpoints, including both the plausibility of toxicity and a note for which structure features triggered the alert. This data can then be easily exported to a .tsv file for viewing in Excel. Toxicity plausibility predictions are classified according to likelihood from "IMPOSSIBLE" to "CERTAIN." If a structural feature does trigger an alert, the software highlights the offending moiety and offers detailed reasoning behind the alert, which may be used to inform design decisions. *Derek* also includes documentation to aid interpretation of predicted toxicophores or outcomes.

Meteor provides metabolism predictions in the form of tree diagrams, including the evidence-based rule for the reaction, a score indicating the likelihood of it occurring, and any intermediates that may be generated along the way.[29,41] Toxic products and intermediates are indicated, as are moieties on the molecule that are most likely to be sites of metabolism. Clicking on a metabolite opens a series of tabs containing information about the series of transformations which lead to it, details about biotransformations, and the "Nearest Neighbors", chemical transformations documented in the literature which most closely resemble the reaction of interest. These predictions are helpful for identifying toxic biotransformation products and can be used to help make the necessary structural modifications to avoid them.

Qikprop[42]
Qikprop is a predictive ADME module within the *Maestro* suite produced by Schrödinger, LLC.[42] It accepts singular .sdf, .mol, and .pdb files, and batch files may be imported in a variety of formats. A variety of output formats, including .csv, are also available. The program offers several predictors for ADME including CNS penetration, QPPCAco (permeability across gut-blood barrier), QPPMDCK (kidney permeability), human oral absorption, Lipinski's Rule of Five, and JM (predicted maximum transdermal transport rate). Qikprop has an option to rank compounds on the basis of how drug-like they are. There is also an option that simplifies molecule-to-molecule comparisons by indicating how many of the predicted ADME values for the molecule fall outside the 95% range of similar values for known drugs – this is known as the "stars" mechanic, where more stars indicates a less drug-like molecule. Like ADMET Predictor, this program includes an array of predictive tools. Additionally, the documentation is extensive and thorough, and provides excellent information on the methods and training sets used to build the predictive models, so domain of applicability questions are easily answered.

OECD QSAR Toolbox[43]
Produced by the Organization for Economic Cooperation and Development (OECD) for evaluating industrial chemicals, this platform includes a blend of decision trees, QSARS, and predictive metabolism modules.[50] Structures may be uploaded as .mol, .sdf, or SMILES files, searched by a name or CAS#, drawn as a structure, and may be uploaded in batch format. Output is in a tree-based layout, indicating violations of various rules and sources of evidence for why the rule violations may contribute to toxicity. Modules for simulating metabolism, oxidation, and hydrolysis are included as well, the latter two of which may be used as indirect measures of environmental persistence. Notably, the software is capable of profiling both the chemical structure entered and the predicted metabolites of that structure in the same run. Attempting to run all decision trees and QSAR predictions on a batch of compounds requires extra time, but can expedite the process of evaluating a set of chemicals. Each output is hot-linked and clicking through links provides information about the QSARS or decision trees used to generate the outputs, their interpretations, and what part of the molecular structure triggered them. If the compound of interest is in the EU database or has a CAS#, it is possible to search for charged compounds or metals – an advantage over medicinal chemistry software suites that do not handle any metals at all.

Toxtree[44]
Toxtree provides basic toxicity evaluations, generally in the form of a binary "Toxic" or "Non-toxic" format or a simple ranked format, usually with three to five hazard rankings.[45] The Toxtree models include the two Cramer's rule sets,[30] Verhaar's schema,[34] the ISS decision trees, structural alerts for the *in vivo* micronucleus assay (a predictor of genetic toxicity), and the binding alerts trees.

Chemaxon Suite[46] *(Marvin Sketch and Metabolizer)*
The *Marvin Sketch* and *Marvin Space* are useful for creating, viewing, and editing chemical structure files. In addition to its drawing function, *Marvin Sketch* also includes models for calculating a variety of physicochemical properties with assessment utility for bioavailability: pKa, logP, logD, aqueous solubility, H-bond Donor/Acceptor, and Polar

surface area. The *Metabolizer* program included in the *Chemaxon Suite* provides metabolism predictions.

Chemicalize[47]
An online tool powered by Chemaxon's predictive algorithms, Chemicalize can be used to generate an array of physicochemical properties and a few drug likeness parameters. The data are easily exported and the layout allows a user to view multiple predictions at the same time. The interface allows the user to quickly modify chemical structures and generate physicochemical parameters.

AIM (Analog Identification Methodology)[48]
Produced and distributed by EPA, this tool accepts CAS#s, chemical names, SMILES, or structural drawings, and returns a .pdf report with links to any information publically available for that compound or its analogs in US or Canadian chemical hazard databases.[48] This can be particularly useful when iterating a compound as well as for gathering information about the chemical space surrounding a compound of interest.

Chemspider[49]
A chemistry search engine with built in physicochemical prediction capabilities. Several options for modeling are available: users may choose between models from ACD, ChemAxon, or EPA's EpiSuite.[49]

Mobyle @RPBS[50]
An online physicochemical profiler managed by the University of Paris Diderot, this modeling tool accepts .sdf or .mol files for individual compounds.[51,52] Predictions can be saved in online user profiles or downloaded as a text file. The various physicochemical property predictive models are nested in a series of menu trees. Mobyle@RPBS provides references to source papers used to program the physicochemical models, which provides transparency to the tool.

Case Study

Improving Hazard Profiles of Fuel Cell Components Using GMD
As an example of how these tools may be put to use, a case study follows in which a collection of compounds with potential as hydrogen storage materials for proton-exchange membrane (PEM) fuel cells[53–56] are evaluated for their potential for human toxicity. The case study is centered on reduction of bioavailability and human hazard. The efficacy and performance of these compounds in PEM fuel cells has not been considered. Eight compounds were chosen for evaluation (Figure 3). The following tools were used to evaluate the compounds of interest: *Lhasa Derek* and *Meteor* Suites, *ADMET Predictor*, *ACD Percepta*.

ACD Percepta and *ADMET Predictor* agreed that all the compounds were predicted to penetrate the BBB and were likely to be orally bioavailable due to favorable LogP values (roughly between 2 and 5), low polarized surface area (all less than 16 $Å^2$), and small size (all less than 400 Da). The *Derek* analysis highlighted *N*-ethylcarbazole as a CERTAIN Ames positive compound, and gave an EQUIVOCAL rating to *N*-

ethyldodecahydrocarbazole and dodecahydrocarbazole for causing phospholipidosis (a tissue-specific lipid metabolism disorder) in multiple species including humans. Indoline was selected for further iteration as it was the only compound among the initial 8 that was without negative predicted effects, was predicted to be Ames negative, and did not inhibit any CYPs. Using indoline as the base structure, a series of molecules was generated using the additions of alcohol, sulfite, or sulfate groups at various positions on the indoline core to generate less toxic and less bioavailable compounds. These groups and sites were chosen after interpreting the metabolism of indoline in humans as predicted by the *Meteor* suite, and with consideration of the metabolism of indole.[57] The second iteration of compounds is indicated in Figure 4. Running the compounds through *ADMET Predictor*, *ACD Profiler*, and *Derek* a second time revealed an improved overall profile for hazard and bioavailability for most of the compounds.

The *Derek* evaluation of the iterated compounds found, no structural alerts for mutagenicity at EQUIVOCAL or PROBABLE levels, although a four of the ten compounds did trigger 12% EC3 predictions (indicating the compound might be a weak skin sensitizer), each compound only had a single EC3 prediction to provide evidence. The second pass at the *ACD Profiler* indicated reduced GI absorption, fewer CYP interactions, reduced BBB absorption, and "hydrophilic" ratings for half of the compounds. An iterative run through *ADMET Predictor* yielded an improved profile for many of the compounds, although some were still predicted to cross the BBB or cause genotoxicity. Compound D, however, had a very promising profile: it was predicted to be poorly permeable to the BBB and the GI, charged at physiological pH, and was hydrophilic, suggesting that it would be cleared quickly. While this is not a complete re-design of the selected compound, it illustrates the value of the predictive tools selected for GMD.

Workflows

Based on an assessment of available tools, 3 workflows are outlined (Figure 5.) to serve as guides for applying a series of predictive tools to improve molecule design. The first step of each workflow is to survey the available literature and databases to ascertain the extent to which information is available for a compound of interest. The following steps depend on both the availability of resources and the chemist's degree of familiarity with biological systems. In all cases, the main purposes of using these tools are twofold:

1. Determine ADME and physicochemical properties of a compound so it can be modified to be less bioavailable
2. Identify known toxicophores and structures of concern so that they may be removed or modified.

Programs that predict metabolism are helpful for identifying potential modifications for successive iterations of molecular design. By understanding how the body is likely to attempt to break down and remove a chemical, it becomes clear which design choices will accelerate the passage of the molecule through the body if not prevent it from being taken up entirely.

It must be noted that these workflows do not account for how such modifications may impact the function of the compounds under development, and we acknowledge that the intended function of some compounds may be at odds with attempts to reduce its activity in biological systems. However, while a workflow may not necessarily provide an obvious solution to balancing functionality and hazard, it will necessarily provide insight into the physicochemical properties of a molecule, which may yet prove useful to the chemist.

Subscription-based workflow

The subscription-based workflow is intended for chemists aiming to collect a trove of information about their molecules of interest and who either have access to toxicology resources or some training in the discipline. This workflow is recommended for the design of chemicals produced and distributed in great volume or whose end use is likely to result in high levels of human exposure. Each of the programs has a high degree of utility and usability as well as other advantages such as more streamlined user interfaces, readily available technical support, and access to proprietary databases. However, since these tools are designed for drug developers, interpretation of the voluminous data sets requires technical knowledge in molecular biology and medicinal chemistry, which may not be readily available. *Derek* is used to identify known or predicted structural alerts, which may disqualify candidate compounds, especially if they are classified as PROBABLE or CERTAIN. *ACD Percepta, ADMET Predictor,* or *QikProp* may be used to collect data for ADME and toxicological risk. These programs provide a preponderance of data that can be used to eliminate compounds with unfavorably high likelihood of absorption, predicted inhibition of metabolic enzymes, or predicted risk of genetic toxicity. *Meteor* provides predictive metabolism and identification of potentially toxic metabolites or intermediates. Metabolism predictions are combined with ADME data from *ACD Percepta, ADMET Predictor,* or *QikProp* to design the next iteration of molecules, which share the features of likely metabolites so that they are less likely to be absorbed, and more likely to be excreted quickly.

Open-source workflow

The Open source workflow balances data volume with cost of effort, and is intended for chemists with an interest in toxicology but fewer financial resources. This workflow recommended for chemicals that will be produced at intermediate volumes and have moderate opportunity for human exposure. Open source options often have fewer features and provide less information about any given endpoint so more programs are needed to collect the same amount of data. In this workflow, the primary emphasis for toxicological risk is placed on the *OECD QSAR Toolbox*. Any compounds which trigger structural alerts in *OECD QSAR Toolbox* should be considered for removal, and kept only of the alert can be considered spurious (due to an erroneous SAR) or if their ADME profile is particularly good. Both *Medchem Designer,* and *OECD QSAR Toolbox* are used to predict physicochemical and ADME parameters. *Chemaxon Metabolizer* is used to predict metabolism, which can be combined with ADME data from *Medchem Designer* and *OECD QSAR Toolbox* to generate the next iteration of structure designs.

Minimalist workflow

The minimalist workflow is most useful for compounds that have highly specialized industrial uses, and which are not likely to be produced in great volume or distributed extensively. The goal of this workflow is to identify structural elements that pose a significant and likely hazard, and to provide basic information about the bioavailability and physicochemical properties of the molecule. Any structural features that trip toxicity alerts in *Toxtree or OECD QSAR Toolbox*, should be either removed or modified considerably. *Medchem Designer* is used to predict logP and the number of Lipinski's Rule of Five violations, indicating likely oral bioavailability. *Chemicalize* and *Mobyle @RPBS* provide additional physicochemical parameters and *Chemicalize* offers suggestions for similar structures, which may be useful for iteration purposes.

Conclusions

An Ideal Green Molecular Design Tool for Chemists

The ideal GMD tool for chemists of varying degrees of experience and knowledge, particularly of toxicological principles, would be an open-access front end with the ability to incorporate subscription-based tools for specific functionalities. This is typically referred to as a "freemium" model. The premium or subscription-based tools would depend on the institution or business where the chemist works or studies. In this chapter, we have highlighted tools available at the University of California, Berkeley. The GMD freemium tool would allow all possible entries of chemical structures, predict species-centric metabolites, recommend multiple compounds with similar structures, and allow instant analysis of all compounds for structural alerts, physicochemical properties, and hazard identification. Based on these evaluations, structural motifs could be modified in the tool to provide multiple options for GMD based on the hazard profiles predicted. It is anticipated that several screens are/or will be available for rapid evaluation of hazard categories. Ultimately, the predicted results and screening results could be used to create local QSAR or QSTR models similar to models based on close structural analogues. These models could be used along with the initial structural alert and physicochemical property predictions. This process would supercede the use of global QSAR models where chemical space and applicability domain can become an issue.

Notably absent from most of these tools (and largely from this review) are methods for designing compounds with reduced environmental impact; there are two primary reasons for this: 1) The greatest effort has been spent developing tools to predict how chemicals may adversely affect human health, and 2) Quantitatively measuring adverse effects on the environment is a complex and challenging task. Tools like *PBT Profiler*[58], *EcoSAR*[59], and *EPIsuite*[60] predict a few endpoints of environmental consequence such as persistence, bioaccumulation, and toxicity values for fish, water fleas, and algae; and the Verhaar scheme[3528] functions similarly to the Cramer scheme, but for evaluating and categorizing environmental pollutants. These tools are helpful for filtering out some of the chemicals with the greatest potential for environmental hazard, but they are not nearly as developed as medicinal chemistry tools, and they generally lack the sophistication to distinguish between mild and moderate environmental toxicants. Although a number of research groups have published models for predicting or describing the physicochemical characteristics of environmental toxicants[61–63] they, too, are bespoke models that only

characterize one or a few environmental health endpoints at a time. A more comprehensive toolbox needs to be developed with expanded capacity for predicting the environmental fate and effects of a large number of chemical classes. The nearest solution presently available is the *OECD QSAR Toolbox* which includes numerous environmental health parameters, but the endpoints examined are far from exhaustive.

The green molecular design of chemicals involves the ability to use several tools both in chemical design and toxicological evaluation. There is a need for a new and innovative tools that allow chemists the ability to add proposed chemical structures in a variety of formats and to automatically calculate and predict key human and environmental health endpoints to aid in new chemical design. A proposed "freemium" model that provides an on-line interface with the ability to be coupled with various open-access and subscription based tools would be ideal. This ideal design tool could be used in academic, government, and industry labs and would provide an exceptional educational model for chemistry students, reinforcing a key principle of GMD: safety and efficacy need not be mutually exclusive.

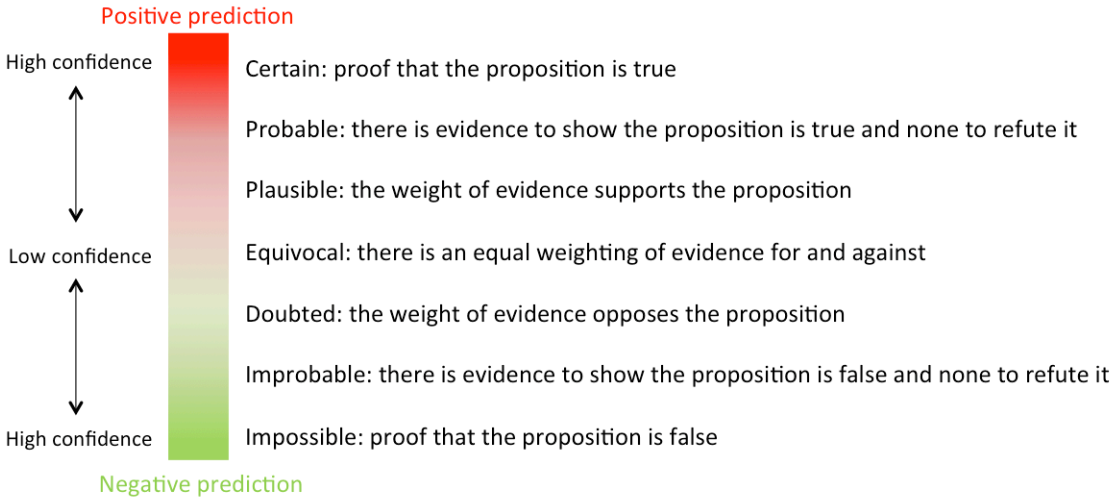


Figure 2.1: Derek Nexus can assess the level of confidence in a prediction To expressed by a likelihood level (from certain to impossible). The darker shades on the colour scale represent the increased confidence of a prediction being correct.

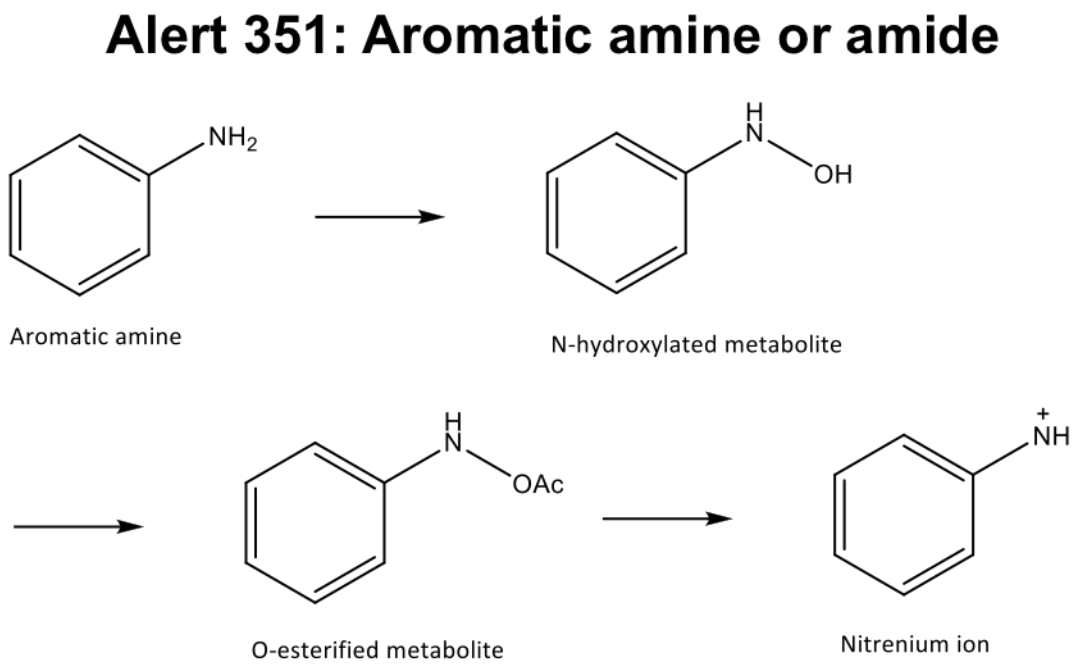


Figure 2.2: Derek Nexus Alert 351 Example Many aromatic amines exhibit mutagenicity in the Ames test. The mechanism of action is generally considered to involve N-hydroxylation, typically mediated by cytochrome P450 1A2, and subsequent O-esterification.[64] The resulting esterified product may then give rise to a reactive nitrenium ion which is capable of binding to cellular nucleophiles such as DNA.

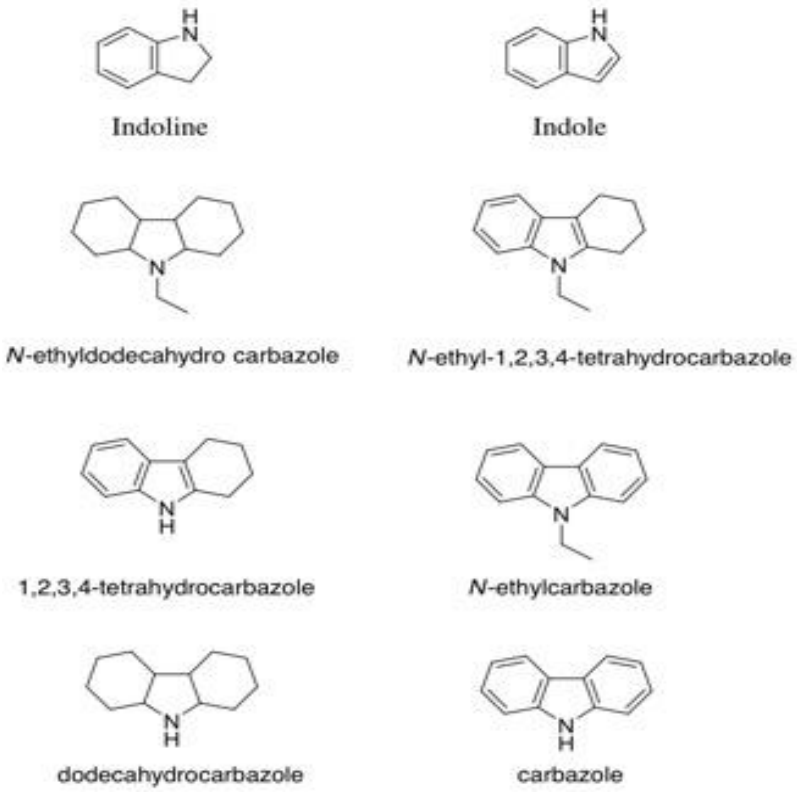


Figure 2.3: Hydrogen Compounds Used in Case Study Hydrogen compounds utilized in case study to evaluate predictive tools and workflow.

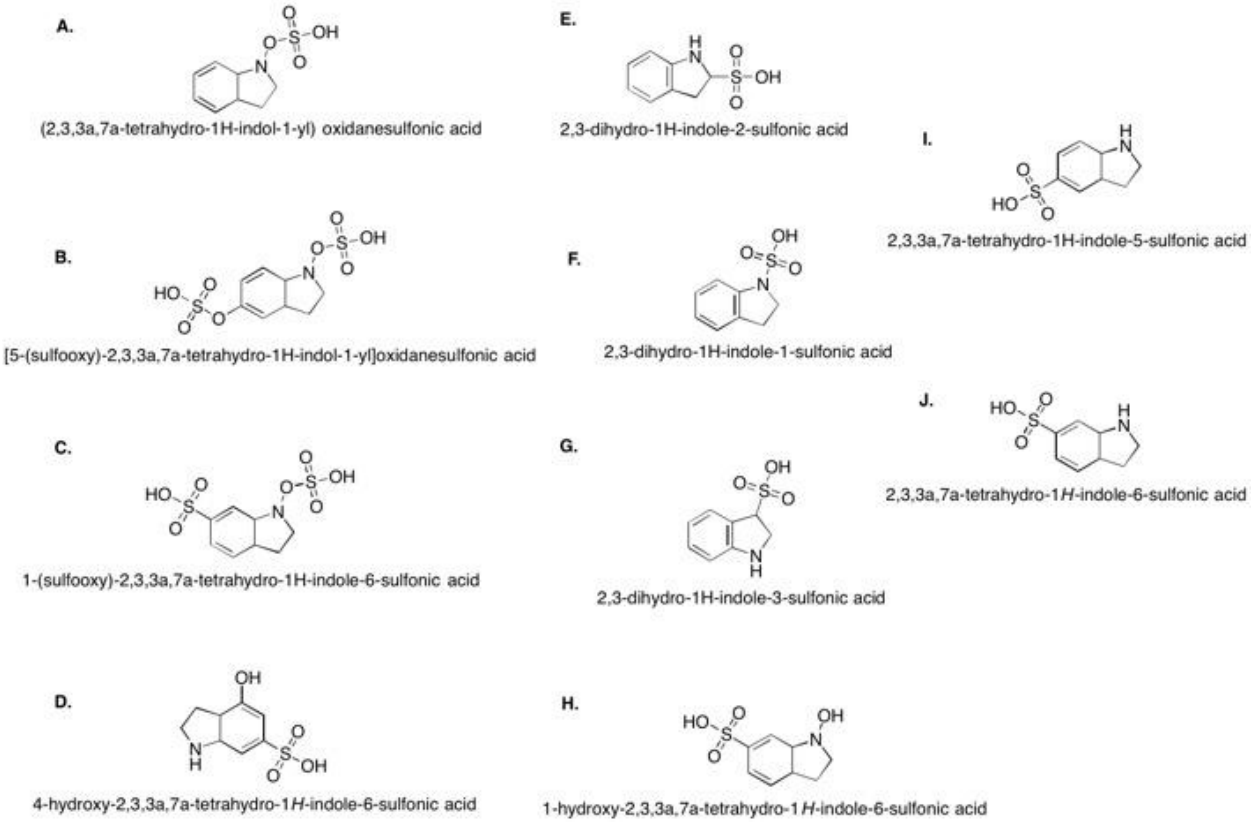


Figure 2.4: Compounds Generated From Second Iteration of Indoline: Compounds generated from second iteration of indoline

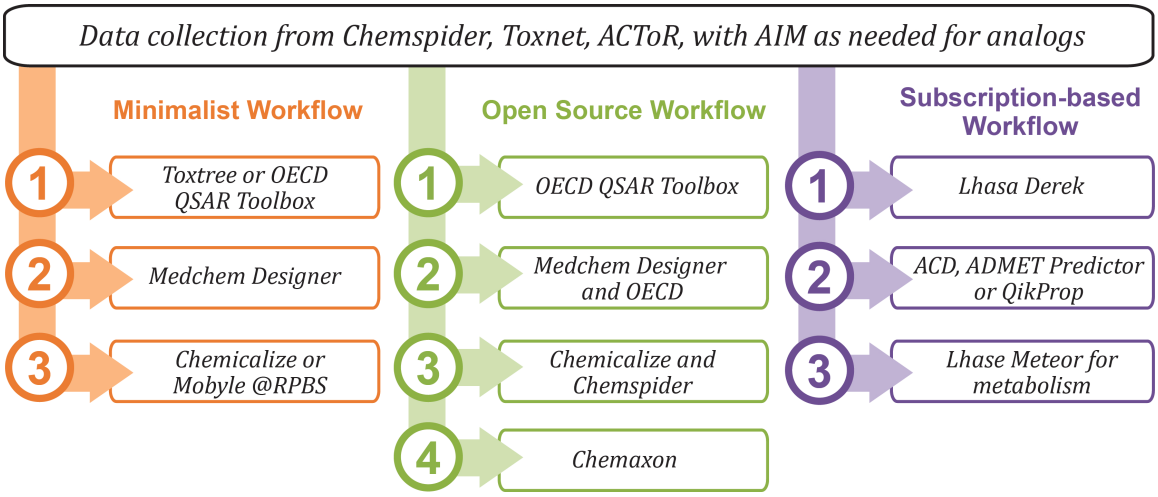


Figure 2.5 Example Workflows for Chemists: Example workflows for chemists.

Sources Cited

1. Dunn, P. J. The importance of green chemistry in process research and development. *Chem. Soc. Rev.* **41,** 1452–1461 (2012).
2. Mulvihill, M. J., Beach, E. S., Zimmerman, J. B. & Anastas, P. T. Green Chemistry and Green Engineering: A Framework for Sustainable Technology Development. *Annu. Rev. Environ. Resour.* **36,** 271–293 (2011).
3. Voutchkova, A. M., Ferris, L. A., Zimmerman, J. B. & Anastas, P. T. Toward molecular design for hazard reduction—fundamental relationships between chemical properties and toxicity. *Tetrahedron* **66,** 1031–1039 (2010).
4. Voutchkova, A. M., Osimitz, T. G. & Anastas, P. T. Toward a Comprehensive Molecular Design Framework for Reduced Hazard. *Chem. Rev.* **110,** 5845–5882 (2010).
5. Andersen, M. E. & Krewski, D. The Vision of Toxicity Testing in the 21st Century: Moving from Discussion to Action. *Toxicol. Sci.* **117,** 17–24 (2010).
6. Krewski, D. *et al.* Toxicity Testing in the 21st Century: A Vision and a Strategy. *J. Toxicol. Environ. Health Part B* **13,** 51–138 (2010).
7. Worth, A. *et al. Alternative methods for regulatory toxicology a state-of-the-art review.* (Publications Office, 2014).
8. Russom, C. L. *et al.* Predicting modes of toxic action from chemical structure: Predicting modes of toxic action from chemical structure. *Environ. Toxicol. Chem.* **32,** 1441–1442 (2013).
9. *Casarett and Doull's toxicology: the basic science of poisons*. (McGraw-Hill, 2008).
10. Bakhtyari, N. G., Raitano, G., Benfenati, E., Martin, T. & Young, D. Comparison of In Silico Models for Prediction of Mutagenicity. *J. Environ. Sci. Health Part C* **31,** 45–66 (2013).
11. Raunio, H. In Silico Toxicology – Non-Testing Methods. *Front. Pharmacol.* **2,** (2011).
12. Hou, T., Wang, J., Zhang, W. & Xu, X. ADME Evaluation in Drug Discovery. 6. Can Oral Bioavailability in Humans Be Effectively Predicted by Simple Molecular Property-Based Rules? *J. Chem. Inf. Model.* **47,** 460–463 (2007).
13. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **64,** 4–17 (1997).
14. Veber, D. F. *et al.* Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *J. Med. Chem.* **45,** 2615–2623 (2002).
15. Irvine, J. D. *et al.* MDCK (Madin-Darby Canine Kidney) Cells: A Tool for Membrane Permeability Screening. *J. Pharm. Sci.* **88,** 28–33 (1999).
16. Ungell, A.-L. B. Caco-2 replace or refine? *Drug Discov. Today Technol.* **1,** 423–430 (2004).
17. Cassano, A. *et al.* Evaluation of QSAR Models for the Prediction of Ames Genotoxicity: A Retrospective Exercise on the Chemical Substances Registered Under the EU REACH Regulation. *J. Environ. Sci. Health Part C* **32,** 273–298 (2014).
18. McCann, J. & Ames, B. N. Detection of carcinogens as mutagens in the Salmonella/microsome test: assay of 300 chemicals: discussion. *Proc. Natl. Acad. Sci.* **73,** 950–954 (1976).
19. Walmsley, R. M. & Billinton, N. How accurate is in vitro prediction of carcinogenicity?: Genotoxicity testing. *Br. J. Pharmacol.* **162,** 1250–1258 (2011).

20. Kirkland, D., Aardema, M., Henderson, L. & Müller, L. Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens. *Mutat. Res. Toxicol. Environ. Mutagen.* **584,** 1–256 (2005).
21. Abbott, N. J. Blood–brain barrier structure and function and the challenges for CNS drug delivery. *J. Inherit. Metab. Dis.* **36,** 437–449 (2013).
22. Aungst, B. J. Absorption Enhancers: Applications and Advances. *AAPS J.* **14,** 10–18 (2012).
23. Obermeier, B., Daneman, R. & Ransohoff, R. M. Development, maintenance and disruption of the blood-brain barrier. *Nat. Med.* **19,** 1584–1596 (2013).
24. Strazielle, N. & Ghersi-Egea, J.-F. Factors affecting delivery of antiviral drugs to the brain. *Rev. Med. Virol.* **15,** 105–133 (2005).
25. Suenderhauf, C., Hammann, F. & Huwyler, J. Computational Prediction of Blood-Brain Barrier Permeability Using Decision Tree Induction. *Molecules* **17,** 10429–10445 (2012).
26. Modi, S., Hughes, M., Garrow, A. & White, A. The value of in silico chemistry in the safety assessment of chemicals in the consumer goods and pharmaceutical industries. *Drug Discov. Today* **17,** 135–142 (2012).
27. Sutter, A. *et al.* Use of in silico systems and expert knowledge for structure-based assessment of potentially mutagenic impurities. *Regul. Toxicol. Pharmacol.* **67,** 39–52 (2013).
28. Low, Y. *et al.* Integrative Chemical–Biological Read-Across Approach for Chemical Hazard Classification. *Chem. Res. Toxicol.* **26,** 1199–1208 (2013).
29. Marchant, C. A., Briggs, K. A. & Long, A. In Silico Tools for Sharing Data and Knowledge on Toxicity and Metabolism: Derek for Windows, Meteor, and Vitic. *Toxicol. Mech. Methods* **18,** 177–187 (2008).
30. Cramer, G. M., Ford, R. A. & Hall, R. L. Estimation of toxic hazard—A decision tree approach. *Food Cosmet. Toxicol.* **16,** 255–276 (1976).
31. Lapenna, S., Worth, A. & Institute for Health and Consumer Protection. *Analysis of the Cramer classification scheme for oral systemic toxicity implications for its implementation in Toxtree.* (Publications Office, 2011).
32. Bhatia, S. *et al.* Comparison of Cramer classification between Toxtree, the OECD QSAR Toolbox and expert judgment. *Regul. Toxicol. Pharmacol.* **71,** 52–62 (2015).
33. Kalkhof, H., Herzler, M., Stahlmann, R. & Gundert-Remy, U. Threshold of toxicological concern values for non-genotoxic effects in industrial chemicals: re-evaluation of the Cramer classification. *Arch. Toxicol.* **86,** 17–25 (2012).
34. Verhaar, H. J. M., van Leeuwen, C. J. & Hermens, J. L. M. Classifying environmental pollutants. *Chemosphere* **25,** 471–491 (1992).
35. Parker, R. E. & Isaacs, N. S. Mechanisms Of Epoxide Reactions. *Chem. Rev.* **59,** 737–799 (1959).
36. Click2Drug. Available at: http://click2drug.org/.
37. *ACD Percepta*. (Advanced Chemistry Development).
38. *ADMET Predictor*. (Simulations Plus Inc.).
39. *Medchem Designer*. (2014).
40. *Derek*. (Lhasa Limited).
41. *Meteor*. (Lhasa Limited).
42. *QikProp*. (Schrodinger).

43. *OECD QSAR Toolbox*. (Organization for Economic Co-operation and Development).
44. *Toxtree*. (Ideaconsult Ltd.).
45. Patlewicz, G., Jeliazkova, N., Safford, R. J., Worth, A. P. & Aleksiev, B. An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR QSAR Environ. Res.* **19,** 495–524 (2008).
46. ChemAxon - Software Solutions and Services for Chemistry & Biology. Available at: https://www.chemaxon.com/. (Accessed: 6th December 2017)
47. Chemicalize. *Chemicalize.org* Available at: http://www.chemicalize.org/.
48. *AIM: Analog Identification Methodology*. (U.S. EPA, Risk Assessment Division).
49. Chemspider. *Chemspider: Search and Share Chemistry* Available at: www.chemspider.com.
50. Mobyle@RPBS. *Mobyle@RPBS* Available at: http://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#welcome.
51. Néron, B. *et al.* Mobyle: a new full web bioinformatics framework. *Bioinformatics* **25,** 3005–3011 (2009).
52. Alland, C. *et al.* RPBS: a web resource for structural bioinformatics. *Nucleic Acids Res.* **33,** W44–W49 (2005).
53. Moores, A., Poyatos, M., Luo, Y. & Crabtree, R. H. Catalysed low temperature H2 release from nitrogen heterocycles. *New J. Chem.* **30,** 1675 (2006).
54. Clot, E., Eisenstein, O. & Crabtree, R. H. Computational structure?activity relationships in H2 storage: how placement of N atoms affects release temperatures in organic liquid storage materials. *Chem. Commun.* 2231 (2007). doi:10.1039/b705037b
55. Araujo, C. M. *et al.* Fuel selection for a regenerative organic fuel cell/flow battery: thermodynamic considerations. *Energy Environ. Sci.* **5,** 9534 (2012).
56. Driscoll, P. F., Deunf, E., Rubin, L., Arnold, J. & Kerr, J. B. Electrochemical Redox Catalysis for Electrochemical Dehydrogenation of Liquid Hydrogen Carrier Fuels for Energy Storage and Conversion. *J. Electrochem. Soc.* **160,** G3152–G3158 (2013).
57. INDOLE-3-ALDEHYDE. *Org. Synth.* **39,** 30 (1959).
58. PBT Profiler. *Persisten, Bioaccumulative, and Toxic Profiles Estimated for Organic Chemicals* Available at: http://www.pbtprofiler.net/.
59. US EPA, O. Ecological Structure Activity Relationships (ECOSAR) Predictive Model. *US EPA* (2015). Available at: https://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model. (Accessed: 7th December 2017)
60. Card, M. L. *et al.* History of EPI Suite™ and future perspectives on chemical property estimation in US Toxic Substances Control Act new chemical risk assessments. *Env. Sci Process. Impacts* **19,** 203–212 (2017).
61. Ceriani, L., Papa, E., Kovarich, S., Boethling, R. & Gramatica, P. Modeling ready biodegradability of fragrance materials: QSAR prediction of ready biodegradability of fragrances. *Environ. Toxicol. Chem.* **34,** 1224–1231 (2015).
62. Barber, M. C. A REVIEW AND COMPARISON OF MODELS FOR PREDICTING DYNAMIC CHEMICAL BIOCONCENTRATION IN FISH. *Environ. Toxicol. Chem.* **22,** 1963 (2003).
63. Cappelli, C. I., Benfenati, E. & Cester, J. Evaluation of QSAR models for predicting the partition coefficient (logP) of chemicals under the REACH regulation. *Environ. Res.* **143,** 26–32 (2015).

64. Colvin, M. E., Hatch, F. T. & Felton, J. S. Chemical and biological factors affecting mutagen potency. *Mutat. Res. Mol. Mech. Mutagen.* **400,** 479–492 (1998).

Chapter 3
Functional Toxicogenomics and Combinatorial Chemistry For Greener Design

This chapter explores the early phases of development for the biofuel 2,5-dimethylfuran (DMF) and how yeast functional toxicogenomics can be used to generate mechanistic toxicity data for DMF and related compounds so as to streamline the hazard assessment process, and to provide valuable insight for chemists developing new furan-based compounds. While the single example of the biofuel DMF is considered here, the general framework and development process used to evaluate it and several related compounds provide an example of how functional toxicogenomics and combinatorial chemistry approaches may be used to provide desperately-needed hazard data to the process of chemical design. With proper application, this functional toxicogenomics/combinatorial chemistry testing paradigm can help academic, commercial, and governmental labs develop greener chemicals much more quickly and effectively.

Energy Resources Are A Critical Element In The Transportation Sector

Since the advent of the industrial era, petroleum and petroleum products have taken on progressively greater importance in our daily lives, not the least of which is their value as a source of fuel. Indeed, the transportation sector accounts for over 60% of the oil consumed on an annual basis, and as the demand for consumer goods and personal transport increases worldwide, fuel demand can only be expected to rise.[1] Concerns over resource availability and security have prompted investigations into alternative liquid fuel sources[2,3], the most prominent of which, bioethanol, has failed to live up to expectations. Compared to gasoline, ethanol has "low energy density (reducing driving distance), high latent heat of vaporization and low vapor pressure (making engine cold start difficult), and water miscibility,"[1] prompting investigations into a more effective alternative.

While electric vehicles are a technology of growing promise, widespread adoption of the technology will likely take several decades[4]. Though desirable to consumers[5], the high cost of electric vehicles[6] and lack of critical infrastructure will delay a fully electrified transportation sector for the foreseeable future[4]. These limitations are further compounded by a resource bottleneck where the cost and availability of lithium and cobalt exert powerful influence on the ability to produce batteries for electric vehicles[7]. None of these problems are insurmountable, and with prudent implementation of transportation policy decisions they are likely to be overcome within the next few decades.

DMF is emerging as a prominent biofuel candidate compound

Until a fully electrified transportation sector can be achieved, however, a "bridge" technology is needed to span the gap between the fossil fuels of the past and our battery-powered future: Biofuels. Several recent advances in the production of DMF suggest it as a promising biofuel candidate: it can be readily and renewably produced from lignocellulose, it is immiscible in water, it can be mixed with gasoline without any additives, and it has a 34% greater energy density than ethanol[3,8]. However, while many research papers have been published about refining DMF production[3] and exploring its effectiveness as a fuel[8], only a handful of studies have investigated the hazard potential of DMF[9–11]. To address the data gaps in the DMF hazard profile, we consider the use of two high throughput

methodologies: metabolomics and yeast toxicogenomics to rapidly elucidate any extant mechanisms of toxicity.

Identifying mechanisms of toxicity for DMF

There is very little information about the metabolism or toxicity of DMF, but the structurally similar compound, furan, is well studied, and provides a starting point for the evaluation of DMF. Furan is classified as a "probable human carcinogen" in the most recent *Report on Carcinogens* from the National Toxicology Program [12] and "possibly carcinogenic to humans" by IARC[13]. It is a highly biologically-active structure [14] which is readily oxidized by CYP2E1[15] to produce toxic epoxide or *cis*-enedione metabolites[16]. Furan studies in rats have shown that its metabolites preferentially bind to lysine residues on cytosolic proteins and mitochondrial proteins, particularly cytochromes[14,17], as well as proteins involved in gluconeogenesis and glucose and fatty acid metabolism[18]. By one report, furan activated to *cis*-2-butene-1,4-dial was shown to be mutagenic in an Ames assay using an aldehyde-sensitive strain[19], but these results have not been independently reproduced[20]. As noted by Moro et al[21], while DNA adducts have been reported in mouse, rat, and turkey egg liver models, there is no consensus as to whether furan and its metabolites cause enough, if any, direct DNA damage to induce genotoxicity, or if the observed toxicity is secondary to mitochondrial damage[17–19,21–27].

Human exposure to and metabolism of furan compounds

Furan and DMF are both formed through the heating of sugars[28] and are common contaminants in cooked foods, especially coffee[29]. They are also found at low doses in cigarette smoke and have can be used as biomarkers for cigarette smoke exposure[30,31]. Metabolism of furans occurs primarily in liver, where they are activated by CYP2E1 and conjugated by glutathione for urinary excretion. (Fig. 2) The half-life of dietary furan is estimated to be less than six hours in rats [21], but no human toxicokinetic data exists for furan, and there is no toxicokinetic data whatsoever for DMF.

As with furan, there is insufficient mechanistic evidence to determine if DMF is genotoxic, and if it is, if the observed toxicity is secondary to another toxic endpoint – i.e. mitochondrial damage in the case of furan. Whatever the case, uncertainty about DMFs toxicological properties is due to a dearth of data, while furan's mechanism of action for carcinogenicity remains unclear in the face of dozens of genotoxicity studies due to inconsistent assay results[21]. Since no additional information is available for the hexene-diketone metabolite of DMF, it is unclear which, if any, modes of toxicity may be expected from DMF or its metabolites. Indeed, as evidenced by the metabolites of *n*-hexane, not all compounds of a chemical class necessarily share the same mechanisms of toxicity. It is likely that if DMF, its metabolites, or related furan compounds have toxicological properties, they are different than those of furan.

Very limited data is available for DMF metabolism

It has been established that the metabolism plays a key role in the toxicity of furan[18,19,32,33], but the importance of bioactivation in DMF toxicity is unclear[9,11]. Significantly less data is available for DMF than for furan: only two *in vitro* toxicity assays have been performed – a mutagenicity assay and a genotoxicity assay. The experiments have provided seemingly

discordant results for DMF: the compound was not observed to be mutagenic in the Ames assay (although it is not clear if DMF was metabolically activated by liver S9 fraction in the study)[34], but work by Fromowitz et al[9] indicated that DMF appeared to have clastogenic properties in an *in vitro* mouse erythropoietic micronucleus assay – with or without metabolic activation. However, this does not indicate that S9-activated DMF does not produce toxic metabolites, only that the metabolites did not affect the genotoxic endpoint being tested. It is worth noting that DMF itself is a metabolite and blood and urinary biomarker of hexane, a known neurotoxicant, although the possible role of DMF in hexane metabolism or mode of action is unknown.[35–37] The primary metabolite of DMF is predicted to be 3-hexene-2,5-dione[16,38], which has not been not been assessed for toxicity. One of the other breakdown products of hexane, hexane 2,5-γ-diketone, demonstrates delayed neuropathy and is a potent protein cross-linking agent in humans, while another hexane breakdown product, a 2,3-α-diketone, is benign and a commonly observed flavoring component in many foods[39]. This suggests that furans and their metabolites likely have very different toxicity profiles, and that they must be investigated individually.

Existing genotoxicity assay results for furan are inconsistent and provide insufficient mechanistic information

Since previous efforts have indicated roles for both furan and DMF as potential genotoxicants, special attention will paid to the assessment of genotoxicity. As noted earlier, the classical genotoxicity tests have provided conflicting or inconclusive evidence as to whether furan is mutagenic, genotoxic, or merely carcinogenic[21]. Furan was subjected to many of the most common *in vitro* mammalian methods for genotoxicity assessment: Chinese hamster ovary chromosomal aberration assay, mouse lymphoma *Tk*$^{+/-}$ gene assay, *Hprt* assay – all of which have been the subject of recent scrutiny for their inability to accurately predict human and rodent carcinogenicity [40–42]. Further, these assays provide little mechanistic information, so by optimizing a compound to minimize genotoxicity as measured by any single assay, one risks enhancing genotoxicity as measured by another assay – what is often referred to as a "regrettable substitution."[43] To combat regrettable substitutions, recent trends in toxicological assessment have turned towards –omics technologies, particularly genomics, to determine toxicological mechanism and to increase throughput[44–46].

A Holistic Approach To Testing

The state of California Code of Regulations (Title 22, Division 4.5, Chapter 54) lists over 40 different hazard endpoints that should be considered during green chemistry research – a truly daunting number that defies any single assay and challenges researchers to develop more comprehensive chemical testing strategies[47]. While some of the endpoints necessitate individual investigation (e.g. it is unlikely that the endpoints for phytotoxicity and neurotoxicity will be effectively probed using the same assay), it may be possible to consolidate the work for some endpoints if a suitable assay can be developed to probe mechanisms of action in eukaryotic cells in an unbiased way. Unbiased (also called "agnostic," or "hazard-generating") test strategies help researchers or regulators narrow their focus in hazard assessments, allowing them to identify adverse outcome pathways and body systems relevant to the chemical at hand. Testing every compound for every endpoint is unrealistic, expensive, and not necessarily informative, but an unbiased screen

for mechanisms of eukaryotic cell toxicity could provide significant insight that would permit fewer and more targeted toxicity tests for chemicals of interest. The ideal testing system would be a simulated human-on-a-chip, and this technology is currently in development[48,49], but it is decidedly low-throughput and likely many years from optimization and widespread adoption. Fortunately, there exists a bridge technology – in the form of functional toxicogenomics – that has tremendous potential as an alternative assay or component of an integrated testing strategy.

Unbiased toxicogenomic yeast screens have demonstrated utility

Toxicogenomic screening methods, i.e. unbiased screening approaches, are advantaged over assays that only allow for the interrogation of a single endpoint in that toxicogenomic screens allow the evaluation of multiple toxic endpoints in a single experiment. Among the toxicogenomic assays, the yeast (*Saccharomyces cerevisiae*) deletion libraries first generated and characterized by Winzeler et al[50] and Giaever et al[51] have demonstrated tremendous utility as functional toxicological screening tools[52–55]. As an unbiased screen, the yeast functional toxicogenomic system offers high-throughput detection of any assayable phenotypes in the yeast system, rather than the singular endpoints measured by the standard battery of test systems. Two yeast screening technologies, the haploinsufficiency profiling (HIP) and the homozygous deletion profiling (HOP) assays have been used to elucidate mechanisms of genetic toxicity and identify genes required for chemical tolerance in humans, due to the high level of genetic conservation between yeast and man[56–61]. The HIP and HOP assays identify which mutant strains demonstrate growth defect in the presence of DMF or its isolated metabolites. Between the HIP and the HOP assays, 97% of the open reading frames (ORFs) in the genome of *S. cerevisiae* have been deleted using barcodes can be tested for their functional importance for growth in the presence of a toxicant using Bar-seq[55]. Clustering patterns of the genes responsible for altered growth in the presence of DMF and its metabolites are then used to elucidate mechanisms of toxicity.

A Combinatorial Chemistry Approach To Biofuel Design

Given that placement of the double bonds in hexane breakdown products has significant effects on the toxicity of the compounds, it is likely that placement of double bonds in DMF breakdown products will be of similar importance, since both sets of breakdown products include diketones. Further, it is expected that the breakdown products of 2,5-dimethylfuran would differ from those of similar furans such as 2,3-dimethylfuran because the altered positions of the methyl groups would sterically hinder oxidation of different carbons. Therefore, we took a combinatorial chemistry approach to the question of biofuel design and used a yeast functional toxicogenomic system to screen 2,5-DMF, 2,3-DMF, 2-methylfuran (2-MF), and 2-ethylfuran (2-EF) to determine the effects of different functionalization schemes of the furan ring. It was anticipated that by shifting the location of the methyl groups or lengthening the alkyl chain, we could manipulate the stability and structure of reactive intermediates and products – thereby altering the degree and mechanisms of toxicity.

Rational molecular design is the notion that chemicals can be developed from the atoms up to ensure that function is balanced with toxicity, and that is the approach that is taken here.

Just as a mechanical engineer may produce several similar, but subtly different prototypes of a part for testing, we propose the use of combinatorial chemistry and functional toxicogenomics to quickly evaluate and define the chemical space around a structure of interest. This data can be collected relatively quickly in a scalable HTS fashion, providing green chemists with valuable information about the effects of their prototype chemicals in a whole-cell eukaryotic test system, and allowing them to make informed decisions about the trade-offs of one molecule over another.

Materials and Methods

Growth curve assays

Growth curves for yeast culture were determined as in North et al 2014[54]. Briefly: Cultures of *S. cerevisiae* with the BY4743 background were grown in YPD media (1% yeast extract, 2% peptone, 2% dextrose) to mid log-phase, diluted to an optical density at 600nm (OD600) of 0.0165, and dispensed into 96-well Greiner polypropylene plates (Sigma-Aldrich M9810) at 100μL per well. Wells were dosed with test compounds prepared in dilutions of DMSO such that each well received no more than 1% DMSO by volume. Compounds tested: 2,5-dimethylfuran (CAS#625-86-5, Santa Cruz Biotech sc-238384), 2-methylfuran (CAS#534-22-5, Tokyo Chemical Industry Co. MFCD00003248), 2-ethylfuran (CAS#3208-16-0, Tokyo Chemical Industry Co. MFCD00003259), 2,3-dimethylfuran (CAS#14920-89-9, Tokyo Chemical Industry Co. MFCD00153893). Each chemical was tested with six replicates at each concentration and compared to control wells dosed with either 1μL water or 1μL DMSO. Doses were tested in the range of 0.1μM - 30μL based on Fromowitz et al 2012[9]. Plates were incubated in Tecan microplate readers[62] set to 30°C with shaking and OD_{595} measurements were taken every 15min for 95 cycles (approximately 24 hrs). Treated wells were compared to DMSO control wells and the area under the curve was calculated to determine the degree of inhibition.

The IC20 value is the dose used in our studies

The concentration of the test compound necessary to reduce yeast growth by 20% as measured by optical density readings (the IC20 value) was determined in the yeast background strain BY4743 to establish a baseline for growth inhibition. IC20 was used rather than IC50 because it increases the sensitivity of the assays, allowing for more effective identification of "sickly" strains that grow more slowly than others due to the importance of their mutated gene for basic cell growth processes[55].

HOP assays to observe differential strain sensitivity

HOP screens are performed using best practices outlined in Pierce et al. 2007[63]using a pool of 4,653 BY4743 strains (ThermoFisher Life Sciences 95401.H1POOL), each with a different homozygous gene deletion, representing the totality of viable knockouts possible for *S. cerevisiae* with this background[64]. Pools are cultured overnight in YPD in a 30°C incubator shaker and then diluted to 0.0625 OD600 in YPD growth media. Pools were exposed to IC20 concentrations of test compounds (10mM for 2-MF, 7mM for 2,3-DMF, 8mM for 2,5-DMF, and 3mM for 2-EF) and samples were saved at 5, 10, and 15 generations, to allow for discrimination between early and late effects of toxicant exposure. Exposure to

DMSO is used as a control, and each well received no more than 1% DMSO by volume. Samples of pools were taken after 5 doubling generations (5G), 10 doubling generations (10G), and 15 doubling generations (15G), to yield three different time points. By analyzing sensitivity or resistance to chemical exposures at different time points, it is possible to determine which genes and biological processes are required for tolerance at acute, sub-chronic, and chronic exposures.

DNA Extraction, Purification, Sequencing

DNA is extracted from each pool using YeaStar Genomic DNA extraction kit (Zymo D2002) following experimental exposure[65]. The barcodes are amplified using PCR according to the protocols described in Robinson et al 2014[66]. Briefly: each well received 1.67μL of 5μM common primer, either

Reverse Up sequence:
CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTGCACGTCAAGACTGTCAAGG or

Forward Down:
CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTCAATCGTATGTGAATGCTGG),

and 2μL of 5 μM INDEX primer sequence:
ACACTCTTTCCCTACACGACGCTCTTCCGATCT+ index sequence as detailed on Appendix 5.

and 3.13μL of 24 ng/μL Genomic DNA, 43.2 μL Invitrogen Platinum PCR Supermix (Catalog #11306-016)[67]. The PCR was run for 3 minutes at 95°C, then 30 cycles of: 30 seconds at 94°C, 30 seconds at 55°C, 30 seconds at 72°C, then 3 minutes at 72°C. Primer product was purified, concentrated, and sequenced by Berkeley's Vincent J. Coates Genomics Sequencing Laboratory on an Illumina HiSeq2500 rapid flowcell. Sequencing data was analyzed as described in Robinson et al 2014[66], with the following comparisons computed using Fastq DESeq and a false discovery rate of (FDR) of 0.05.

Bar-seq analysis

Samples from the four furan test compounds were compared with DMSO controls at each of the three time points: 5G, 10G, and 15G. The enhancement or defect of growth was assessed using differential strain sensitivity analysis, and growth was measured by $\log_2$ fold change ($\log_2$ FC) in number of strain-specific barcodes compared to DMSO control, using a false discovery rate of >0.05[68] and pathways were constructed from statistically significant (P value >0.01 with Bonferroni correction) genes using the Funspec[69], AmiGo 2[70], and Saccharomyces Genome Database[71] web-based gene clustering tools to identify enriched gene ontology (GO) pathways. GO pathways marked as "enriched" by Funspec, but which only included a single gene from our dataset were not considered significantly enriched in this analysis.

Results

2,5-DMF Gene Enrichment Analysis

As noted earlier, the primary metabolite of DMF is known to be 3-hexene-2,5-dione[16,38] (3-HD), a compound without any available toxicity data. A similar compound, 2,5-hexanedione, causes neurotoxicity in rodents models[35], likely through oxidative stress of neurons[39,72].

Null Mutants With Resistance to 2,5-DMF

No mutants were consistently resistant to 2,5-DMF across all time points. Five strains with known gene deletions and two with deletions of genes of unknown function were resistant at the 5G time point. The genes did not appear to be related and are involved in various functions including tRNA export (Sol2)[73], ubiquitin-dependent endocytosis (Rog3), budding (Bud9 and Prm10), and sterol synthesis regulation (Spt23)[74]. The deleted gene in the *BUD9*Δ strain is associated with resistance to aneuploidy, a condition which might be expected in the event of oxidative insult to the DNA[75]. The resistant strains at the 10G and 15G time points displayed only minor increases in $\log_2$ fold change and Funspec analysis did not indicate any enriched gene ontologies.

Null Mutants With Sensitivity to 2,5-DMF

Three sensitive strains at 10G were mutants for genes involved in response to oxidative stress: *YAP1Δ*, *IMP2Δ* (involved in preventing DNA damage against oxidants), *OYE2Δ* (a flavin mononucleotide NADPH oxidoreductase) and *MNR2Δ* (a metal ion transporter, the overexpression of which is necessary for resistance to manganese and other oxidative metals[76]). Other genes required for tolerance to 2,5-DMF at this time point were involved in DNA replication and cell cycle: LTE1 regulates mitotic spindle formation and Csm1 is vital for homolog segregation. Both Tat1 and Stp1 are important for cell amino acid transport[77,78], which may be expected given likely GSH flux[33,79] in the presence of known oxidizers like 2,5-DMF and 3-HD.

At the "chronic exposure" 15G time point, we observed that among the genes required for 2,5-DMF tolerance, there was GO enrichment of genes coding for oxidative stress response (*YAP1* and *WHI2*) and the often-related pathways of DNA repair (*WSS1*, *RAD57*, and *APN1*) and chromosome segregation (*CSM1Δ* and *LRS4Δ*).[80] (Table 3.1) Cadet and Davies note that Wss1 (homologous to the human Spartan enzyme), is necessary for tolerance to formaldehyde, A canonical DNA cross-linking agent[81], and Allam et al report that Wss1 is vital for repairing DNA-protein cross-links and maintaining genomic integrity.[82] *RAD57* belongs to the *RAD51* subgroup of genes required for successful non-homologous end-joining repair of double strand breaks.[83] Apn1 is a major enzyme in the base excision repair pathway, responsible for mending apurinic lesions in DNA that result from alkylation, oxidation, or hydrolysis of DNA bases.[84] Taken together with the sensitivity observed in the *CSM1Δ* and *LRS4Δ*, our data suggests that chronic 2,5-DMF exposure results in DNA damage and genomic instability in yeast.

DNA damage is not the only mechanism of toxicity we observe, however: GO pathways for protein recycling were also enriched with susceptible mutants at 15G (*IRS4Δ, DOA1Δ,*

FIS1Δ, VID28Δ, and *VTC1Δ*). Loss of inter- and intra-cellular transport proteins (Get1, Sys1, Tlg2, Vac7, Vtc1) and amino acid transport proteins such as Tat1 and Thr4 confers sensitivity to 2,5-DMF as well, likely due to increased need for raw materials to replaced damaged proteins and other macromolecules.

The only strain consistently sensitive to 2,5-DMF across all three time points was the *YAP1*Δ strain (Table 3.6). Yap1 is an important transcription factor controlling a regulon of dozens of genes critical for oxidative stress tolerance[85]. It is activated by indirectly by oxidizing compounds, relying on peroxidases such as Gpx3 to decouple it from Crm1[86], which facilitates Yap1 export from the nucleus in the absence of oxidative stress.[87,88] This suggests that the primary mechanism of toxicity of 2,5-DMF is oxidative stress, which may be induced by its reaction with cellular proteins, likely lysines and cysteines[33,79,89]. However, the loss of representation among genes related to DNA repair at the 15G time point supports the findings of Fromowitz et al[9]. Significant loss of representation among the *APN1Δ, RAD57Δ*, and *WSS1Δ* strains suggest that chronic exposure to 2,5-DMF or its primary metabolite 3-HD causes non-specific DNA-protein adducts. Such adducts would also likely cause genomic instability and double strand breaks, which accounts for the sensitivity of the *WSS1Δ, CSM1Δ,* and *LRS4Δ* strains. At sub-chronic and chronic exposure levels, we observed that protein synthesis and recycling pathways became important for 2,5-DMF tolerance as well, suggesting that acute 2,5-DMF exposures cause oxidative damage, and that over time that damage begins to affect protein function and eventually DNA integrity.

2,3-DMF Gene Enrichment Analysis

Null Mutants With Sensitivity to 2,3-DMF

Taking the metabolism of 2,5-DMF as a guide, we may anticipate that the primary metabolite of 2,3-DMF to be 2-Methyl-4-oxo-2-pentenal. The structure of 2-Methyl-4-oxo-2-pentenal (2-MOP) is consistent with the reaction schema put forth by Peterson et al[16], although a review of the literature indicates that the primary synthesis route for 2-MOP is through the oxidation of *m*-xylene[90]. Unfortunately, there is a paucity of literature on 2,3-DMF, and even less on 2-MOP. No toxicity data exists for either compound.

MIPS Functional classification yields several significantly enriched pathways wherein the loss of certain genes promotes resistance to 2,3-DMF at the 5G time point: Peroxisomal transport; lipid, fatty acid, and isoprenoid metabolism; and protein synthesis. However, we interpret these pathways with caution because only five of the 44 resistant strains showed a $\log_2$ FC in growth greater than 1.5: *ACE2Δ* ($\log_2$ FC 1.665), *FPS1Δ* ($\log_2$ FC 1.801), *GIN4Δ* ($\log_2$ FC 2.010), *MMS2Δ* ($\log_2$ FC 2.552), and *SWR1Δ* ($\log_2$ FC 1.759848). (Table 3.9)

There are no clear mechanistic reasons based on the chemistry of simple furans that indicate why loss of these genes results in increased cellular growth and proliferation. Acb1 is normally only expressed under starvation conditions, and the null mutant has been observed to replicate extensively under these conditions.[91] Increased proliferation here may occur by a similar mechanism, with cell stress resulting from furan exposure rather than starvation conditions. Deletion of *FPS1* reduces glycerol permeability at the plasma

membrane, alters plasma membrane composition, improves xylose fermentation, and enhances cell resistance to osmotic shock.[92,93]. Two resistant genes, Mms2 and Doa1, work together to play an important role in ubiquitin-mediated proliferating cell nuclear antigen (PCNA) degradation – a crucial pathway for DNA damage response under conditions of cell stress[94].

While 81 ORFs were determined to be statistically significantly overrepresented at the 10G time point, only 68 identified genes were among them, and of those, only genes for protein-lysine N-methyltransferase activity were enriched for this condition and time point, and only two strains, *FPR2Δ* and *VAC14Δ* showed $\log_2$ FC in growth greater than 1.5 (1.727 and 2.001, respectively). (Appendix 2) Fpr2 is an ER membrane protein that exacerbates cell sensitivity to stress caused by protein aggregates[95] – as might be formed upon treatment with a reactive furan compound. The absence of a consistent pattern or pathway suggests that the mechanism of action is non-specific.

Intriguingly, loss of a variety of genes related to homeostasis of iron confers resistance at the latest time point: *ARN2Δ, FET3Δ, FTR1Δ* strains all showed growth improvement at 15G.[96] GO enrichment for genes related to iron and other metals is seen at the level of biological processes, cellular components, molecular function, and MIPS classification. This is consistent with the hypothesis that 2,3-DMF is a potent oxidizing agent within the cell, as extended exposure to 2,3-DMF may shift redox conditions within the cell to such a point as to make iron and other metals to be too reactive and dangerous within the cell. Other researchers have noted the potency of iron as an oxidizing agent within yeast cells[68].

Only sixteen mutants were resistant to 2,3-DMF across all three time points (Appendix 2), of which, four corresponded to unverified or uncharacterized ORFs, leaving twelve genes whose null genotype confers resistance. No meaningful GO enrichment is observed among these genes, although a few potential mechanisms of resistance may be inferred from gene function. *DOA1*, *FPR2*, and *LMO1* all relocate to different cellular compartments under oxidative or DNA replication stress conditions, and as noted earlier *ACB1*'s null phenotype is prone to increased proliferation during periods of cell stress. Similarly, the null mutant for BSC2 is known to have an increased translation rate, which may promote resistance through more rapid replacement of damaged proteins[97]. *PRY3* codes for a cysteine-rich cell wall protein[98], so the null mutant may have more cysteine on hand to replace sulfur-switches damaged through oxidative stress, improving cell viability. Mechanisms for the resistant phenotype observed in the remaining null mutants are not readily apparent.

Null Mutants With Sensitivity to 2,3-DMF

Enriched GO pathways for Cytoplasm-to-vacuole (CVT) pathway, ER membrane proteins, Golgi membrane proteins, N-terminal protein acylation, and cellular protein localization and transport are observed in the 5G time point, and consistently across later time points. (Table 3.10) Get1, Get2, Cog5, Cog6, Trs95, Sys1, and Tlg2 are all involved at different stages of the trans-Golgi network[99,100] Trs85 and Irs4 both play roles in the formation of the autophagosome, an important protein recycling structure[101,102], both proteins also work with Cog5, Cog6, and Tlg2 to facilitate the CVT pathway – the vacuole being the other major protein recycling organelle.[103]

The list of 2,3-DMF-sensitive mutants at the 10G time point is very contains many of the same strains as that of the 5G time point – only expanded to include yet more genes related to oxidative stress and protein metabolism. Cog7, Cog8, Get3, Sec22, Sso2, and Snc2 are added to the lists of CVT, trans-Golgi network and vesicle-mediated transport proteins whose null mutants experience adverse growth in the presence of 2,3-DMF. This seems to indicate that at sub-chronic exposure, these pathways become significantly more sensitive to 2,3-DMF exposure. The *YAP1Δ* strain also demonstrated decreased viability at this time point. (Appendix 2)

After 15G growth, the sensitive strains at this time point cover a similar range of cellular responses to oxidative stress and general protein production to earlier time points. Null mutants for genes related to Golgi and ER structure and function are significant here, as they were in the 10G time point, and mechanisms that disrupt transit between them or the integration of membrane proteins are significantly in the GO for biological processes. (Table 3.12) Additionally, pathways involved inter- and intra-cellular transport continue to be significant for 2,3-DMF tolerance. One possible explanation is that these pathways are vital to transporting damaged proteins for recycling and importing the raw materials necessary to replace them. GO for pathways related to cell cycle progression were not observed among to the sensitive mutants of earlier time points, but are found here, with noticeable growth defects in the *ACE2Δ* and *MSA1Δ* strains. The Ace2 transcription factor is necessary for cell cycle progression and septum destruction following cytokinesis. It has been noted by other groups that Ace2 is necessary for tolerance to furural and hydroxyfurural, and that overexpression of Ace2 can confer resistance to yeast prompted to produce these fuel compounds in bioreactors – although no mechanistic explanation exists as to why.[104,105]

Thirty mutants were found to be sensitive to 2,3-DMF across all time points, and of those, 6 ORFs were uncharacterized or unverified and did not correspond to known genes. The remaining 24 genes belonged to the previously observed pathways for protein recycling and transport, ER and Golgi function, and oxidative stress response. These findings were consistent with the hypothesis that the primary mechanism of 2,3-DMF cellular toxicity is through oxidative protein damage. Strains lacking proteins related to oxidative stress pathways (Apj1, Dbf2, Skn7, and Mga2) and protein synthesis (Get1, Get2, Irs4, Mak10, Mak3, Sys1, Tlg2, Trs85) and recycling (Rpn4) pathways were sensitive to 2,3-DMF across all three time points (Appendix 2). Although genes for oxidative stress and protein metabolism pathways are consistent across time points, some variation in the particular strains that display resistance is observed.

2-MF Gene Enrichment Analysis

Previous work by Ravindranath and Boyd indicates that the primary metabolite of 2-MF is acetylacrolein, and that acetylacrolein preferentially binds cysteine moieties over the oxidant-scavenging glutathione, although it binds both groups of molecules with great efficacy.[29]

Of the many strains with increased representation at the 5G time point for 2-MF exposure, only *PRM2Δ, STP1Δ, ELO3Δ, TUP1Δ, VPS27Δ,* and *YCK3Δ* had a $\log_2$ FC increase greater than 1.5. (Appendix 3) *PRM2* is poorly characterized but seems to have something to do with nuclear fusion in mating yeast[106]. *ELO3* encodes a fatty acid elongase involved in sphingolipid metabolism, and the null mutant accumulates inositol phosphoceramide[107]. Perhaps the accumulated fatty acid precursors act as a sponge to react with the 2-MF, preventing additional damage to cellular components. *TUP1* encodes a transcriptional repressor[108,109] that suppresses cell cycle progression in response to DNA damage. Vps27p is an endosomal protein necessary for sorting and recycling proteins[110], and Yck3p is responsible for regulating vesicle fusion to the vacuole[111] – both proteins play important roles in the alkaline phosphatase pathway as well as ESCRT complex endosomal recruitment[112].

At this and the other time points, enriched Gene ontologies for cellular components including the ESCRT II complex, endosome membrane, endosome, ESCRT III complex, and general membrane proteins were observed among the resistant strains. ESCRT complexes have been linked to yeast survival in conditions with altered glutathione availability.[113,114] Comparing lists of resistant strains across all time points, we see that although there are considerably more mutants with enhanced tolerance to 2-MF than we see among the other test compounds, the $\log_2$ FC increases tend to be small, and the same GO pathways appear across the three time points. (Tables 3.15, 3.18, 3.21, 3.24) Null mutants for genes involved in protein processing, targeting, and transport to the vacuole; as well as genes involved in cellular response to stress, anoxia, and pH; carbohydrate metabolism, and a handful of transcription repressors are consistently resistant to 2-MF.

Among the sensitive mutants, a few GO pathways are enriched for cellular features related to oxidative stress (DBF2, YAP1) and cellular protein synthesis (CVT pathway, glutathione synthesis pathway, and Golgi transport) – pathways also observed among the strains sensitive to other test compounds. The mutants sensitive to 2-MF include strains lacking genes involved in phosphate homeostasis, signal transduction, chromatin silencing at telomeres, and nuclear deacetylation. Chromatin structure maintenance appears to be significant for tolerance across all time points. (Table 3.38) Relatedly, genes related to ribonucleotide processing (mRNA export from the nucleus, negative regulation of transcription from RNA Pol I promoter, purine ribonucleotide monophosphate biosynthesis), are necessary to thrive under 2-MF exposure.

2-EF Gene Enrichment Analysis

Based on the furan degradation pathways worked out by Peterson et al[16], the likely breakdown product of 2-EF in a cellular milieu would be expected to be 2E)-4-oxo-2-hexenal (4-OHE), and indeed other researchers have confirmed that this is the case.[115] A 2010 review by Long and Picklo[116] describes 4-OHE as a potent oxidizing agent with similar properties to *trans*-4-hydroxy-2-nonenal (HNE), a well-characterized lipid peroxidation product. 4-OHE can be generated through oxidation of polyunsaturated fatty acids, and is suspected to be more reactive than HNE. Although mutagenic in *S. typhimurium*[117], oral administration of tritiated 4-OHE to rats revealed no detectable DNA

adducts, although high levels of liver protein adduct formation were detected up to 16 hours after exposure[118].

Our findings support the hypothesis that 2-EF toxicity is mediated through a mechanism of non-specific protein adduct formation and oxidative stress due to depleted GSH stores. No gene deletions or pathways demonstrated improved tolerance to 2-EF across any time point. However, mutants for genes related to oxidative stress response and amino acid uptake were sensitive across all time points and showed decreased fitness when exposed to 2-EF. Strains lacking the oxidative stress response factors Skn7 and Yap1 consistently demonstrated decreased fitness (Table 3.49), which is consistent with the evidence of 2-EF and 4-OHE as oxidative agents. At the 5G and 10G time points, genes involved in Golgi function and membrane trafficking such as Cog5, Cog6, and Sys1 were seen to be necessary for 2-EF tolerance (Tables 3.41 and 3.45), however, slightly different Golgi membrane and transport genes were required for tolerance at the 15G time point: Drs2, Erv14, Get1, Gyp1, Trs65, Trs85, Tlg2, Snc2, and Vps13. (Appendix 4) Similarly, a fully functioning Cytosol-vesicle transport pathway was necessary for 2-EF tolerance at all time points, but again, different sets of genes appeared to be necessary for tolerance at early (5 and 10G) versus late (15G) time points. (Appendix 4). Protein transport between the Golgi and other cellular compartments and between the cytosol and the vacuole is critical for protein synthesis, maturation, and recycling in the cell.[119–121]

These data indicates that systems related to protein flux through the cell are particularly sensitive to 2-EF and its likely metabolite, 4-OHE, and that these systems are adaptive – the yeast are responding to chronic exposure. By the final time point, the same three oxidative stress gene mutants (*SKN7Δ, YAP1Δ, IRS4Δ*) showed loss of fitness as in the other time points, but a larger and different suite of Golgi proteins were highlighted along with a smattering of genes for ER structure. Nine mutants for genes related to the structure of the Golgi apparatus were significantly less abundant, as well as roughly half a dozen genes for trans-Golgi network (TGN) vesicle trafficking including SNARE and TRAPP complexes. (Tables 3.48-3.51)

Discussion

These experiments demonstrated that very minor changes to chemical structure – in this case, the re-positioning of a methyl group – could result in significant changes to the mechanisms of action. Although the cellular pathways required for resistance to 2,5-DMF, 2,3-DMF, and 2-EF (oxidative stress response and protein metabolism pathways) were broadly similar, there were nuances in their suites of affected genes that seem to account for the different IC20 values of the various compounds: 7mM for 2,3-DMF, 8mM for 2,5-DMF, and 3mM for 2-EF.

There were, as expected quite a few similarities among the tolerance/susceptibility profiles of the different compounds, particularly among 2,5-DMF, 2,3-DMF, and 2-EF. Null mutants for ER and Golgi structural proteins related to vesicle transport are particularly vulnerable to 2,3-DMF exposure, while deletions among the proteins involved in the CVT pathway consistently produced sensitivity to 2-EF. The Cog proteins are evolutionarily conserved to

a significant degree[99] and defects in these proteins are associated with developmental and neurological disorders[100]. This is particularly interesting given the similarity between 3-hexene-2,5-dione and hexane 2,5-γ-diketone, the later of which is associated with delayed neuropathies. Interestingly, loss of the Cog genes did not result in sensitivity to 2,5-DMF exposure, although null mutants for Cog genes were sensitive to both 2,3-DMF and 2-EF. There are no obvious mechanistic reasons, however as to why mutants for vacuole transport proteins are more tolerant to 2-MF.

Kim and Hahn 2013[104] and Thi My Nguyen et al[122] both examine the importance of *YAP1* for furfural tolerance, so it is not surprising to see that *YAP1* is vital for tolerance to these furans as well – even in the case of 2-MF. Indeed, only the *YAP1*Δ strains showed reduced representation across all furan treatment conditions and all time points compared to control. An excellent review by Witz notes that while α,β-unsaturated aldehydes tend to bind glutathione or other protein sulfhydryl moieties, these compounds are also capable of causing significant lipid peroxidation as well[123], which may account for some of the differences we observe between the different furan compounds.

Conclusion

Replacing the petroleum-based transportation economy with a renewable biofuel is a laudable goal, and certainly worthy of further exploration; however, supplanting such a large industry requires prudence, and it is prudent to identify undesired activity in a proposed replacement compound so as to avoid the scenario of a regrettable substitution. 2,5-dimethylfuran has many enticing properties, but as demonstrated in this report, a rather significant drawback as well. Using the principles of green molecular design and the power of functional toxicogenomics, we have tested a series of potential alternatives to 2,5-DMF, which are potentially similar enough so as to be produced through similar synthesis pathways, but which are structurally different enough so as to not have the same hazard profile. By shifting the position of a single methyl group, we have changed the toxicological properties of our chemical of interest – in the case of 2-MF, rather dramatically.

While the doses we tested were high by the standards of most environmental chemicals, they may be considered fit for purpose in this case, since the interest of 2,5-DMF as a biofuel presupposes that the chemical would become widely available, and that if so adopted, exposures to high concentrations of the compound are not only possible, but likely. That being the case, it is prudent to consider the possible consequences of mass-producing a compound that appears to produce oxidative DNA damage under certain exposure conditions. The genetic evidence we have collected points to a possible mechanism of 2,5-DMF toxicity through DNA damage through non-specific oxidative damage leading to DNA-protein adduct formation. 2,5-DMF, for example, appears to be a more potent toxicant at 15G to strains lacking various DNA repair proteins, while none of the other furans tested here share that feature. These results suggest two conclusions, the first being that there is toxicogenomic support for the role of 2,5-DMF as a DNA damaging agent, and the second being that the pairing of functional toxicogenomics and combinatorial chemistry can yield powerful and informative results.

Tables

2,5-DMF 15G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
protein localization to nucleolar rDNA repeats [GO:0034503]	0.000148	*CSM1, LRS4*	2	5
homologous chromosome segregation [GO:0045143]	0.000221	*CSM1, LRS4*	2	6
rDNA condensation [GO:0070550]	0.000309	*CSM1, LRS4*	2	7
negative regulation of gluconeogenesis [GO:0045721]	0.000527	*UBC8, VID28*	2	9
response to heat [GO:0009408]	0.001955	*YAP1, WHI2*	2	17
chromatin silencing at rDNA [GO:0000183]	0.002990	*LRS4, IRS4*	2	21
apoptosis [GO:0006915]	0.003281	*OYE2, FIS1*	2	22
endocytosis [GO:0006897]	0.003906	*THR4, TLG2, WHI2*	3	82
proteasomal ubiquitin-dependent protein catabolic process [GO:0043161]	0.006877	*UBC8, VID28*	2	32
CVT pathway [GO:0032258]	0.009124	*IRS4, TLG2*	2	37

Table 3.1: 2,5-DMF 15G Sensitive Strains GO Biological Process

2,5-DMF 15G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
monopolin complex [GO:0033551]	8.90E-05	*CSM1, LRS4*	2	4
trans-Golgi network [GO:0005802]	0.004230	*SYS1, TLG2*	2	25

Table 3.2: 2,5-DMF 15G Sensitive Strains GO Cellular Component

2,5-DMF 15G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
regulation of glycolysis and gluconeogenesis [02.01.03]	0.002446	*UBC8, VID28*	2	19

Table 3.3: 2,5-DMF 15G Sensitive Strains MIPS Functional Classification

2,3-DMF 5G Resistant Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
peroxisomal transport [20.09.03]	0.000248	*DJP1, CAT2, ANT1*	3	19
lipid, fatty acid and isoprenoid metabolism [01.06]	0.001777	*SUR2, ACB1, DOA1, FPS1, CAT2*	5	133
PROTEIN SYNTHESIS [12]	0.009211	*SRO9, PET130*	2	22

Table 3.4: 2,3-DMF 5G Resistant Strains MIPS Functional Classification MIPS Functional classification yields a handful of significantly enriched pathways: Peroxisomal transport; lipid fatty acid, and isoprenoid metabolism; and protein synthesis. There are no clear mechanistic reasons, based on likely 2,3-DMF metabolites that indicate why loss of these genes results in increased cellular growth and proliferation. ACB1 is normally only expressed under starvation conditions, and the null mutant has been observed to replicate extensively under these conditions. Increased proliferation here may occur by a similar mechanism, with cell stress resulting from furan exposure rather than starvation conditions. DOA1 has an important role in ubiquitin-mediated protein degradation, deletion of FPS1 improves xylose fermentation, and CAT2 may play a role in other fermentation-related shunt pathways.

2,3-DMF 10G Resistant Strains GO Molecular Function				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
protein-lysine N-methyltransferase activity [GO:0016279]	0.002926	*RKM4, SEE1*	2	7

Table 3.5: 2,3-DMF 10G Resistant Strains GO Molecular Function While 81 ORFs were determined to be statistically significantly overrepresented, only 68 identified genes were among them, and of those, only genes for protein-lysine N-methyltransferase activity were enriched for this condition and time point.

2,3-DMF 15G Resistant Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
iron ion homeostasis [GO:0055072]	0.000120	*YDR506C, SIT1, FTR1, ARN2, FET3*	5	26
iron assimilation by reduction and transport [GO:0033215]	0.000372	*FTR1, FET3*	2	2
high-affinity iron ion transport [GO:0006827]	0.002180	*FTR1, FET3*	2	4
response to copper ion [GO:0046688]	0.003588	*YCR102C, FET3*	2	5
ion transport [GO:0006811]	0.004572	*YDR506C, SIT1, FTR1, ARN2, TOK1, FET3, ANT1*	7	107
siderophore transport [GO:0015891]	0.009671	*SIT1, ARN2*	2	8

Table 3.6: 2,3-DMF 15G Resistant Strains GO Biological Process

2,3-DMF 15G Resistant Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
high affinity iron permease complex [GO:0033573]	0.000372	*FTR1, FET3*	2	2

Table 3.7: 2,3-DMF 15G Resistant Strains GO Cellular Component

2,3-DMF 15G Resistant Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
siderophore-iron transport [20.01.01.01.01.01]	5.91E-05	*SIT1, FTR1, ARN2, FET3*	4	12
homeostasis of metal ions (Na, K, Ca etc.) [34.01.01.01]	0.000584	*YDR506C, SIT1, FTR1, TOS8, ARN2, TOK1, FET3, IZH2*	8	98
protein folding and stabilization [14.01]	0.009095	*HSP26, HSP78, FPR2, DJP1, JJJ1, RBL2*	6	93

Table 3.8: 2,3-DMF 15G Resistant Strains MIPS Functional Classification Loss of a variety of genes related to homeostasis of iron and other metals confers resistance at this time point. GO enrichment for genes related to iron and other metals is seen at the level of biological processes, cellular components, molecular function, and MIPS classification. This is consistent with the hypothesis that 2,3-DMF is a potent oxidizing agent within the cell, as extended exposure to 2,3-DMF may shift redox conditions within the cell to such a point as to make iron and other metals to be too reactive and dangerous within the cell. Other researchers have noted the potency of iron as an oxidizing agent within yeast cells[68].

2,3-DMF 5G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
modification by acetylation, deacetylation [14.07.04]	0.005381	*MAK10, SPT8, SAP30, MAK3*	4	69
cAMP/cGMP mediated signal transduction [30.01.09.07]	0.006452	*CYR1, RAS2*	2	12

Table 3.9: 2,3-DMF 5G Sensitive Strains MIPS Functional Classification Only a 16 mutants were resistant to 2,3-DMF across all three time points, of which, 4 corresponded to unverified or uncharacterized ORFs, leaving 12 genes whose null genotype confers resistance. No meaningful GO enrichment is observed among these genes, although a few potential mechanisms of resistance may be inferred from gene function. DOA1, FPR2, and LMO1 all relocate to different cellular compartments under oxidative or DNA replication stress conditions, and as noted earlier ACB1's null phenotype is prone to increased proliferation. Similarly, the null mutant for BSC2 is known to have an increased translation rate, which may promote resistance through more rapid replacement of damaged proteins. PRY3 codes for a cysteine-rich protein, so the null mutant may have more cysteine on hand to replace sulfur-switches damaged through oxidative stress, improving cell viability. Mechanisms for the resistant phenotype observed in the remaining null mutants are not readily apparent.

2,3-DMF 5G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
CVT pathway [GO:0032258]	3.37E-05	*TRS85, IRS4, COG6, COG5, TLG2*	5	37
protein insertion into ER membrane [GO:0045048]	0.001024	*GET2, GET1*	2	5
N-terminal protein amino acid acetylation [GO:0006474]	0.001526	*MAK10, MAK3*	2	6
Golgi apparatus [GO:0005794]	9.28E-08	*TRS85, VPS52, GET2, GET1, SYS1, COG6, COG5, TLG2*	8	213
Golgi membrane [GO:0000139]	2.01E-05	*GET2, GET1, SYS1, COG6, COG5*	5	117
GET complex [GO:0043529]	2.61E-05	*GET2, GET1*	2	3
Golgi transport complex [GO:0017119]	0.000241	*COG6, COG5*	2	8
pre-autophagosomal structure [GO:0000407]	0.002122	*TRS85, IRS4*	2	23
trans-Golgi network [GO:0005802]	0.002508	*SYS1, TLG2*	2	25
cellular protein localization [GO:0034613]	0.0028114	*GET2, GET1*	2	8
transport [GO:0006810]	0.0029983	*ATP1, TRS85, RAV2, YDR338C, VPS52, GET2, GET1, YOR1, SYS1, MOG1, FRE8, PHO84, COG6, COG5, BRE5, TLG2, YOL075C*	17	815

Table 3.10: 2,3-DMF 5G Sensitive Strains GO Biological Process

2,3-DMF 10G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
CVT pathway [GO:0032258]	6.39E-07	*TRS85, COG7, IRS4, COG8, COG6, COG5, TLG2*	7	37
vesicle-mediated transport [GO:0016192]	2.40E-05	*LTE1, GET3, TRS85, GET2, GET1, SEC22, SSO2, TLG2, GYP1, SNC2*	10	140
protein insertion into ER membrane [GO:0045048]	2.65E-05	*GET3, GET2, GET1*	3	5
intra-Golgi vesicle-mediated transport [GO:0006891]	0.000315	*COG7, COG8, COG6, COG5*	4	24
response to arsenic-containing substance [GO:0046685]	0.000411	*RPN4, GET3, YAP1*	3	11
retrograde vesicle-mediated transport, Golgi to ER [GO:0006890]	0.000582	*GET3, GET2, GET1, SEC22*	4	28
vesicle fusion [GO:0006906]	0.000582	*SEC22, SSO2, TLG2, SNC2*	4	28
response to singlet oxygen [GO:0000304]	0.001156	*SKN7, YAP1*	2	4
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	0.001156	*SKN7, YAP1*	2	4
positive regulation of transcription from RNA polymerase II promoter [GO:0045944]	0.002734	*STP1, PHO4, AFT1, MGA2, SPT8, SAP30*	6	100
response to metal ion [GO:0010038]	0.002837	*GET3, YAP1*	2	6
N-terminal protein amino acid acetylation [GO:0006474]	0.002837	*MAK10, MAK3*	2	6
cellular protein localization [GO:0034613]	0.005200	*GET2, GET1*	2	8

Table 3.11: 2,3-DMF 10G Sensitive Strains GO Biological Process

2,3-DMF 15G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
vesicle-mediated transport [GO:0016192]	5.87E-06	*APM3, FEN1, GET3, ARF1, TRS85, GET2, GET1, APS3, SEC22, SSO2, TLG2, GYP1, RUD3, SNC2*	14	140
response to arsenic-containing substance [GO:0046685]	9.36E-05	*RPN4, GET3, HOG1, YAP1*	4	11
protein insertion into ER membrane [GO:0045048]	0.000132	*GET3, GET2, GET1*	3	5
response to heat [GO:0009408]	0.000603	*GET3, YAP1, WSC2, LSP1*	4	17
ER to Golgi vesicle-mediated transport [GO:0006888]	0.000631	*ARF1, TRS85, TCA17, BST1, EMP47, ERV14, SEC22, RUD3*	8	80
vesicle organization [GO:0016050]	0.002568	*TRS85, BST1, SYS1*	3	12
response to singlet oxygen [GO:0000304]	0.003349	*SKN7, YAP1*	2	4
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	0.003349	*SKN7, YAP1*	2	4
retrograde vesicle-mediated transport, Golgi to ER [GO:0006890]	0.004222	*GET3, GET2, GET1, SEC22*	4	28
vesicle fusion [GO:0006906]	0.004222	*SEC22, SSO2, TLG2, SNC2*	4	28
response to metal ion [GO:0010038]	0.008110	*GET3, YAP1*	2	6
negative regulation of transcription from RNA polymerase I promoter [GO:0016479]	0.008110	*SAP30, PHO23*	2	6
regulation of transcription involved in G1 phase of mitotic cell cycle [GO:0000114]	0.008110	*ACE2, MSA1*	2	6
N-terminal protein amino acid acetylation [GO:0006474]	0.008110	*MAK10, MAK3*	2	6
osmosensory signaling pathway [GO:0007231]	0.008110	*PBS2, HOG1*	2	6

Table 3.12: 2,3-DMF 15G Sensitive Strains GO Biological Process

2,3-DMF Sensitive Strains Across All Time Points GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
protein insertion into ER membrane [GO:0045048]	0.000197	*GET2, GET1*	2	5
N-terminal protein amino acid acetylation [GO:0006474]	0.000295	*MAK10, MAK3*	2	6
cellular protein localization [GO:0034613]	0.000549	*GET2, GET1*	2	8
CVT pathway [GO:0032258]	0.000592	*TRS85, IRS4, TLG2*	3	37
vesicle organization [GO:0016050]	0.001280	*TRS85, SYS1*	2	12
Ras protein signal transduction [GO:0007265]	0.001508	*CYR1, RAS2*	2	13
vesicle-mediated transport [GO:0016192]	0.003456	*TRS85, GET2, GET1, TLG2*	4	140

Table 3.13: 2,3-DMF Sensitive Strains Across All Time Points GO Biological Process

2,3-DMF Sensitive Strains Across All Time Points GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
NatC complex [GO:0031417]	5.97E-05	*MAK10, MAK3*	2	3
GET complex [GO:0043529]	5.97E-05	*GET2, GET1*	2	3
Golgi apparatus [GO:0005794]	0.002457	*TRS85, GET2, GET1, SYS1, TLG2*	5	213
pre-autophagosomal structure [GO:0000407]	0.004758	*TRS85, IRS4*	2	23
trans-Golgi network [GO:0005802]	0.005610	*SYS1, TLG2*	2	25

Table 3.14: 2,3-DMF Sensitive Strains Across All Time Points GO Cellular Component

2-MF 5G Resistant Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway [GO:0043162]	2.14E-08	*STP22, VPS25, SNF7, VPS36, VPS20, SNF8*	6	15
negative regulation of transcription from RNA polymerase II promoter by glucose [GO:0000433]	7.70E-05	*VPS25, VPS36, SNF8*	3	7
late endosome to vacuole transport [GO:0045324]	0.000121	*STP22, SNF7, VPS20, VPS27*	4	20
protein targeting to vacuole [GO:0006623]	0.000192	*STP22, VPS25, VPS36, VPS27, SNF8*	5	41
cellular response to anoxia [GO:0071454]	0.001035	*RIM101, SNF7*	2	4
response to pH [GO:0009268]	0.001710	*BPH1, RIM101*	2	5
intralumenal vesicle formation [GO:0070676]	0.001710	*SNF7, VPS20*	2	5
negative regulation of transcription, DNA-dependent [GO:0045892]	0.007372	*TUP1, RGT1*	2	10
protein retention in Golgi apparatus [GO:0045053]	0.008932	*VPS36, VPS27*	2	11

Table 3.15: 2-MF 5G Resistant Strains GO Biological Process

2-MF 5G Resistant Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ESCRT II complex [GO:0000814]	2.29E-06	*VPS25, VPS36, SNF8*	3	3
endosome membrane [GO:0010008]	9.93E-05	*VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	6	57
endosome [GO:0005768]	0.000550	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	7	108
internal side of plasma membrane [GO:0009898]	0.001035	*STP22, RIM8*	2	4
ESCRT III complex [GO:0000815]	0.001035	*SNF7, VPS20*	2	4
cytosolic large ribosomal subunit [GO:0022625]	0.006171	*RPL35B, RPL34A, RPL26B, RPL16A, RPL40B*	5	88
membrane [GO:0016020]	0.007320	*BAP2, TAT1, YBR090C, RTC2, STP22, HSP30, BPH1, ARE1, YET3, YDR282C, STP1, MNN1, ISC1, YCK3, GEP7, TIM21, ANS1, PRM2, YIL067C, PRY3, IML2, YJL132W, VPS25, BCH2, CAF4, SNF7, YLR050C, SUR4, VPS36, VPS20, TGL3, VPS27, SNF8*	33	1671

Table 3.16: 2-MF 5G Resistant Strains GO Cellular Component

2-MF 5G Resistant Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcription repression [11.02.03.04.03]	2.89E-05	*TUP1, RIM101, VPS25, VPS36, SNF8*	5	28
pH response [34.11.03.11]	0.000522	*BPH1, RIM101*	2	3
regulation of C-compound and carbohydrate metabolism [01.05.25]	0.001374	*PCL7, VPS25, RGT1, HAP4, SNF7, VPS36, SNF8*	7	126
vacuolar/lysosomal transport [20.09.13]	0.004132	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	7	153
pH stress response [32.01.04]	0.004667	*BPH1, RIM101*	2	8
lipid, fatty acid and isoprenoid metabolism [01.06]	0.008385	*DPL1, ISC1, ACB1, YJL132W, TGL3, SPS19*	6	133

Table 3.17: 2-MF 5G Resistant Strains MIPS Functional Classification More surprising are the enriched Gene ontologies for cellular components including the ESCRT II complex, endosome membrane, endosome, ESCRT III complex, and general membrane proteins. ESCRT complexes have been linked to yeast survival in conditions with altered glutathione availability[109,110] GO enrichment is also observed in several MIPS functional classifications: transcription repression, pH response, carbohydrate metabolism regulation, pH stress response, and lipid metabolism. These observations run counter to the general mechanisms of protein flux that are required for tolerance to other furans.

2-MF 10G Resistant GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway [GO:0043162]	4.18E-09	*STP22, VPS25, SNF7, VPS36, VPS20, SNF8, VPS28*	7	15
protein processing [GO:0016485]	5.39E-05	*RIM8, RIM13, DFG16, RIM20,*	4	12
response to pH [GO:0009268]	6.44E-05	*BPH1, RIM101, RIM20*	3	5
protein targeting to vacuole [GO:0006623]	0.000106	*STP22, VPS25, VPS36, VPS27, SNF8, VPS28*	6	41
negative regulation of transcription from RNA polymerase II promoter by glucose [GO:0000433]	0.000219	*VPS25, VPS36, SNF8*	3	7
late endosome to vacuole transport [GO:0045324]	0.000468	*STP22, SNF7, VPS20, VPS27*	4	20
cellular response to anoxia [GO:0071454]	0.002080	*RIM101, SNF7*	2	4
protein targeting to membrane [GO:0006612]	0.002551	*AST1, STP22, VPS28*	3	15
intralumenal vesicle formation [GO:0070676]	0.003424	*SNF7, VPS20*	2	5
protein catabolic process [GO:0030163]	0.005142	*BLM10, YSP3, SUE1*	3	19

Table 3.18: 2-MF 10G Resistant GO Biological Process

2-MF 10G Resistant GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
cytosolic large ribosomal subunit [GO:0022625]	5.34E-06	*RPL35B, RPL27B, RPL34A, RPL24A, RPL9A, RPL16A, RPL17B, RPL40B, RPL16B, RPL33B*	10	88
ESCRT II complex [GO:0000814]	6.63E-06	*VPS25, VPS36, SNF8*	3	3
endosome membrane [GO:0010008]	0.000672	*VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	6	57
endosome [GO:0005768]	0.000953	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8, VPS28*	8	108
ESCRT I complex [GO:0000813]	0.002080	*STP22, VPS28*	2	4
internal side of plasma membrane [GO:0009898]	0.002080	*STP22, RIM8*	2	4
ESCRT III complex [GO:0000815]	0.002080	*SNF7, VPS20*	2	4
large ribosomal subunit [GO:0015934]	0.003097	*RPL16A, RPL17B, RPL16B*	3	16
late endosome membrane [GO:0031902]	0.007841	*STP22, GTR2, VPS28*	3	22

Table 3.19: 2-MF 10G Resistant GO Cellular Component

2-MF 10G Resistant MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcription repression [11.02.03.04.03]	0.000155	*TUP1, RIM101, VPS25, VPS36, SNF8*	5	28
vacuolar/lysosomal transport [20.09.13]	0.000601	*GEM1, STP22, PIB2, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8, VPS28*	10	153
pH response [34.11.03.11]	0.001053	*BPH1, RIM101*	2	3
ribosomal proteins [12.01.01]	0.006692	*RPL35B, RPL27B, RPL34A, RPL24A, DBP3, RPL9A, RPL16A, RPL17B, RPL40B, RPL16B, RPL33B*	11	246
pH stress response [32.01.04]	0.009238	*BPH1, RIM101*	2	8
C-compound and carbohydrate metabolism [01.05]	0.009422	*AAD3, GPD1, DLD2, LYS20, MNN1, RGT1, BCH2, YKR096W, GAL80, SUR1*	10	223

Table 3.20: 2-MF 10G Resistant MIPS Functional Classification Significant enrichment is observed in mutant strains lacking genes related to ubiquitin-dependent protein catabolism, the ESCRT complexes I-III, cytosolic large ribosomal subunit, and intracellular transit, particularly endosomal and vacuolar transport. Strains with mutations in genes for transcription repression and pH stress response were also significantly overrepresented compared to control.

2-MF 15G Resistant GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway [GO:0043162]	6.81E-09	*STP22, VPS25, SNF7, VPS36, VPS20, SNF8, VPS28*	7	15
protein processing [GO:0016485]	7.08E-05	*RIM8, RIM13, DFG16, RIM20*	4	12
response to pH [GO:0009268]	7.93E-05	*BPH1, RIM101, RIM20*	3	5
protein targeting to vacuole [GO:0006623]	0.000156	*STP22, VPS25, VPS36, VPS27, SNF8, VPS28*	6	41
negative regulation of transcription from RNA polymerase II promoter by glucose [GO:0000433]	0.000269	*VPS25, VPS36, SNF8*	3	7
late endosome to vacuole transport [GO:0045324]	0.000610	*STP22, SNF7, VPS20, VPS27*	4	20
negative regulation of transcription from RNA polymerase II promoter by pheromones [GO:0046020]	0.000627	*DIG2, ITC1, ISW2*	3	9
chromatin silencing at telomere [GO:0006348]	0.001063	*DPB4, ITC1, DLS1, RTT106, ISW2, HAT1*	6	58
glycosphingolipid biosynthetic process [GO:0006688]	0.001210	*CSG2, SUR1*	2	3
trehalose catabolic process [GO:0005993]	0.001210	*NTH1, ATH1*	2	3
cellular response to anoxia [GO:0071454]	0.002388	*RIM101, SNF7*	2	4
regulation of transcription, DNA-dependent [GO:0006355]	0.002897	*TBS1, TUP1, DPB4, JHD1, TOS8, ITC1, NUT1, RIM101, SKN7, VHR1, DLS1, RGT1, BAS1, CHS5, RTT106, HAL9, HIR2, ISW2, NTO1, ROX1*	20	507
regulation of cell size [GO:0008361]	0.002950	*GPA2, SSF1, SKN7, PTK2*	4	30
sphingolipid biosynthetic process [GO:0030148]	0.003109	*LCB3, SUR4, SUR1*	3	15
intralumenal vesicle formation [GO:0070676]	0.003927	*SNF7, VPS20*	2	5
cellular process [GO:0009987]	0.008029	*MAP2, ARX1*	2	7

Table 3.21: 2-MF 15G Resistant GO Biological Process

2-MF 15G Resistant GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
chromatin accessibility complex [GO:0008623]	1.62E-07	*DPB4, ITC1, DLS1, ISW2*	4	4
cytosolic large ribosomal subunit [GO:0022625]	1.64E-07	*RPL21A, RPL35B, ARX1, RPL27B, RPL34A, RPL29, RPL9A, RPL26B, RPL24B, RPL40B, RPL26A, RPL16B,*	12	88
ESCRT II complex [GO:0000814]	8.18E-06	*VPS25, VPS36, SNF8*	3	3
endosome membrane [GO:0010008]	0.000969	*VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	6	57
endosome [GO:0005768]	0.001495	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8, VPS28*	8	108
internal side of plasma membrane [GO:0009898]	0.002388	*STP22, RIM8*	2	4
ESCRT I complex [GO:0000813]	0.002388	*STP22, VPS28*	2	4
ESCRT III complex [GO:0000815]	0.002388	*SNF7, VPS20*	2	4
large ribosomal subunit [GO:0015934]	0.003771	*RPL26B, RPL26A, RPL16B*	3	16
exomer complex [GO:0034044]	0.003927	*BCH2, CHS5*	2	5
ribosome [GO:0005840]	0.009908	*RPL21A, RPL35B, RPL27B, RPL34A, RPL29, RPL9A, RPL26B, YGR054W, RPL24B, RPL40B, YKR096W, RPL26A, RPL16B*	13	310

Table 3.22: 2-MF 15G Resistant GO Cellular Component

2-MF 15G Resistant MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcription repression [11.02.03.04.03]	1.63E-05	*TUP1, RIM101, VPS25, VPS36, SNF8, ROX1*	6	28
ribosomal proteins [12.01.01]	0.000432	*RPL21A, RPL35B, RPL27B, RPL34A, RPL29, DBP3, RPL9A, RPL26B, YGR054W, RPL24B, SSF1, RPL40B, RPL26A, RPL16B*	14	246
vacuolar/lysosomal transport [20.09.13]	0.001035	*GEM1, STP22, LST7 VPS25, SNF7, VPS36, VPS20, VPS27, SNF8, VPS28*	10	153
pH response [34.11.03.11]	0.001210	*BPH1, RIM101*	2	3

Table 3.23: 2-MF 15G Resistant MIPS Functional Classification Similar to the 10G time point, pathways related to ubiquitin-dependent protein catabolism, intracellular transport, and ribosomal function were significantly enriched. All three ESCRT complexes were enriched. Additional GO enrichment was observed in protein processing, chromatin silencing, sphingolipid biosynthesis, glycosphingolipid biosynthesis, trehalose catabolism, cell size regulation, and various plasma membrane proteins. Mechanistic explanations for this remain elusive.

2-MF Resistant Across All Time Points GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway [GO:0043162]	2.19E-10	*STP22, VPS25, SNF7, VPS36, VPS20, SNF8*	6	15
protein targeting to vacuole [GO:0006623]	5.16E-06	*STP22, VPS25, VPS36, VPS27, SNF8*	5	41
late endosome to vacuole transport [GO:0045324]	6.37E-06	*STP22, SNF7, VPS20, VPS27*	4	20
negative regulation of transcription from RNA polymerase II promoter by glucose [GO:0000433]	8.23E-06	*VPS25, VPS36, SNF8*	3	7
cellular response to anoxia [GO:0071454]	0.000235	*RIM101, SNF7*	2	4
response to pH [GO:0009268]	0.000390	*BPH1, RIM101*	2	5
intralumenal vesicle formation [GO:0070676]	0.000390	*SNF7, VPS20*	2	5
negative regulation of transcription, DNA-dependent [GO:0045892]	0.001721	*TUP1, RGT1*	2	10
protein retention in Golgi apparatus [GO:0045053]	0.002094	*VPS36, VPS27*	2	11
protein transport [GO:0015031]	0.002305	*STP22, YET3, VPS25, BCH2, SNF7, VPS36, VPS20, SNF8*	8	379
protein processing [GO:0016485]	0.002503	*RIM8, RIM13*	2	12

Table 3.24: 2-MF Resistant Across All Time Points GO Biological Process

2-MF Resistant Across All Time Points GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
ESCRT II complex [GO:0000814]	2.39E-07	*VPS25, VPS36, SNF8*	3	3
endosome membrane [GO:0010008]	1.30E-06	*VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	6	57
endosome [GO:0005768]	4.34E-06	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	7	108
ESCRT III complex [GO:0000815]	0.000235	*SNF7, VPS20*	2	4
internal side of plasma membrane [GO:0009898]	0.000235	*STP22, RIM8*	2	4

Table 3.25: 2-MF Resistant Across All Time Points GO Cellular Component

2-MF Resistant across all time points MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcription repression [11.02.03.04.03]	7.19E-07	*TUP1, RIM101, VPS25, VPS36, SNF8*	5	28
vacuolar/lysosomal transport [20.09.13]	4.27E-05	*STP22, VPS25, SNF7, VPS36, VPS20, VPS27, SNF8*	7	153
pH response [34.11.03.11]	0.000118	*BPH1, RIM101*	2	3
pH stress response [32.01.04]	0.001079	*BPH1, RIM101*	2	8
regulation of C-compound and carbohydrate metabolism [01.05.25]	0.001129	*VPS25, RGT1, SNF7, VPS36, SNF8*	5	126
development of asco- basidio- or zygospore [43.01.03.09]	0.003790	*ARE1, GPA2, RIM101, SNF7, RIM13*	5	166
protein processing (proteolytic) [14.07.11]	0.007292	*MAP2, RIM8, RIM13*	3	63
protein targeting, sorting and translocation [14.04]	0.008120	*STP22, VPS25, SNF7, VPS36, VPS27, SNF8*	6	281

Table 3.26: 2-MF Resistant across all time points MIPS Functional Classification in contrast to other test compounds, mutants for genes related to protein synthesis and degradation confers tolerance to 2-MF across all time points. Additional GO enrichment was observed in negative regulation of RNA Pol II (as opposed to in sensitive strains, in which loss of positive RNA Pol II regulation reduced fitness), anoxic response, pH response, and Golgi-related transport pathways. Loss of transcription repression and loss of spore-formation pathways also appears to improve resistance to 2-MF.

2-MF 5G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
CVT pathway [GO:0032258]	0.000346	*IRS4, ATG16, COG6, COG5*	4	37
negative regulation of transcription from RNA polymerase I promoter [GO:0016479]	0.001229	*SAP30, PHO23*	2	6
glutathione metabolic process [GO:0006749]	0.003602	*GTT2, GTO3*	2	10
response to arsenic-containing substance [GO:0046685]	0.004376	*RPN4, YAP1*	2	11
negative regulation of chromatin silencing at telomere [GO:0031939]	0.004376	*SAP30, PHO23*	2	11
Ras protein signal transduction [GO:0007265]	0.006133	*CYR1, RAS2*	2	13

Table 3.27: 2-MF 5G Sensitive Strains GO Biological Process

2-MF 5G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
histone deacetylase complex [GO:0000118]	0.001711	*SAP30, PHO23*	2	7
Golgi transport complex [GO:0017119]	0.002268	*COG6, COG5*	2	8
Rpd3L complex [GO:0033698]	0.007113	*SAP30, PHO23*	2	14

Table 3.28: 2-MF 5G Sensitive Strains GO Cellular Component

2-MF 5G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
homeostasis of phosphate [34.01.03.03]	0.002898	*NPP1, PHO84*	2	9
intra Golgi transport [20.09.07.05]	0.003360	*SEC22, COG6, COG5*	3	33
nuclear division [10.03.04.07]	0.004376	*TOM1, DBF2*	2	11
cAMP/cGMP mediated signal transduction [30.01.09.07]	0.005220	*CYR1, RAS2*	2	12

Table 3.29: 2-MF 5G Sensitive Strains MIPS Functional Classification

2-MF 10G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
negative regulation of transcription from RNA polymerase I promoter [GO:0016479]	0.000159	*SAP30, PHO23, UME1*	3	6
mRNA export from nucleus [GO:0006406]	0.000213	*NUP170, NPL3, SGF73, THP2, NUP2, MFT1*	6	43
apoptosis [GO:0006915]	0.000918	*OYE2, FIS1, CPR3, NMA111*	4	22
negative regulation of chromatin silencing at telomere [GO:0031939]	0.001222	*SAP30, PHO23, UME1*	3	11
purine ribonucleoside monophosphate biosynthetic process [GO:0009168]	0.002423	*YBR284W, AMD1*	2	4
cellular carbohydrate metabolic process [GO:0044262]	0.004607	*MAL33, IMP2', YMR099C*	3	17
response to heat [GO:0009408]	0.004607	*NBP2, YAP1, WHI2*	3	17
ascospore formation [GO:0030437]	0.005926	*SPT3, NEM1, VID28, RAS2, MCK1*	5	57
chromatin modification [GO:0016568]	0.008568	*SPT3, SGF73, SPT8, SAP30, PHO23, EAF7, UME1*	7	114

Table 3.30: 2-MF 10G Sensitive Strains GO Biological Process

2-MF 10G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
nucleoplasmic THO complex [GO:0000446]	0.002423	*THP2, MFT1*	2	4
THO complex part of transcription export complex [GO:0000445]	0.002423	*THP2, MFT1*	2	4
Rpd3L complex [GO:0033698]	0.002579	*SAP30, PHO23, UME1*	3	14
Rpd3L-Expanded complex [GO:0070210]	0.006372	*SAP30, PHO23, UME1*	3	19
SAGA complex [GO:0000124]	0.007385	*SPT3, SGF73, SPT8*	3	20
histone deacetylase complex [GO:0000118]	0.008146	*SAP30, PHO23*	2	7

Table 3.31: 2-MF 10G Sensitive Strains GO Cellular Component

2-MF 10G Sensitive Strains GO MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
modification by acetylation, deacetylation [14.07.04]	0.000479	*SPT3, MAK10, SGF73, SPT8, SAP30, PHO23, EAF7*	7	69
regulator of transcription factor [18.02.09]	0.000854	*SPT3, OPI1, IMP2', SPT8, UME1*	5	37
cAMP/cGMP mediated signal transduction [30.01.09.07]	0.001606	*SOK1, CYR1, RAS2*	3	12
development of asco- basidio- or zygospore [43.01.03.09]	0.006740	*FEN1, TRS85, SPT3, NEM1, CYR1, SAP30, RAS2, SLZ1, MCK1*	9	166

Table 3.32: 2-MF 10G Sensitive Strains GO MIPS Functional Classification GO enrichment is observed in pathways and cellular components for genes related to ribonucleotide processing (mRNA export from the nucleus, negative regulation of transcription from RNA Pol I promoter, purine ribonucleotide monophosphate biosynthesis), oxidative stress, an chromatin structure maintenance (negative regulation of chromatin silencing at telomere, chromatin modification). Genes involved in acylation were also significantly enriched.

2-MF 15G Sensitive GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
positive regulation of transcription from RNA polymerase II promoter [GO:0045944]	0.000393	*OPI1, VID28, SUB1, SAP30. PHO23*	5	100
ascospore formation [GO:0030437]	0.000438	*SPT3, NEM1, VID28, MCK1*	4	57
negative regulation of transcription from RNA polymerase I promoter [GO:0016479]	0.000583	*SAP30, PHO23*	2	6
regulation of transcription from RNA polymerase II promoter in response to stress [GO:0043618]	0.001382	*MSN2, SUB1*	2	9
negative regulation of chromatin silencing at telomere [GO:0031939]	0.002095	*SAP30, PHO23*	2	11
response to stress [GO:0006950]	0.002590	*MSN2, OCA1, MCK1, IRA2, WHI2*	5	152
M phase of mitotic cell cycle [GO:0000087]	0.002299	*NAP1, MIH1*	2	3
metal ion transport [GO:0030001]	0.006271	*BSD2, PCA1, COX19*	3	14
cAMP-mediated signaling [GO:0019933]	0.007384	*SOK1, PDE2*	2	5
protein insertion into ER membrane [GO:0045048]	0.007384	*GET2, GET1*	2	5

Table 3.33: 2-MF 15G Sensitive GO Biological Process

2-MF 15G Sensitive GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
histone deacetylase complex [GO:0000118]	0.000813	*SAP30, PHO23*	2	7
NatC complex [GO:0031417]	0.002299	*MAK10, MAK3*	2	3
GET complex [GO:0043529]	0.002299	*GET2, GET1*	2	3
retromer complex, inner shell [GO:0030906]	0.002299	*PEP8, VPS35*	2	3
nuclear envelope [GO:0005635]	0.007292	*RRT12, OPI1, WSS1*	3	63
nuclear membrane [GO:0031965]	0.008293	*NUP170, OPI1, NEM1*	3	66
Golgi apparatus [GO:0005794]	0.006613	*ATG15, ARF1, TCA17, GET2, GET1, SYS1, GEF1, SEC22, VPS38, NPR1, TLG2, GYP1, LDB19*	13	213
retromer complex [GO:0030904]	0.007384	*PEP8, VPS35*	2	5

Table 3.34: 2-MF 15G Sensitive GO Cellular Component

2-MF 15G Sensitive MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcriptional control [11.02.03.04]	0.000255	*SPT3, PGD1, DBF2, OPI1, ZAP1, DAT1, MSN2, SUB1, SAP30, PHO23*	10	426
regulator of transcription factor [18.02.09]	0.001600	*SPT3 OPI1 SUB1*	3	37
metabolism of cyclic and unusual nucleotides [01.03.10]	0.005057	*PUS2 IRA2*	2	17
regulation of glycolysis and gluconeogenesis [02.01.03]	0.006307	*VID28 FBP26*	2	19
modification by acetylation, deacetylation [14.07.04]	0.009371	*SPT3 SAP30 PHO23*	3	69

Table 3.35: 2-MF 15G Sensitive MIPS Functional Classification surprisingly few GO-enriched pathways and cell components compared to other compounds at the 15G time point. Primarily, sensitivity genes are clustered in pathways for RNA Polymerase I and II transcriptional regulation, cell stress response, cell cycle, cell signaling pathways, and cell cycle. Loss of some structural components of the Golgi and the nucleus were found to reduce tolerance to 2-MF. A unifying mechanistic explanation for these observations is not forthcoming.

2-MF Sensitive across all time points GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
negative regulation of transcription from RNA polymerase I promoter [GO:0016479]	0.000239	*SAP30, PHO23*	2	6
response to arsenic-containing substance [GO:0046685]	0.000865	*RPN4, YAP1*	2	11
negative regulation of chromatin silencing at telomere [GO:0031939]	0.000865	*SAP30, PHO23*	2	11
regulation of transcription, DNA-dependent [GO:0006355]	0.003458	*RPN4, TOM1, SPT8, YAP1, MFT1, SAP30, PHO23*	7	507
transcription, DNA-dependent [GO:0006351]	0.004919	*RPN4, TOM1, SPT8, YAP1, MFT1, SAP30, PHO23*	7	540
histone deacetylation [GO:0016575]	0.005306	*SAP30, PHO23*	2	27
regulation of cell size [GO:0008361]	0.006527	*YCR061W, TOM1*	2	30
positive regulation of transcription from RNA polymerase II promoter [GO:0045944]	0.007573	*SPT8, SAP30, PHO23*	3	100

Table 3.36; 2-MF Sensitive across all time points GO Biological Process

2-MF Sensitive across all time points GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
histone deacetylase complex [GO:0000118]	0.000333	*SAP30, PHO23*	2	7
Rpd3L complex [GO:0033698]	0.001421	*SAP30, PHO23*	2	14
Rpd3L-Expanded complex [GO:0070210]	0.002638	*SAP30, PHO23*	2	19

Table 3.37: 2-MF Sensitive across all time points GO Cellular Component

2-MF Sensitive across all time points MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
nuclear division [10.03.04.07]	0.000865	*TOM1, DBF2*	2	11
modification by acetylation, deacetylation [14.07.04]	0.002669	*SPT8, SAP30, PHO23*	3	69
DNA conformation modification (e.g. chromatin) [10.01.09.05]	0.006686	*IRS4, SPT8, SAP30, PHO23*	4	188

Table 3.38: 2-MF Sensitive across all time points MIPS Functional Classification As noted before, significantly fewer genes conferred susceptibility to 2-MF compared to other strains. Strains with mutations for genes in pathways for positive regulation of transcription from RNA Pol II promoter, histone deacetylation, and DNA packaging/modification pathways were more sensitive across all time points for 2-MF.

2-EF 5G Resistant Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
vesicular transport (Golgi network, etc.) [20.09.07]	0.008462	*CHS6, VPS17*	2	72

Table 3.39: 2-EF 5G Resistant Strains MIPS Functional Classification

2-EF 5G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
response to singlet oxygen [GO:0000304]	2.50E-05	*SKN7, YAP1*	2	4
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	2.50E-05	*SKN7, YAP1*	2	4
CVT pathway [GO:0032258]	5.65E-05	*IRS4, COG6, COG5*	3	37
response to arsenic-containing substance [GO:0046685]	0.000227132	*RPN4, YAP1*	2	11
positive regulation of transcription, DNA-dependent [GO:0045893]	0.000776125	*RPN4, STP1*	2	20
intra-Golgi vesicle-mediated transport [GO:0006891]	0.00112197	*COG6, COG5*	2	24

Table 3.40: 2-EF 5G Sensitive Strains GO Biological Process

2-EF 5G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
Golgi transport complex [GO:0017119]	0.000116	*COG6, COG5*	2	8
Golgi membrane [GO:0000139]	0.001711	*SYS1, COG6, COG5*	3	117
Golgi apparatus [GO:0005794]	0.009259	*SYS1, COG6, COG5*	3	213

Table 3.41: 2-EF 5G Sensitive Strains GO Cellular Component

2-EF 5G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
transcription activation [11.02.03.04.01]	8.30E-05	*RPN4, THI3, ACE2*	3	42
cellular signalling [30.01]	0.002123	*SKN7, IRS4*	2	33
intra Golgi transport [20.09.07.05]	0.002123	*COG6, COG5*	2	33
oxidative stress response [32.01.01]	0.005814	*SKN7, YAP1*	2	55
cell wall [42.01]	0.009259	*RAD23, INP51, IRS4*	3	213

Table 3.42: 2-EF 5G Sensitive Strains MIPS Functional Classification

2-EF 10G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	3.29E-05	*SKN7, YAP1*	2	4
response to singlet oxygen [GO:0000304]	3.29E-05	*SKN7, YAP1*	2	4
CVT pathway [GO:0032258]	8.63E-05	*IRS4, COG8, COG6*	3	37
intra-Golgi vesicle-mediated transport [GO:0006891]	0.001472	*COG8, COG6*	2	24

Table 3.43: 2-EF 10G Sensitive Strains GO Biological Process

2-EF 10G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
Golgi transport complex [GO:0017119]	0.000152	*COG8, COG6*	2	8
Golgi membrane [GO:0000139]	0.002565	*SYS1, COG8, COG6*	3	117

Table 3.44: 2-EF 10G Sensitive Strains GO Cellular Component

2-EF 10G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
intra Golgi transport [20.09.07.05]	0.002782	*COG8, COG6*	2	33
cellular signalling [30.01]	0.002782	*SKN7, IRS4*	2	33
metabolism of secondary products derived from primary amino acids [01.20.17]	0.007252	*MAK10*	1	3
oxidative stress response [32.01.01]	0.007585	*SKN7, YAP1*	2	55

Table 3.45: 2-EF 10G Sensitive Strains MIPS Functional Classification Although sixteen mutants were significantly underrepresented at this time point, three of those lacked proteins whose functions are unknown. Pathway enrichment for sensitive mutants is very similar to the 5G time point for 2-EF exposure, except that the fold-change is more severe, and the *STP1Δ* strain appears as well. The Stp1 is an important transcription factor for pathways related to import of extracellular amino acids.

2-EF 15G Sensitive Strains GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
CVT pathway [GO:0032258]	0.000147	*TRS85, SNX4, IRS4, TLG2*	4	37
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	0.000320	*SKN7, YAP1*	2	4
response to singlet oxygen [GO:0000304]	0.000320	*SKN7 ,YAP1*	2	4
endocytosis [GO:0006897]	0.000324	*DRS2, THR4, TLG2, WHI2, SNC2*	5	82
vesicle-mediated transport [GO:0016192]	0.000538	*FEN1, TRS85, GET1, TLG2, GYP1, SNC2*	6	140
ascospore formation [GO:0030437]	0.000792	*ERV14, NEM1, VID28, VPS13*	4	57
mitochondrion degradation [GO:0000422]	0.001225	*PTC6, SNX4, WHI2*	3	29
response to arsenic-containing substance [GO:0046685]	0.002843	*RPN4, YAP1*	2	11
vacuolar protein catabolic process [GO:0007039]	0.003395	*VID28, VPS13*	2	12
vesicle organization [GO:0016050]	0.003395	*TRS85, SYS1*	2	12
response to heat [GO:0009408]	0.006833	*YAP1, WHI2*	2	17

Table 3.46: 2-EF 15G Sensitive Strains GO Biological Process

2-EF 15G Sensitive Strains GO Cellular Component				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
trans-Golgi network [GO:0005802]	8.68E-07	*DRS2, TRS65, SYS1, TLG2, SNC2*	5	25
Golgi apparatus [GO:0005794]	2.14E-05	*DRS2, TRS85, GET1, ERV14, TRS65 SYS1, VPS13, TLG2, GYP1*	9	213
pre-autophagosomal structure [GO:0000407]	0.000612	*TRS85, SNX4, IRS4*	3	23
SNARE complex [GO:0031201]	0.000787	*MSO1, TLG2, SNC2*	3	25
TRAPP complex [GO:0030008]	0.002843	*TRS85, TRS65*	2	11
integral to endosome membrane [GO:0031303]	0.007420	*TLG2*	1	1
endosome [GO:0005768]	0.008151	*SNX4, VPS13, TLG2, SNC2*	4	108
endoplasmic reticulum membrane [GO:0005789]	0.008451	*FEN1, YPS7, GET1, CWH41, ERV14, NEM1, RCE1*	7	318

Table 3.47: 2-EF 15G Sensitive Strains GO Cellular Component

2-EF 15G Sensitive Strains MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
protein/peptide degradation [14.13]	0.000377	*RRT12, AFG3, SNX4, RCE1*	4	47
transcription activation [11.02.03.04.01]	0.003598	*RPN4, ACE2, MSN1*	3	42
development of asco- basidio- or zygospore [43.01.03.09]	0.007381	*FEN1, TRS85, ERV14, NEM1, MSO1*	5	166
endocytosis [20.09.18.09.01]	0.009330	*THR4, WHI2, SNC2*	3	59

Table 3.48: 2-EF 15G Sensitive Strains MIPS Functional Classification Oxidative stress response pathways as well as Golgi structure and trafficking components are once again the most significantly enriched GO pathways, cellular components, and functional classifications. At this time point, the same three oxidative stress genes were activated as in the other time points, but a larger and different suite of Golgi proteins were highlighted along with a smattering of genes for ER structure. Nine mutants for genes related to the structure of the Golgi apparatus were significantly less abundant, as well as roughly half a dozen genes for trans-Golgi network (TGN) vesicle trafficking including SNARE and TRAPP complexes.

2-EF Sensitive Strains Across All Time Points GO Molecular Function				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
sequence-specific DNA binding [GO:0043565]	0.000289	*SKN7, ACE2, YAP1*	3	165
DNA binding [GO:0003677]	0.005349	*SKN7, ACE2, YAP1*	3	449
sequence-specific DNA binding transcription factor activity [GO:0003700]	0.006156	*SKN7, YAP1*	2	138

Table 3.49: 2-EF Sensitive Strains Across All Time Points GO Molecular Function

2-EF Sensitive Strains Across All Time Points GO Biological Process				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
regulation of transcription from RNA polymerase II promoter in response to oxidative stress [GO:0043619]	4.13E-06	SKN7 YAP1	2	4
response to singlet oxygen [GO:0000304]	4.13E-06	SKN7 YAP1	2	4

Table 3.50: 2-EF Sensitive Strains Across All Time Points GO Biological Process

2-EF Sensitive Strains Across All Time Points MIPS Functional Classification				
Category	**p-value**	**In Category from Cluster**	**k**	**f**
cellular signaling [30.01]	0.000358	SKN7 IRS4	2	33
oxidative stress response [32.01.01]	0.001000	SKN7 YAP1	2	55

Table 3.51: 2-EF Sensitive Strains Across All Time Points MIPS Functional Classification Mutants for genes related to oxidative stress response and amino acid uptake were sensitive across all time points and showed decreased fitness when exposed to 2-EF. This observation, coupled by the observed increased potency of 2-EF suggests that the compound and its likely metabolites cause oxidative damage to the cell. That mutants for genes involved in protein synthesis and recycling were also consistently sensitive to 2-EF suggests that the targets of 2-EF's oxidative attack are cell proteins and that the damage is non-specific. Since the pathways involved in protein flux are consistently sensitive to 2-EF, but few genes are consistently sensitive across all time points.

References

1. Wei, H. *et al.* Experimental investigation on the combustion and emissions characteristics of 2-methylfuran gasoline blend fuel in spark-ignition engine. *Appl. Energy* **132,** 317–324 (2014).
2. Costagliola, M. A., De Simio, L., Iannaccone, S. & Prati, M. V. Combustion efficiency and engine out emissions of a S.I. engine fueled with alcohol/gasoline blends. *Appl. Energy* **111,** 1162–1171 (2013).
3. Qian, Y., Zhu, L., Wang, Y. & Lu, X. Recent progress in the development of biofuel 2,5-dimethylfuran. *Renew. Sustain. Energy Rev.* **41,** 633–646 (2015).
4. Yildiz, B., Hagos, M., De, S., Kim, J. & Daim, T. Technology Forecasting: Case of Electric Vehicle Technology. in *Research and Development Management* 125–136 (Springer, Cham, 2017). doi:10.1007/978-3-319-54537-0_8
5. Jansson, J., Pettersson, T., Mannberg, A., Bränlund, R. & Lindgren, U. Adoption of alternative fuel vehicles: Influence from neighbors, family and coworkers. *Transp. Res. Part Transp. Environ.* **54,** 61–73 (2017).
6. Adepetu, A. & Keshav, S. The relative importance of price and driving range on electric vehicle adoption: Los Angeles case study. *Transportation* **44,** 353–373 (2017).
7. Olivetti, E. A., Ceder, G., Gaustad, G. G. & Fu, X. Lithium-Ion Battery Supply Chain Considerations: Analysis of Potential Bottlenecks in Critical Metals. *Joule* **1,** 229–243 (2017).
8. Feller, D. & Simmie, J. M. High-Level ab Initio Enthalpies of Formation of 2,5-Dimethylfuran, 2-Methylfuran, and Furan. *J. Phys. Chem. A* **116,** 11768–11775 (2012).
9. Fromowitz, M. *et al.* Bone marrow genotoxicity of 2,5-dimethylfuran, a green biofuel candidate. *Environ. Mol. Mutagen.* **53,** 488–491 (2012).
10. Phuong, J., Kim, S., Thomas, R. & Zhang, L. Predicted toxicity of the biofuel candidate 2,5-dimethylfuran in environmental and biological systems. *Environ. Mol. Mutagen.* **53,** 478–487 (2012).
11. Simmie, J. M. & Würmel, J. Harmonising Production, Properties and Environmental Consequences of Liquid Transport Fuels from Biomass-2,5-Dimethylfuran as a Case Study. *ChemSusChem* **6,** 36–41 (2013).
12. NTP, (National Toxicology Program). *Report on Carcinogens, Thirteenth Edition*. (Departmenth of Health And Human Services, Public Health Service, 2014).
13. *Dry Cleaning, Some Chlorinated Solvents And Other Industrial Chemicals*. 393–407 (World Health Organization International Agency For Research On Cancer).
14. Taxak, N., Kalra, S. & Bharatam, P. V. Mechanism-Based Inactivation of Cytochromes by Furan Epoxide: Unraveling the Molecular Mechanism. *Inorg. Chem.* **52,** 13496–13508 (2013).
15. Gates, L. A., Lu, D. & Peterson, L. A. Trapping of cis-2-Butene-1,4-dial to Measure Furan Metabolism in Human Liver Microsomes by Cytochrome P450 Enzymes. *Drug Metab. Dispos.* **40,** 596–601 (2012).
16. Peterson, L. A. Reactive Metabolites in the Biotransformation of Molecules Containing a Furan Ring. *Chem. Res. Toxicol.* **26,** 6–25 (2013).
17. Phillips, M. B., Sullivan, M. M., Villalta, P. W. & Peterson, L. A. Covalent Modification of Cytochrome *c* by Reactive Metabolites of Furan. *Chem. Res. Toxicol.* **27,** 129–135 (2014).

18. Moro, S. *et al.* Identification and Pathway Mapping of Furan Target Proteins Reveal Mitochondrial Energy Production and Redox Regulation as Critical Targets of Furan Toxicity. *Toxicol. Sci.* **126,** 336–352 (2012).
19. Peterson, L. A., Naruko, K. C. & Predecki, D. P. A Reactive Metabolite of Furan, *cis* -2-Butene-1,4-dial, Is Mutagenic in the Ames Assay. *Chem. Res. Toxicol.* **13,** 531–534 (2000).
20. Kellert, M., Brink, A., Richter, I., Schlatter, J. & Lutz, W. K. Tests for genotoxicity and mutagenicity of furan and its metabolite cis-2-butene-1,4-dial in L5178Y tk+/− mouse lymphoma cells. *Mutat. Res. Toxicol. Environ. Mutagen.* **657,** 127–132 (2008).
21. Moro, S. *et al.* Furan in heat-treated foods: Formation, exposure, toxicity, and aspects of risk assessment. *Mol. Nutr. Food Res.* **56,** 1197–1211 (2012).
22. Durling, L., Svensson, K. & Abramssonzetterberg, L. Furan is not genotoxic in the micronucleus assay in vivo or in vitro. *Toxicol. Lett.* **169,** 43–50 (2007).
23. Jeffrey, A. M., Brunnemann, K. D., Duan, J.-D., Schlatter, J. & Williams, G. M. Furan induction of DNA cross-linking and strand breaks in turkey fetal liver in comparison to 1,3-propanediol. *Food Chem. Toxicol.* **50,** 675–678 (2012).
24. McDaniel, L. P. *et al.* Genotoxicity of furan in Big Blue rats. *Mutat. Res. Toxicol. Environ. Mutagen.* **742,** 72–78 (2012).
25. Ding, W. *et al.* In vivo genotoxicity of furan in F344 rats at cancer bioassay doses. *Toxicol. Appl. Pharmacol.* **261,** 164–171 (2012).
26. Neuwirth, C. *et al.* Furan carcinogenicity: DNA binding and genotoxicity of furan in rats in vivo. *Mol. Nutr. Food Res.* **56,** 1363–1374 (2012).
27. Terrell, A. N. *et al.* Mutagenicity of furan in female Big Blue B6C3F1 mice. *Mutat. Res. Toxicol. Environ. Mutagen.* **770,** 46–54 (2014).
28. Mariotti, M. S., Granby, K., Rozowski, J. & Pedreschi, F. Furan: a critical heat induced dietary contaminant. *Food Funct.* **4,** 1001 (2013).
29. Ravindranath, V. & Boyd, M. R. Metabolic activation of 2-methylfuran by rat microsomal systems. *Toxicol. Appl. Pharmacol.* **78,** 370–376 (1985).
30. Mochalski, P. *et al.* Blood and breath levels of selected volatile organic compounds in healthy volunteers. *The Analyst* **138,** 2134 (2013).
31. Perbellini, L., Princivalle, A., Cerpelloni, M., Pasini, F. & Brugnone, F. Comparison of breath, blood and urine concentrations in the biomonitoring of environmental exposure to 1,3-butadiene, 2,5-dimethylfuran, and benzene. *Int. Arch. Occup. Environ. Health* **76,** 461–466 (2003).
32. Peterson, L. A. Electrophilic Intermediates Produced by Bioactivation of Furan*. *Drug Metab. Rev.* **38,** 615–626 (2006).
33. Lu, D. & Peterson, L. A. Identification of Furan Metabolites Derived from Cysteine− *cis* -2-Butene-1,4-dial−Lysine Cross-Links. *Chem. Res. Toxicol.* **23,** 142–151 (2010).
34. Zeiger, E., Anderson, B., Haworth, S., Lawlor, T. & Mortelmans, K. Salmonella mutagenicity tests: V. Results from the testing of 311 chemicals. *Environ. Mol. Mutagen.* **19,** 2–141 (1992).
35. Couri, D. & Milks, M. Toxicity and metabolism of the neurotoxic hexacarbons n-hexane, 2-hexanone, and 2,5-hexanedione. *Annu. Rev. Pharmacol. Toxicol.* **22,** 145–166 (1982).
36. Kovacic, P. & Somanathan, R. Nervous About Developments in Electron Transfer-Reactive Oxygen Species-Oxidative Stress Mechanisms of Neurotoxicity? in *Systems*

Biology of Free Radicals and Antioxidants (ed. Laher, I.) 1925–1944 (Springer Berlin Heidelberg, 2014).
37. LoPachin, R. M. & Gavin, T. Toxic neuropathies: Mechanistic insights based on a chemical perspective. *Neurosci. Lett.* (2014). doi:10.1016/j.neulet.2014.08.054
38. Wang, K., Zheng, L., Peng, Y., Song, J. & Zheng, J. Selective and Sensitive Platform for Function-Based Screening of Potentially Harmful Furans. *Anal. Chem.* **86,** 10755–10762 (2014).
39. Iwasaki, K. & Tsuruta, H. Molecular mechanism of hexane neuropathy: Significant differences in pharmacokinetics between 2.3-, 2.4-, and 2.5-hexanedione. *Ind. Health* **22,** 177–187 (1984).
40. Kirkland, D. *et al.* Can in vitro mammalian cell genotoxicity test results be used to complement positive results in the Ames test and help predict carcinogenic or in vivo genotoxic activity? I. Reports of individual databases presented at an EURL ECVAM Workshop. *Mutat. Res. Toxicol. Environ. Mutagen.* **775–776,** 55–68 (2014).
41. Ellis, P. *et al.* Where will genetic toxicology testing be in 30 years' time? Summary report of the 25th Industrial Genotoxicity Group Meeting, Royal Society of Medicine, London, November 9, 2011. *Mutagenesis* **29,** 73–77 (2014).
42. Fowler, P. *et al.* Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type. *Mutat. Res. Toxicol. Environ. Mutagen.* **742,** 11–25 (2012).
43. Wilson, M. P. & Schwarzman, M. R. Toward a New U.S. Chemicals Policy: Rebuilding the Foundation to Advance New Science, Green Chemistry, and Environmental Health. *Environ. Health Perspect.* **117,** 1202–1209 (2009).
44. Waters, M. D., Jackson, M. & Lea, I. Characterizing and predicting carcinogenicity and mode of action using conventional and toxicogenomics methods. *Mutat. Res. Mutat. Res.* **705,** 184–200 (2010).
45. Chen, M., Zhang, M., Borlak, J. & Tong, W. A Decade of Toxicogenomic Research and Its Contribution to Toxicological Science. *Toxicol. Sci.* **130,** 217–228 (2012).
46. Gusenleitner, D. *et al.* Genomic Models of Short-Term Exposure Accurately Predict Long-Term Chemical Carcinogenicity and Identify Putative Mechanisms of Action. *PLoS ONE* **9,** e102579 (2014).
47. Browse - California Code of Regulations. Available at: https://govt.westlaw.com/calregs/Browse/Home/California/CaliforniaCodeofRegulations?guid=I6E0E45C032A411E186A4EF11E7983D17&originationContext=documenttoc&transitionType=Default&contextData=(sc.Default). (Accessed: 16th November 2017)
48. Fatehullah, A., Tan, S. H. & Barker, N. Organoids as an in vitro model of human development and disease. *Nat. Cell Biol.* **18,** 246–254 (2016).
49. Bongartz, R. *et al.* Living Cell Microarrays: An Overview of Concepts. *Microarrays* **5,** 11 (2016).
50. Winzeler, E. A. Functional Characterization of the S. cerevisiae Genome by Gene Deletion and Parallel Analysis. *Science* **285,** 901–906 (1999).
51. Giaever, G. *et al.* Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418,** 387–391 (2002).
52. Lee, A. Y. *et al.* Mapping the Cellular Response to Small Molecules Using Chemogenomic Fitness Signatures. *Science* **344,** 208–211 (2014).

53. Hoepfner, D. *et al.* High-resolution chemical dissection of a model eukaryote reveals targets, pathways and gene functions. *Microbiol. Res.* **169,** 107–120 (2014).
54. North, M. *et al.* Genome-Wide Functional Profiling Identifies Genes and Processes Important for Zinc-Limited Growth of Saccharomyces cerevisiae. *PLoS Genet.* **8,** e1002699 (2012).
55. Giaever, G. & Nislow, C. The Yeast Deletion Collection: A Decade of Functional Genomics. *Genetics* **197,** 451–465 (2014).
56. Gaytán, B. D. & Vulpe, C. D. Functional toxicology: tools to advance the future of toxicity testing. *Front. Genet.* **5,** (2014).
57. Steinmetz, L. M. *et al.* Systematic screen for human disease genes in yeast. *Nat. Genet.* (2002). doi:10.1038/ng929
58. McHale, C. M., Smith, M. T. & Zhang, L. Application of toxicogenomic profiling to evaluate effects of benzene and formaldehyde: from yeast to human: Yeast and human toxicogenomic approaches. *Ann. N. Y. Acad. Sci.* **1310,** 74–83 (2014).
59. Mahadevan, B. *et al.* Genetic toxicology in the 21st century: Reflections and future directions. *Environ. Mol. Mutagen.* **52,** 339–354 (2011).
60. McHale, C. M., Zhang, L., Hubbard, A. E. & Smith, M. T. Toxicogenomic profiling of chemically exposed humans in risk assessment. *Mutat. Res. Mutat. Res.* **705,** 172–183 (2010).
61. Li, H.-H., Aubrecht, J. & Fornace, A. J. Toxicogenomics: Overview and potential applications for the study of non-covalent DNA interacting chemicals. *Mutat. Res. Mol. Mech. Mutagen.* **623,** 98–108 (2007).
62. Ltd, T. G. Microplate Readers. Available at: https://lifesciences.tecan.com/microplate-readers. (Accessed: 17th November 2017)
63. Pierce, S. E., Davis, R. W., Nislow, C. & Giaever, G. Genome-wide analysis of barcoded Saccharomyces cerevisiae gene-deletion mutants in pooled cultures. *Nat. Protoc.* **2,** 2958–2974 (2007).
64. 95401.H1POOL (yeast deletion pools) - Thermo Fisher Scientific. Available at: https://www.thermofisher.com/order/catalog/product/95401.H1POOL. (Accessed: 17th November 2017)
65. YeaStar™ Genomic DNA Kit - Bacterial & Fungal DNA - Microbial & Environmental DNA Isolation - DNA. Available at: https://www.zymoresearch.com/dna/microbial-environmental-dna-isolation-1/bacterial-fungal-dna/yeastar-genomic-dna-kit. (Accessed: 17th November 2017)
66. Robinson, D. G., Chen, W., Storey, J. D. & Gresham, D. Design and Analysis of Bar-seq Experiments. *G358 GenesGenomesGenetics* **4,** 11–18 (2014).
67. Platinum PCR SuperMix High Fidelity - Thermo Fisher Scientific. Available at: https://www.thermofisher.com/order/catalog/product/12532016. (Accessed: 18th November 2017)
68. Jo, W. J. *et al.* Comparative Functional Genomic Analysis Identifies Distinct and Overlapping Sets of Genes Required for Resistance to Monomethylarsonous Acid (MMAIII) and Arsenite (AsIII) in Yeast. *Toxicol. Sci.* **111,** 424–436 (2009).
69. Robinson, M. D., Grigull, J., Mohammad, N. & Hughes, T. R. FunSpec: a web-based cluster interpreter for yeast. *BMC Bioinformatics* **3,** 35 (2002).
70. AmiGO 2: Welcome. Available at: http://amigo.geneontology.org/amigo/landing. (Accessed: 20th November 2017)

71. Saccharomyces Genome Database | SGD. Available at: https://www.yeastgenome.org/. (Accessed: 20th November 2017)
72. Abolaji, A. O. *et al.* Exposure to 2,5-hexanedione is accompanied by ovarian and uterine oxidative stress and disruption of endocrine balance in rats. *Drug Chem. Toxicol.* **38,** 400–407 (2015).
73. Stanford, D. R. *et al.* Division of Labor Among the Yeast Sol Proteins Implicated in tRNA Nuclear Export and Carbohydrate Metabolism. *Genetics* **168,** 117–127 (2004).
74. Rape, M. *et al.* Mobilization of Processed, Membrane-Tethered SPT23 Transcription Factor by CDC48UFD1/NPL4, a Ubiquitin-Selective Chaperone. *Cell* **107,** 667–677 (2001).
75. Torres, E. M. *et al.* Identification of Aneuploidy-Tolerating Mutations. *Cell* **143,** 71–83 (2010).
76. Kim, H. J., Ishidou, E., Kitagawa, E., Momose, Y. & Iwahashi, H. A Yeast DNA Microarray for the Evaluation of Toxicity in Environmental Water Containing Burned Ash. *Environ. Monit. Assess.* **92,** 253–272 (2004).
77. Omnus, D. J. & Ljungdahl, P. O. Latency of transcription factor Stp1 depends on a modular regulatory motif that functions as cytoplasmic retention determinant and nuclear degron. *Mol. Biol. Cell* **25,** 3823–3833 (2014).
78. Schmidt, A., Hall, M. N. & Koller, A. Two FK506 resistance-conferring genes in Saccharomyces cerevisiae, TAT1 and TAT2, encode amino acid permeases mediating tyrosine and tryptophan uptake. *Mol. Cell. Biol.* **14,** 6597–6606 (1994).
79. Wang, K., Li, W., Chen, J., Peng, Y. & Zheng, J. Detection of cysteine- and lysine-based protein adductions by reactive metabolites of 2,5-dimethylfuran. *Anal. Chim. Acta* **896,** 93–101 (2015).
80. Brito, I., Monje-Casas, F. & Amon, A. The Lrs4-Csm1 monopolin complex associates with kinetochores during anaphase and is required for accurate chromosome segregation. *Cell Cycle* **9,** 3611–3618 (2010).
81. Cadet, J. & Davies, K. J. A. Oxidative DNA damage & repair: An introduction. *Free Radic. Biol. Med.* **107,** 2–12 (2017).
82. Allam, W. R., Ashour, M. E., Waly, A. A. & El-Khamisy, S. Role of Protein Linked DNA Breaks in Cancer. in *Personalised Medicine* 41–58 (Springer, Cham, 2017). doi:10.1007/978-3-319-60733-7_3
83. Tsutsui, Y., Morishita, T., Iwasaki, H., Toh, H. & Shinagawa, H. A Recombination Repair Gene of Schizosaccharomyces pombe, rhp57, Is a Functional Homolog of the Saccharomyces cerevisiae RAD57 Gene and Is Phylogenetically Related to the Human XRCC3 Gene. *Genetics* **154,** 1451–1461 (2000).
84. Morris, L. P. *et al.* Saccharomyces cerevisiae Apn1 mutation affecting stable protein expression mimics catalytic activity impairment: Implications for assessing DNA repair capacity in humans. *DNA Repair* **11,** 753–765 (2012).
85. Lee, J. *et al.* Yap1 and Skn7 Control Two Specialized Oxidative Stress Response Regulons in Yeast. *J. Biol. Chem.* **274,** 16040–16046 (1999).
86. Apel, K. & Hirt, H. REACTIVE OXYGEN SPECIES: Metabolism, Oxidative Stress, and Signal Transduction. *Annu. Rev. Plant Biol.* **55,** 373–399 (2004).
87. Yan, C., Lee, L. H. & Davis, L. I. Crm1p mediates regulated nuclear export of a yeast AP-1-like transcription factor. *EMBO J.* **17,** 7416–7429 (1998).

88. Herrero, E., Ros, J., Bellí, G. & Cabiscol, E. Redox control and oxidative stress in yeast cells. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1780,** 1217–1235 (2008).
89. McIntyre, J. & Woodgate, R. Regulation of translesion DNA synthesis: Posttranslational modification of lysine residues in key proteins. *DNA Repair* (2015). doi:10.1016/j.dnarep.2015.02.011
90. Zhao, J., Zhang, R., Misawa, K. & Shibuya, K. Experimental product study of the OH-initiated oxidation of m-xylene. *J. Photochem. Photobiol. Chem.* **176,** 199–207 (2005).
91. Gaigg, B. *et al.* Depletion of Acyl-Coenzyme A-Binding Protein Affects Sphingolipid Synthesis and Causes Vesicle Accumulation and Membrane Defects inSaccharomyces cerevisiae. *Mol. Biol. Cell* **12,** 1147–1160 (2001).
92. Sutherland, F. C. *et al.* Characteristics of Fps1-dependent and -independent glycerol transport in Saccharomyces cerevisiae. *J. Bacteriol.* **179,** 7790–7795 (1997).
93. Toh, T.-H. *et al.* Implications of FPS1 deletion and membrane ergosterol content for glycerol efflux from Saccharomyces cerevisiae. *FEMS Yeast Res.* **1,** 205–211 (2001).
94. Lis, E. T. & Romesberg, F. E. Role of Doa1 in the Saccharomyces cerevisiae DNA Damage Response. *Mol. Cell. Biol.* **26,** 4122–4133 (2006).
95. Willingham, S., Outeiro, T. F., DeVit, M. J., Lindquist, S. L. & Muchowski, P. J. Yeast Genes That Enhance the Toxicity of a Mutant Huntingtin Fragment or α-Synuclein. *Science* **302,** 1769–1772 (2003).
96. Yun, C.-W., Tiedeman, J. S., Moore, R. E. & Philpott, C. C. Siderophore-Iron Uptake in Saccharomyces cerevisiae IDENTIFICATION OF FERRICHROME AND FUSARININE TRANSPORTERS. *J. Biol. Chem.* **275,** 16354–16359 (2000).
97. Samanfar, B. *et al.* A global investigation of gene deletion strains that affect premature stop codon bypass in yeast, Saccharomyces cerevisiae. *Mol. Biosyst.* **10,** 916–924 (2014).
98. Choudhary, V. & Schneiter, R. Pathogen-Related Yeast (PRY) proteins and members of the CAP superfamily are secreted sterol-binding proteins. *Proc. Natl. Acad. Sci.* **109,** 16882–16887 (2012).
99. Smith, R. D. & Lupashin, V. V. Role of the conserved oligomeric Golgi (COG) complex in protein glycosylation. *Carbohydr. Res.* **343,** 2024–2031 (2008).
100. Climer, L. K., Hendrix, R. D. & Lupashin, V. V. Conserved Oligomeric Golgi and Neuronal Vesicular Trafficking. in *SpringerLink* 1–21 (Springer, Berlin, Heidelberg, 2017). doi:10.1007/164_2017_65
101. Bugnicourt, A., Mari, M., Reggiori, F., Haguenauer-Tsapis, R. & Galan, J.-M. Irs4p and Tax4p: Two Redundant EH Domain Proteins Involved in Autophagy. *Traffic* **9,** 755–769 (2008).
102. Lynch-Day, M. A. *et al.* Trs85 directs a Ypt1 GEF, TRAPPIII, to the phagophore to promote autophagy. *Proc. Natl. Acad. Sci.* **107,** 7811–7816 (2010).
103. Lynch-Day, M. A. & Klionsky, D. J. The Cvt pathway as a model for selective autophagy. *FEBS Lett.* **584,** 1359–1366 (2010).
104. Kim, D. & Hahn, J.-S. Roles of Yap1 transcription factor and antioxidants in yeast tolerance to furfural and 5-hydroxymetylfurfural that function as thiol-reactive electrophiles generating oxidative stress. *Appl. Environ. Microbiol.* AEM.00643-13 (2013). doi:10.1128/AEM.00643-13

105. Parnell, E. J. *et al.* The Rts1 Regulatory Subunit of PP2A Phosphatase Controls Expression of the HO Endonuclease via Localization of the Ace2 Transcription Factor. *J. Biol. Chem.* **289,** 35431–35437 (2014).
106. Heiman, M. G. & Walter, P. Prm1p, a Pheromone-Regulated Multispanning Membrane Protein, Facilitates Plasma Membrane Fusion during Yeast Mating. *J. Cell Biol.* **151,** 719–730 (2000).
107. Oh, C.-S., Toke, D. A., Mandala, S. & Martin, C. E. ELO2 and ELO3, Homologues of theSaccharomyces cerevisiae ELO1 Gene, Function in Fatty Acid Elongation and Are Required for Sphingolipid Formation. *J. Biol. Chem.* **272,** 17376–17384 (1997).
108. Huang, M., Zhou, Z. & Elledge, S. J. The DNA Replication and Damage Checkpoint Pathways Induce Transcription by Inhibition of the Crt1 Repressor. *Cell* **94,** 595–605 (1998).
109. Zhang, Z. & Reese, J. C. Redundant Mechanisms Are Used by Ssn6-Tup1 in Repressing Chromosomal Gene Transcription in Saccharomyces cerevisiae. *J. Biol. Chem.* **279,** 39240–39250 (2004).
110. Robinson, J. S., Klionsky, D. J., Banta, L. M. & Emr, S. D. Protein sorting in Saccharomyces cerevisiae: isolation of mutants defective in the delivery and processing of multiple vacuolar hydrolases. *Mol. Cell. Biol.* **8,** 4936–4948 (1988).
111. Anand, V. C., Daboussi, L., Lorenz, T. C. & Payne, G. S. Genome-wide Analysis of AP-3–dependent Protein Transport in Yeast. *Mol. Biol. Cell* **20,** 1592–1604 (2009).
112. Morvan, J., Rinaldi, B. & Friant, S. Pkh1/2-dependent phosphorylation of Vps27 regulates ESCRT-I recruitment to endosomes. *Mol. Biol. Cell* **23,** 4054–4064 (2012).
113. Thorsen, M. *et al.* Genetic basis of arsenite and cadmium tolerance in Saccharomyces cerevisiae. *BMC Genomics* **10,** 105 (2009).
114. Perrone, G. G., Grant, C. M. & Dawes, I. W. Genetic and Environmental Factors Influencing Glutathione Homeostasis in Saccharomyces cerevisiae. *Mol. Biol. Cell* **16,** 218–230 (2005).
115. Hidalgo, F. J., Alcón, E. & Zamora, R. Reactive Carbonyl-Scavenging Ability of 2-Aminoimidazoles: 2-Amino-1-methylbenzimidazole and 2-Amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP). *J. Agric. Food Chem.* **62,** 12045–12051 (2014).
116. Long, E. K. & Picklo, M. J. Trans-4-hydroxy-2-hexenal, a product of n-3 fatty acid peroxidation: Make some room HNE…. *Free Radic. Biol. Med.* **49,** 1–8 (2010).
117. Kasai, H. & Kawai, K. 4-Oxo-2-hexenal, a mutagen formed by ω-3 fat peroxidation: Occurrence, detection and adduct formation. *Mutat. Res. Mutat. Res.* **659,** 56–59 (2008).
118. Grasse, L. D., Lamé, M. W. & Segall, H. J. In vivo covalent binding of trans-4-hydroxy-2-hexenal to rat liver macromolecules. *Toxicol. Lett.* **29,** 43–49 (1985).
119. Dinter, A. & Berger, E. G. Golgi-disturbing agents. *Histochem. Cell Biol.* **109,** 571–590 (1998).
120. Oka, T. & Krieger, M. Multi-Component Protein Complexes and Golgi Membrane Trafficking. *J. Biochem. (Tokyo)* **137,** 109–114 (2005).
121. Losev, E. *et al.* Golgi maturation visualized in living yeast. *Nature* **441,** 1002 (2006).
122. Nguyen, T. T. M., Iwaki, A., Ohya, Y. & Izawa, S. Vanillin causes the activation of Yap1 and mitochondrial fragmentation in Saccharomyces cerevisiae. *J. Biosci. Bioeng.* **117,** 33–38 (2014).
123. Witz, G. Biological interactions of α,β-unsaturated aldehydes. *Free Radic. Biol. Med.* **7,** 333–349 (1989).

124. Birrell, G. W. *et al.* Transcriptional response of Saccharomyces cerevisiae to DNA-damaging agents does not identify the genes that protect against these agents. *Proc. Natl. Acad. Sci.* **99,** 8778–8783 (2002).

Chapter 4
Using CRISPR To Enhance The Metabolic Capacity Of Cell Lines Commonly Used In High Throughput Screening

Adoption of High-Throughput Methods For Regulatory Use

Regulatory agencies both in the US and abroad have been shifting to High Throughput Screening (HTS) strategies for chemical evaluation over the last few decades[1–6] to both accelerate the pace of and reduce the cost of hazard testing for industrial chemicals. From large-scale international regulatory schema such as the European Union's Registration, Evaluation, Authorization, and Restriction of Chemicals (REACH)[7] program, to the recent amendments to the United States Toxic Substances Control Act (the 2016 Frank R. Lautenberg Chemical Safety for the 21st Century Act)[8], regulations banning the use of animal testing under certain circumstances (as is the case of REACH) or encouraging the use of in vitro and alternative methods (both REACH and the Lautenberg Act) have provided the statutory impetus for regulatory bodies to more fully embrace HTS. This is not to say that regulators haven't recognized the power and necessity of HTS for some time: the National Research Council's landmark 2007 report *Toxicity Testing in the 21st Century: A Vision and a Strategy*[9] spawned the massive Tox21[10] and ToxCast[11] initiatives led by the US Environmental Protection Agency within years of its publication. These programs, as well as the more narrowly-focused Endocrine Disruptor Screening Program (EDSP)[6,12] represent a fundamental paradigm shift in the field of toxicology from being reactive to being proactive. This shift was made possible through the marriage of *in vitro* testing strategies based around intellectual advancements in molecular toxicology with the substantial technological groundwork laid by the pharmaceutical industry over the last two decades.

Since the technology boom began in the 1990s, pharmaceutical companies had been developing tools that combined automation, microfluidics, multi-well plates, and *in vitro* assays to identify drug candidate compounds and generate safety data; the commercialization of these high-throughput systems lead to their adaption – and eventually, adoption – for the purpose of screening industrial chemicals. Drug companies also pushed for the acceptance of many *in vitro* tests for regulatory purposes by submitting so much *in vitro* assay data to various regulatory agencies in the late 90s that the FDA was prompted to issue standardized guidance for *in vitro* safety data[13]. Recognizing the potential opportunities of these new technologies and the breakneck pace at which the field of molecular biology was advancing, in 2004, the NIH established dozens of programs and spent billions of dollars on its' Roadmap initiatives, many of which offered awards for the development of HTS and *in vitro* testing methodologies[14,15]. In 2005, the Government Accounting Office estimated the number of chemicals in US commerce listed on the TSCA inventory to be around 82,000, a sobering figure which prompted researchers and regulators to emphasize the development of test methods to prioritize chemicals for additional testing[16]. (The TSCA inventory lists roughly 85,000 as of this writing[17]) The 2007 NRC report correctly recognized HTS as not only the most effective strategy for prioritizing and assessing industrial chemicals, but also the only practical means of

doing so. Hence, the ToxCast program prioritized an initial list of about 10,000 "Antimicrobials, [*sic*] pesticidal inerts, high production volume (HPV: > 1 million lbs/year) chemicals, inventory update rule (> 10,000 lbs/year, < 1 million lbs/year) chemicals, and drinking water contaminant candidate list chemicals,"[16] which lack extensive toxicological data to be run through a battery of high-throughput *in vitro* assays.

As of this writing, slightly fewer than half of the 10,000 chemicals prioritized by ToxCast have been run through the program's HTS battery, but even these data have generated myriad and various analyses[11,5,18,19,12,6,20]. Over the course of the programs' existence, the number and type of ToxCast assays has shifted[21], but the HTS battery has yet to effectively incorporate the metabolism of test compounds as an aspect of chemical toxicity.

Relevance Of *in vitro* Hepatic Models To High-Throughput Screens

Immortalized cell lines, useful though they may be for many biological applications, are noted to have wildly different, and in most cases, deficient, levels of Phase I and Phase II metabolic enzymes[22–24]. Depending on the application and the *in vitro* assay one wishes to perform, this may not present a significant barrier, but in the case of HTS toxicity testing, the problem is grave indeed. The liver is the primary metabolic organ and is principally responsible for detoxification of xenobiotic compounds, so any *in vitro* liver system unable to facilitate biotransformation or detoxification of test compounds is liable to either over- or under-estimate the potency of a compound for a given endpoint[24]. A myriad of possible solutions to this problem have been explored, and there are some very promising *in vitro* liver assay technologies under development. However, existing liver cell lines and *in vitro* liver systems are inadequate solutions for HTS either due to insufficient metabolic capacity or due to practical limitations.

An excellent review *in vitro* liver technologies is provided by Soldatow et al.[25], but a brief overview of historical and contemporary systems and their applicability to HTS is provided here.

Liver slices are perhaps the oldest, most metabolically-relevant, and most straightforward *in vitro* liver model, but due to their cost and complexity they are also the lowest throughput and least common.[25] Primary liver cells harvested from donors provide the most physiologically-relevant metabolic and toxicity data for *in vitro* experiments. However, primary liver cells are not only infeasible due to the logistical challenges of obtaining them in large quantity, there are also questions of generalizability of the results due to inter-individual variability in enzyme expression.[26] Further, unlike the liver slice model, primary liver cells lose the liver metabolic phenotype within 24-48 hours of extraction, expressing significantly fewer Phase I and Phase II enzymes, diminishing their predictive power for drugs, industrial chemicals, and environmental contaminants[27].

A growing variety of *in vitro* techniques have been developed in an attempt to recapitulate the spatial environments that hepatocytes experience in the body, as there is ample research that mechanical forces and spatial orientation can affect the phenotype of cultured hepatocytes[26]. Substrates such as murine sarcoma extract and human collagen have been used to provide a three-dimensional (3-D) matrix for cultured hepatocytes, but are both recognized as being of limited use in hepatocyte cell culture at scale, albeit for different reasons: the former, being of animal tumor origin and immunogenic, the latter having physicochemical properties that limit its utility[28]. Sandwich cultures, which are *in vitro* configurations wherein a layer of hepatocytes is cultured between two layers of collagen, help cultured hepatocytes maintain their polarity and many phenotypic features, including metabolism[28–30]. However, sandwich cultures are costly, low-throughput, and eventually lose metabolic parity with primary cells[25].

Spheroid and 3-D cultures capitalize on our growing understanding of the importance of cellular microenvironments for the regulation of cell metabolism and homeostasis. Through these techniques, researchers have demonstrated improved metabolic competence of cultured cells, but the resources required to produce the culture render them impractical as part of an HTS strategy. Related technologies such as organs-on-a-chip and liver organoid cultures similarly employ 3-D culture techniques as well as microfluidics and nanofabricated substrates to restore "primary-like" metabolic capacity, but these, too, are not (as of this writing) suitably scalable or cost-effective for HTS.

Metabolic Competence In Hepatoma Cell Lines

The HepaRG and HepG2 cell lines are two of the commonest hepatoma cell lines, each seeing broad use across a range of biomedical applications. But while both cell lines demonstrate expression of some Phase I and Phase II detoxification enzymes, neither achieves expression levels approaching those of primary cells.

The USEPA considers the HepG2 cell line "metabolically competent" for the purposes of HTS studies, even though HepG2 cells are notably deficient in key the phase I and Phase II metabolic enzymes compared to fresh hepatocytes[31,32]. HepaRG cells have been shown to be more sensitive to model hepatotoxicants and have more "primary-like" expression levels for xenobiotic metabolism enzymes, but they are more sensitive to culture conditions and are much more time-consuming to work with than HepG2, requiring weeks of DMSO treatment before they differentiate towards a hepatocyte phenotype[24,33].

While some of *in vitro* liver technologies described thus far may eventually prove scalable to the point of adoption for HTS, cost and logistical considerations for the present regulatory environment necessitate a solution using a metabolically active immortalized cell line. Aside from treating HepaRG cells with DMSO for extended periods of time and growing them in the aforementioned 3D systems, other, more traditional molecular biology techniques have been used to induce Phase I and/or Phase II metabolic enzymes in hepatoma cell lines, primarily those in the

cytochrome P450 (CYP) family of monooxygenases. Adenoviral[34,35] and SV40 transfection vectors[36,37] with CYP-containing plasmids, *piggyBac* transposition and monoclonal expansion[38], and simple transformation protocols using cloned human CYP cDNA[39] have all been used with varying success to coax hepatoma cell lines to express more Phase I enzymes. Unfortunately, these techniques tend to be labor-intensive, limited to the activation of one or a handful of genes, and when successful, generally lead to constitutive massive overexpression of the genes in question, limiting their utility for screening purposes.

Ultimately, these technologies do not constitute a holistic solution to the problem, and therefore, many research groups have approached this challenge of metabolic insufficiency in cultured hepatocytes using a range of tactics, from simple adjuvant treatments to total cell reprogramming. Classic inducers of metabolic enzymes in HepG2 cells include beta-naphthophenone, phenobarbital, and rifampicin[40], but concerns about how their mechanisms of action might confound toxicity tests have limited their use in regulatory and HTS contexts. Fully differentiated hepatocytes expressing near-physiological levels of CYPs, UGTs, and other nuclear receptors have been derived from iPSCs[41]. Mitani et al[42] recently demonstrated that modulation of the Wnt pathway allowed them to generate iPSC hepatocytes with zone-specific phenotypes. Though promising, these methods are still decidedly low throughput due to cost and effort, and the challenge of maintaining consistent enzyme levels across differentiated cultures remains[25,43]. Other groups have explored the relationships of between induction nuclear receptors CAR, PXR, RXR, VDR, FXR, LXR and PPAR and the activation of CYPs. The result of this work is a model wherein agonists or antagonists could be specific compounds in nuclear receptor signaling pathways to induce CYP expression without producing as many off-target effects as existing adjuvants[44,45]. Epigenetic factors such as methylation content have been shown to affect profoundly affect CYP expression[46], and there is evidence that treatment with demethylating agents such as 5-azacytidine can significantly enhance expression in cultured hepatocytes[47]. Our work aims to build on these myriad strategies and adopt the aspects of each which are most effective so as to develop a cheap, scalable, minimally-invasive method for enhancing CYP expression in cultured hepatocytes.

CRISPRa as a Strategy to Improve Metabolic Competence in Cell Lines

CRISPR is a powerful gene-editing tool responsible for the current renaissance in cell biology, with a sprawling range of applications, many of which are still being discovered. We have adapted the technology to tackle the problem of metabolic competence in HTS *in vitro* testing and develop a minimally-invasive protocol for improving CYP expression using the tandem catalytically-deactivated Cas9 enzyme and transcriptional activator system described in Konerman et al (2014)[48]. This system, called CRISPR activation system (CRISPRa) takes advantage of the Cas9 enzyme's potent targeting ability to bind the promoter sequences of various genes for xenobiotic-metabolizing enzymes in the CYP and UGT families. Three components are used to activate the genes: a Cas9 enzyme (dCas9-VP64-Blast, or dCas9) modified to retain its sequence-specific binding capabilities but lack DNA

cleaving functionality, a synergistic activation mediator (hereafter called "SAM," MS2-P65-HSF1-Hygro) and guide RNA (sgRNA or "activation guides") specific to the promoter region of any gene one wishes to activate. Plasmids containing the sequences for the dCas9 and the SAM complex are available from Addgene, and any cloning lab can make sgRNA plasmids quickly and cheaply. A simple transfection with these three components produces transient activation of the gene or genes specified by the sgRNA, with activation peaking at about 48 hours.

In the initial paper, this method was used to activate up to 12 genes simultaneously, raising exciting possibilities for restoring metabolic competence to cell lines. While this discussion primarily concerns the expression of CYPs in hepatocytes, CRISPRa technology, once optimized, could be used to generate metabolic "mimics" of other tissue types, expanding the range of *in vitro* applications. Since this method requires only a single transfection step that could easily be incorporated into existing HTS with no additional requirement for inducers or adjuvants, it is an appealing strategy for improving the efficacy of HTS data gathering by inducing more "primary-like" metabolism.

We chose target enzymes for activation in our pilot study based on literature evidence for their role in xenobiotic metabolism and selected CYP1A1/1A2 (these enzymes have a bidirectional promoter, so induction of one or the other is somewhat trivial, given their closely related functions[49]), CYP3A4, CYP2E1, CYP2B6, and UGT1A6 given the importance of these enzymes for Phase I and Phase II transformations in the liver [50–52]. Our initial experiments focused on HepG2 cells, but we found that these cells were not as responsive to transfection as desired, so we continued our work in HEK293T cells and used these experiments to optimize our transfection strategy and to probe mechanisms of CYP regulation in cultured cells.

Materials and Methods

Designing the CRISPRa component plasmids

Plasmids containing the CRISPRa components were acquired through two sources: Addgene and the Berkeley q3 macrolab. The dCas9-VP64 (addegene# 61425) and MS2-P65-HSF1 activator helper complex (Addgene# 61426) were purchased from Addgene, while the guide RNA plasmid was generated by the macrolab by cloning the promoter target sequence for the gene of interest into a lenti sgRNA(MS2)_zeo backbone (Addgene#61427). Promoter target sequences were generated using the Cas9 Activator tool created by the Zhang Lab[48,53]. Three sgRNA sequences were generated for each gene of interest and tested individually by RT-qPCR to identify the most effective sgRNA for gene activation. Activation was measured

Generation of CRISPRa component plasmids

Each guide was cloned individually in the lenti sgRNA(MS2)_zeocin plasmid (Addgene# 61427) using the Golden-Gate sgRNA method described in detail in Konermann S et al. 2014.[48] After selection, bacteria were cultured flasks of liquid

LB-Amp and on LB-Amp agar plates. Maxi Preps of plasmid were made from bacteria cultured in flasks of liquid LB-AMP using Qiagen Plasmid kits[54].

Production of transiently transfected HEK293T cells
HEK293T cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) with 10% FBS and 1% Penicillin/Streptomycin. The day before transfection, cells were trypsinized and counted. Cells were seeded in a 24-well at 1.25x105 cells per well in 0.5 ml of complete growth medium. Cell density was 50-80% confluent on the day of transfection. Using a 24-well plate, 0.5 µg of dCas9 plasmid, 0.5 µg of SAM plasmid, and up to 0.5 µg of sgRNA plasmid, not exceeding a total of 1.5 µg per well for a 24-well plate, was added to a volume of Optimem equal to 50 µL for each well to be transfected. This mixture and additional reagents were prepared as directed by the Lipofectamine 3000 kit (Thermofisher Cat. L3000015). Cells were then incubated at 37 C in a CO2 incubator.

Production of stably transfected HEK293T cells
Stable HEK293T cells for the SAM component were generated by lentiviral transduction. Lentiviruses were produced for each of the SAM components and packaged by co-transfecting using lipofectamine 3000 and each of the following lentiviral plasmids: Addegene# 61425 (dCas9-VP64), Addgene# 61426 (MS2-P65-HSF1 activator helper complex), or guide RNA plasmid (Addegene# 61427 harboring a cloned CYP activation sgRNA), with the psPAX2 packaging and pMD2.G envelope plasmids. The day after the transfection, the cell culture medium was replaced with fresh medium (DMEM, 10% FBS). After 48H of culture, the medium (containing the lentivirus) was collected, filtered using 0.45 µM sterile syringe filters. The 293T cells were seeded in 6-well plates and infected with 250 µl of the lentiviral solution in the presence of 8µg/ml polybrene to enhance the infection rate. After 24H, the media was changed and the cells were allowed to grow for another 24H before starting the selection against the appropriate antibiotics. HEK293T cells were transduced either with dCas9-VP64 alone to generate 293T-Cas9-VP64 (blasticidin selection) or with both dCas9-VP64 and MS2-P65-HSF1 lentiviruses (blasticidin and hygromycin B selection) to generate 293T-SAM. To generate 293T-CYP3A5/1A1/3A4-SAM cells, 293T cells were co-transduced with the two previous components and with lentiviruses for three different CYP3A5 activation guides (blasticin, hygromycin B, and zeocin selection). The cells were then transduced as described above with CY3A4 and CY1A1 lentiviruses (for all guides gRNA).

Gene expression by real-time PCR:
Gene expression was determined at 48 hours post-transfection using RT-qPCR. RNA was extracted from cells using Qiagen RNeasy Mini Kit (Cat No./ID: 74104) using manufacturer's instructions. RNA was quantified on Infinite 200 NanoQuant plate reader (Tecan). cDNA was obtained using a Bio-Rad T100 Thermal Cycler and iScript™ Reverse Transcription Supermix kit (Bio-Rad 1708841) with 1 µg of RNA per sample (volume varied by sample, due to different volumes of RNA extracted), 4µL of iScript RT Supermix and enough water to reach 20µL per reaction. Following the manufacturer's instructions, reactions were run at 5 minutes at 25°C, 20

minutes at 46°C, and 1 minute at 95°C. Using the generated cDNA, we ran qPCR experiments on a Bio-Rad CFX Connect Real-Time System using SsoFast™ EvaGreen® Supermix (Bio-Rad 1725201). Each reaction consisted of 10µL of Ssofast EvagGreen supermix, 0.1µL forward primer, 0.1µL reverese primer, 2µL cDNA, and 7.8µL DNase-free water and was run for 35 cycles. The program initialized for 30 seconds at 95°C and each cycle had the following steps: 5 seconds at 98°C and 5 seconds at 65°C. Expression data was normalized (ΔΔCq) to RPL19, and we compared the fold expression of enzyme transcripts against no treatment controls and controls receiving Cas9 and SAM components but no sgRNA plasmids. Initial testing of the sgRNA plasmids used three sgRNA sequence variants and the plasmid with the sequence that generated the highest expression level as measured by RT-qPCR was used for subsequent transfection experiments.

Generation of stable 293T cells for the SAM components
Stable 293T cells for the SAM component were generated by lentiviral transduction. Lentiviruses were produced for each of the SAM components: they were packaged by co-transfecting each of the following lentiviral plasmids: Addegene# 61425 (dCas9-VP64), Addgene# 61426 (MS2-P65-HSF1 activator helper complex), or guide RNA plasmid (addegene#61427 harboring a cloned CYP activation gRNA), with the psPAX2 packaging and pMD2.G envelope plasmids.

Cytotoxicity Assay
Ten 25cc flasks with 4 mL DMEM containing 1x10^6 HEK293T cells, where N is the number of conditions to be evaluated in the cytotoxicity assay, transfect each flask using the "6-well plate" parameters as outlined in the manufacturer's instructions for Lipofectamine 3000 kit (Thermofisher Cat. L3000015). Briefly: 800 ng of dCas9, SAM, and sgRNA plasmids are added to 10µL P3000 Reagent and 250µL of Optim-MEM. To this mixture, 250µL Optim-MEM and 11.25µL Lipofectamine was added, and the larger mixture was allowed to incubate for 15 minutes. After 15 minutes, 500µL of the Lipofectamine/DNA/Optim-MEM mixture was added to a well in the 6-well plate. In the 6-well plate, one well received no transfection (No-treatment control), one well received only the dCas9 and SAM plasmids (SAM control), and one well each received dCas9, SAM, and an sgRNA plasmid for CYP1A2, CYP2E1, or CYP3A4. After each transfection, the plate was agitated gently by hand and then placed in the cell culture incubator. 24 hours after transfection, cells from each well were trypsinized and seeded into separate 384-well plates at 3500 cells/25 µL of fresh DMEM per well, cells were allowed to settle for 24 hours. DMEM negative control media was prepared with 1% DMSO (2X dosing concentration). Test chemicals were prepared at 2X dosing concentrations of 500 µM, 250 µM, 125 µM, 62.5 µM, 31.25 µM, 15.625 µM, 7.81 µM, and 3.90 µM. Test chemicals: Benzo[a]pyrene (CAS #: 50-32-8), Aflatoxin B1 (CAS # 1162-65-8), Cyclophosphamide monohydrate (CAS #: 6055-19-2), 2-naphthylamine (CAS #: 91-51-8), Acrylamide (CAS #: 79-06-1), Doxorubicin hydrochloride (CAS #: 25316-40-9), 6-aminochyrsene (CAS #: 2642-98-0), Methoxypsoralen (CAS #: 298-81-7), and 4-nitrophenol (CAS #: 100-02-7). Negative control wells and background wells were

dosed with 25 µL DMEM control media with 1% DMSO (final DMSO concentration: 0.5%; 5 replicates per control). Treatment wells were dosed with 25 µL of appropriate 2X chemical dose in triplicates (final dosing concentrations of 250 µM, 125 µM, 62.5 µM, 31.25 µM, 15.625 µM, 7.81 µM, and 3.90 µM, and 1.95 µM). After dosing, plates were incubated at 37°C/5% CO_2 for 24 hours. After 24 hours, cells were removed from incubator and allowed to equilibrate to room temperature for 30 minutes. 40 µL of CellTiter-Glo (Promega) was added to each well (~1:1 ratio with media) and mixed by shaking. The plate was incubated at room temperature for ~15 minutes. Luminescence was recorded on a BioTek Synergy H1 Hybrid Reader. The cell viability was expressed as percentage of the control (DMSO vehicle control).

Results

HepG2 was an obvious choice for using CRISPRa, since it is a hepatic cell line, and the aim was to improve metabolic competence in HTS cell lines. However, this cell line has not been known to be easily transfected, and we have confirmed that here. As indicated in Figure 1, we were unable to induce expression of CYPs beyond the cell line's meager background levels, which are significantly lower than in the liver standard. We therefore used HEK293T because it is readily transfected and is known to effectively express transfected CYP proteins (Figures 2 and 3)[55,56]. As recommended in Konermann et al 2014[48] and in general in the CRISPR literature, we tested several variants of the sgRNA sequences to identify the most effective guides for transfection. Not all the sgRNA variants we tested were equally effective for inducing CYPs and for the CYP2B6 and UGT1A6, two rounds of sgRNA design were necessary before functional guides were identified (Figure 2A, Figure 3).

Unfortunately, we observed that we were not able to achieve expression on par with the liver standard, even in the HEK293T cell line, so we took the aim of exploring methods to enhance CYP expression in our system. To do this, we focused on the CYP3A4 construct that produced the greatest induction of CYP3A4 mRNA transcripts in HEK293T cells (Figure 2A). We observed improved expression efficiency for CYP3A4 when we co-transfected with sgRNA for CYP1A2, CYP3A4, and CYP2E1 simultaneously, versus CYP3A4 alone (Figure 4).

Since each CRISPRa transfection required transfection with plasmids for at least three different components: the dCas9 enzyme, the SAM complex, and the sgRNA for the gene of interest, we created a cell line that stably expressed the dCas9 and SAM protein products so that we would only need to transfect with the sgRNA of interest. This would both save effort in future experiments and had the potential to improve expression in our cell lines. We created the HEK-SAM cell line through lentiviral transduction and were able to successfully induce transfection with sgRNA plasmids alone (Figure 5). However, this system did not produce improved mRNA expression as we hoped. This could be the result of a poor integration location or low expression of the dCas9 and SAM components.

Regulation of CYP450s occurs at multiple layers and there are many co-regulatory proteins that are involved in the expression of these Phase I enzymes. We selected a handful of these co-regulators in an attempt to stimulate production of important proteins the CYP3A4 regulatory pathway: FOXA3, SREBF, CEBPA, HNF1A, PXR, HNF4A, and PPARγ[57–62] to determine if up-regulating one of these factors would improve CYP3A4 expression. (Figures 6-8) We also included in our experiments a single treatment condition wherein we transfected HEK293T cells with sgRNA guides for CYP3A4, the CRISPRa components (in the HEK293T cell line, HEK-SAM cells only received CYP3A4 plasmid) and each of the different regulatory proteins. This condition was referred to as the "Big Mix" condition due to its myriad contents (Figures 6 and 7). We did not observe meaningful up-regulation of CYP3A4 expression in any of these treatment conditions, for either the HEK293T cells or the HEK-SAM cell line. Among the regulators we transfected the cells with, only HNF4A demonstrated increased expression (Figure 8), and this was only in the HEK293T cell line. It is not clear whether this is the result of poor sgRNA targeting or if additional regulatory mechanisms are involved. In the initial Konermann paper, the researchers activated up to twelve genes simultaneously with this method of gene activation[48], so we do not believe that the poor expression is the result of overwhelming the cells with transfection material.

Since DMSO co-treatment is vital to the successful induction of Phase I and Phase II enzymes in HepaRG cells[33], we co-treated HEK293T cells with DMSO and transfected them with CRISPRa and sgRNA for CYP3A4. The results of this experiment (Figures 9 and 10) indicated no improved expression for CYP3A4, but rather, a dose-response decrease of expression. This could have been attributable to DMSO-induced cytotoxicity or to some suppressive mechanism. Intrigued, we repeated the experiment using the sgRNA for CYP1A2 and observed a similar pattern, although the cells transfected with CYP1A2 sgRNA demonstrated greater resistance to DMSO treatment than cells transfected with CYP3A4 sgRNA. From our results, it is possible to conclude that treatment with DMSO decreases CYP3A4 expression in HEK293T cells. If future studies bear this out, then it may be advisable to reconsider *in vitro* chemical testing strategies that use DMSO as a solvent.

Several studies have noted the role of DNA methylation in the epigenetic regulation of CYP450 enzymes[59,63–65], and this may account for the rapid and significant phenotypic changes observed in cell lines compared to primary cells. We hypothesized that treatment with a methylation inhibitor azacytidine (AZA) may facilitate greater CYP expression, so we co-treated HEK293T cells with doses of 1, 3, 5, 10, and 20 mM of AZA and a CRISPRa transfection to activate CYP3A4 (Figure 11). We observed a dose-dependent increase, then decrease in CYP3A4 mRNA, with the highest levels of induction occurring at 1 mM AZA, and lower levels of induction at all higher doses. The improvement in expression was not to the level of primary liver cells (Figure 11B), but the results of this experiment suggest that methylation may be a target for future experiments to improve the expression of Phase I and Phase II xenobiotic metabolism enzymes in cultured cells.

Having had some, albiet limited, success improving CYP expression in previous experiments, we decided to perform a few functional examinations to see if the CYP levels that we were inducing were sufficient to affect the sensitivity of the cell lines to canonical bioactivated toxins. (Figures 12-18) Although the CYP expression levels that we were able to induce in HEK293T cells were not at physiological levels, we suspected that due to the potent enzymatic activity of CYPs. Cells transfected with CYP1A2 and CYP2E1 did not indicate clear differences in cytotoxicity across any treatment concentrations. However, CYP3A4-transfected HEK293T cells treated with 4-nitrophenol demonstrated a noticeable decrease in viability compared to control at the 15.63μM dose, and the same was true for cells treated with Aflatoxin B1 or 8-methoxypsoran for all doses above 15.63μM (Figures 17 and 18).

Discussion

High-throughput screening methods are a vital component in drug development, regulatory toxicology, and industrial chemicals testing, particularly in their use to predict liver injury. As the major metabolic organ, the liver is the subject of intense interest for scientists working in all areas of the modern chemical ecosystem, and many talented researchers have devoted themselves to generating *in vitro* liver models to provide faster, cheaper, and more accurate testing. Common practice in current test methods, however, have not successfully merged HTS screening with physiologically-relevant metabolism in cultured cells, and the test batteries used by the EPA currently suffer from a lack of metabolic competence. A relatively small number of enzymes are responsible for a substantial portion (but by no means all) of the xenobiotic metabolism that occurs in the human liver, so inducing cultured cell lines to express these CYPs and UGTs at physiological levels has tremendous potential for improving the rate and accuracy of HTS for industrial chemicals.

Using a novel CRISPRa-based transfection method, we attempted to improve CYP and UGT expression in HepG2 cells, but found that the cell line did not seem to accept the transfection well and we could not induce CYP expression much higher than the cell line background. (Figure 1) We began experimenting with the HEK293T cell line, as that cell line has been demonstrated to readily accommodate transfections and express exogenous plasmids. Through transfections with the CRISPRa components, we observed considerable up-regulation of CYPS and UGTs above background levels in the HEK293T cell line. (Figures 2-4) We then used lentiviral transduction to create a stably-transfected cell line that constitutively expressed the dCas9 enzyme and the MS2-P65-HSF1 activator helper complex that we called HEK-SAM. (Figure 5) The HEK-SAM cell line only needed to be transfected with the sgRNA plasmid of interest to achieve CRISPRa-modulated gene expression.

Unfortunately, even in our most successful transfections, we could not achieve the level of expression observed in primary liver cells. To further explore the mechanisms involved in CYP suppression in cultured cells, we focused on the HEK293T cell line and the CYP3A4 enzyme.

Co-transfection with CYP1A2, CYP2E1, and CYP3A4 sgRNAs resulted in increased activation for all those genes, suggesting a synergistic effect. However, co-transfecting with a range of CYP3A4 regulatory genes did not appear to improve expression of CYP3A4, and in fact seems to have reduced expression, likely as a result of the many regulatory feedback mechanisms involved in CYP3A4 expression. Additionally, this may indicate that loss of metabolic competence in cultured cells is the result of other epigenetic or post-transcriptional mechanisms of suppression are at work, a possibility which we explored by treating the CYP3A4-transfected HEk293T cells with 5-azacytidine.

Figure 11 indicates that 5-azacytidine treatment increased expression of CYP3A4 over that of control at doses of 1,3, 5, and 10 mM, but not at 20 mM. This experiment suggests that the role of epigenetic methylation-based silencing mechanisms are more significant than the cell signaling and transcriptional activator pathways (as explored in Figures 6-8) in the phenotypic shift observed in cultured liver cells. This finding confirms in the HEK293T cell line observations other researchers have previously described, with respect to the suppressive effects of methylation on the CYP3A4 promoter in HepG2 cells[66]. It has also been demonstrated that inhibition of the DNA methyltransferase 1 (DNMT1) in HepG2 cells with the compound Zebularine results in up-regulated CYP expression[67], so the next step in this investigation will likely involve using CRISPR technology to knock down DNMT1. Rather than using CRISPR to excise the DNMT1 gene, or dose the cells with anti-cancer drugs, it may be possible to achieve DNMT1 knockdown by inducing microRNA expression. DNMT1 suppression has been accomplished using miR-29b[68], which affects DNMT1 indirectly, and through the synergistic use of miR-152 and miR-185, which targets DNMT1 indirectly[69]. Transfecting cells with the CRISPRa components, the CYP3A4 sgRNA plasmids, and the appropriate sgRNA plasmids to induce these microRNAs, and lead to DNMT1 suppression, which would in turn result in up-regulation of our desired CYPs. Either stable or transient transfections would be possible, and both would provide an improvement on existing HTS cell cultures. And based on the cytotoxicity data shown in Figures 17 and 18, our efforts may be further along than we initially realized.

Although we've yet to achieve physiological levels of CYP or UGT expression, this level of enzyme induction may not be necessary to improve the cell line metabolic competence to the point of usefulness. Despite the apparently minor change in CYP3A4 levels compared to primary liver cells, Figures 17 and 18 provide evidence that our CRISPRa methods resulted in increased sensitivity to canonical liver toxicants. This raises an interesting question: how much induction do we need to create "physiologically-relevant" metabolism? In the case of HEK293T cells, it seems that even rather modest increases in relative enzyme quantity improve sensitivity to an observable level. If we are then able to further enhance CYP production through the addition of DNMT1 inhibitors, the result may be a technique for conferring the metabolic phenotype of hepatocytes to HEK293T cells, with respect to bioactivation-induced cytotoxicity.

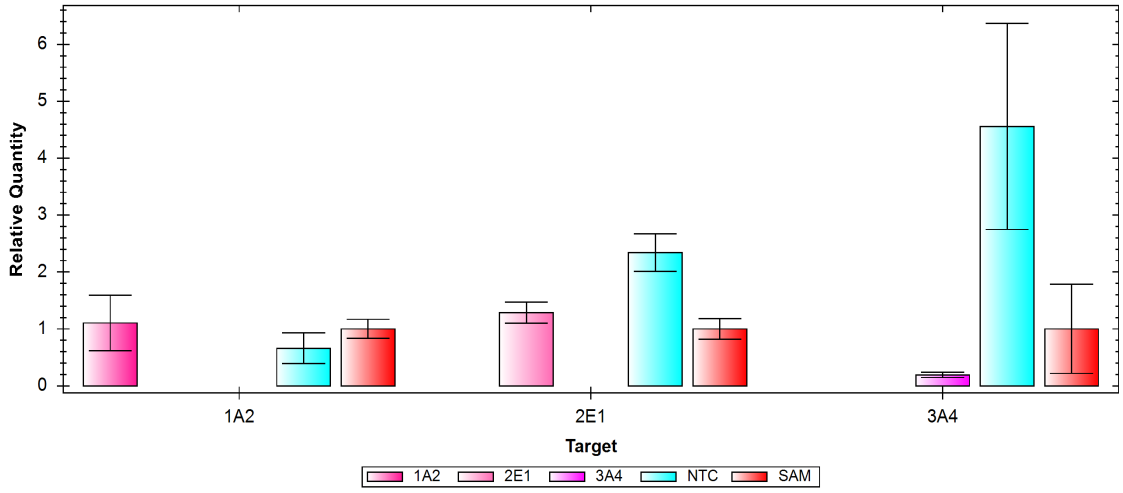
Conclusion

High Throughput Screening is the necessary and inevitable future of chemicals research and regulation. It is only through rapid, scalable technologies that we will be able to populate the gaps in our chemical hazard data sets and make informed decisions about the use of industrial chemicals. Of course, for these screens to be useful, they must demonstrate biological, and therefore, metabolic parity with the low throughput systems that they are intended to replace. The liver's role as the principle metabolic organ in mammals has given it a special status among screening systems, unfortunately, the development of *in vitro* hepatic HTS has not been commensurate with its necessity, and the HepG2 cell lines used in large scale industrial chemical screening programs such as the EPA's ToxCast program do not express CYPS or UGTs, key Phase I and II enzymes, at levels comparable to primary hepatocytes. As a result, use of these cell lines for cytotoxicity testing results in either an overestimation of toxicity in the case of chemicals which would otherwise be detoxified in the liver, or underestimation of toxicity, as in the case of chemicals which would otherwise be bioactivated.

We used a variant of CRISPR technology called CRISPRa to target a transcriptional activator to metabolic genes of interest in HepG2 and HEK293T cell lines. Although we saw very little induction of CYPs or UGTs in the HepG2 cells (Figure 1), we observed impressive induction in HEK23T cell lines (Figures 2 and 4). Simultaneous transfection with multiple sgRNA sequences targeting multiple CYPs resulted in enhanced this effect. Co-treatment of transiently transfected cells with components to induce CYP3A4 as well as the demethylating agent 5-azacytidine also enhanced CYP3A4 expression along a dose curve (Figure 11), while co-treatment with DMSO reduced CYP3A4 in a linear dose-dependent fashion (Figure 9). We were able to generate a stably-transfected HEK293T cell line that constitutively expressed dCas9 and the MS2-P65-HSF1 activator helper complex; we named this cell line HEK-SAM (Figure 5). Although HEK-SAM did not display as great of an increase in CYP expression as the HEK293T cell line, a series of experiments with canonical bioactivated toxins indicated that the modest increases in CYP expression we achieved in HEK293T cells were sufficient to meaningfully alter the metabolic behavior of those cells. This pilot study illustrates the potential of CRISPRa technology as a drop-in modification for HTS to improve metabolic competence. Although we were unable to induce CYPs in HepG2 cells, the malleability of the HEK293T cell line makes it an excellent candidate cell line to refine and optimize this technique, which may then be readily applied to HTS systems, improving the quality of data collection.

Figures

A



B

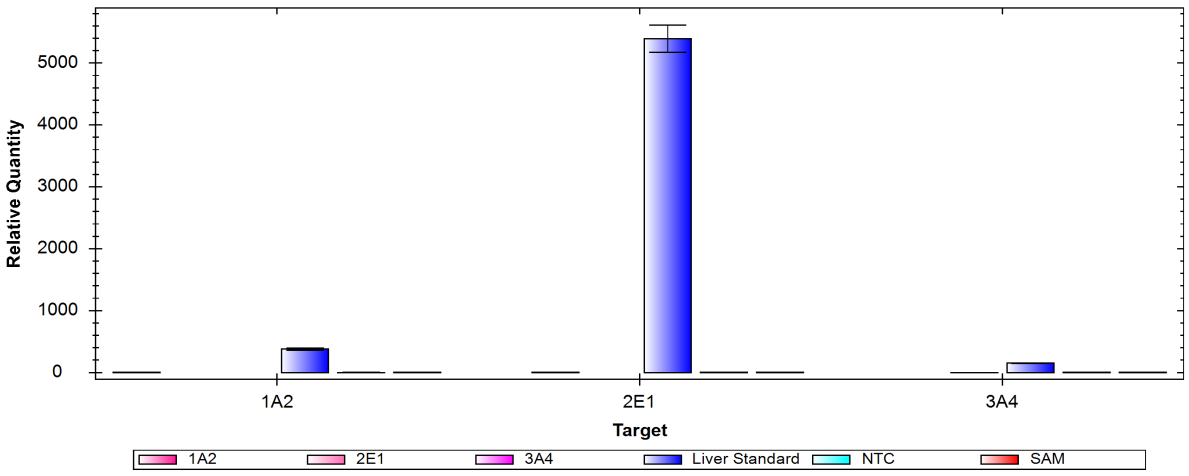
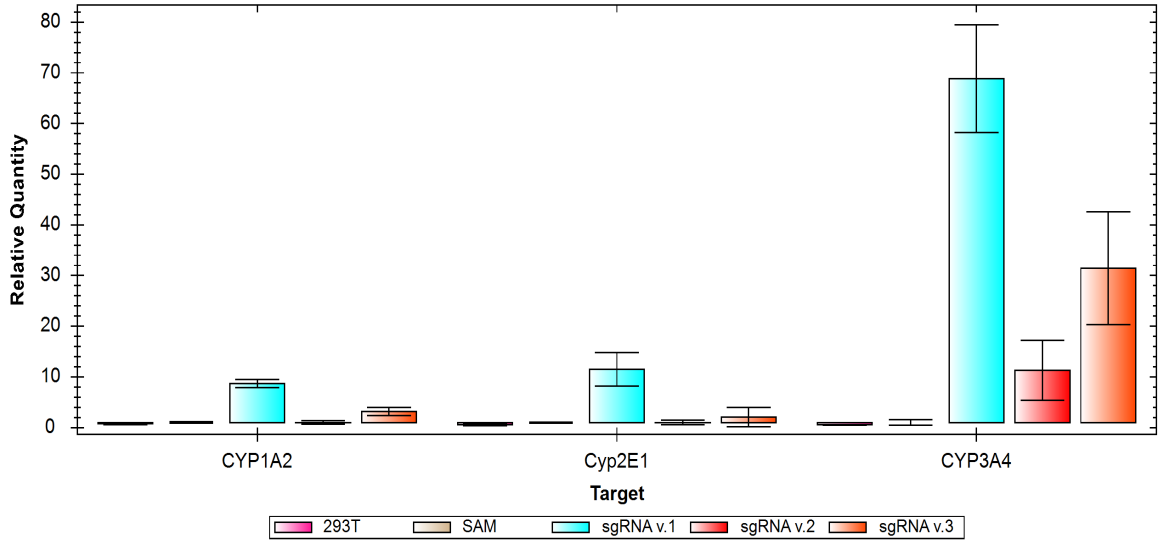


Figure 4.1: CRISPRa Tests in HepG2 Cells Transfection of HepG2 cell line with dCas9, MS2-P65-HSF1, and sgRNAs for different individual CYP1A2, CYP3A4, and CYP2E1 guide sequences. Presented are data from RT-qPCR experiments (using RPL19 to normalize expression) comparing mRNA transcripts of these CYPs versus a) the untreated cell line and b) versus RNA extracted from a primary liver cell sample. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

A



B

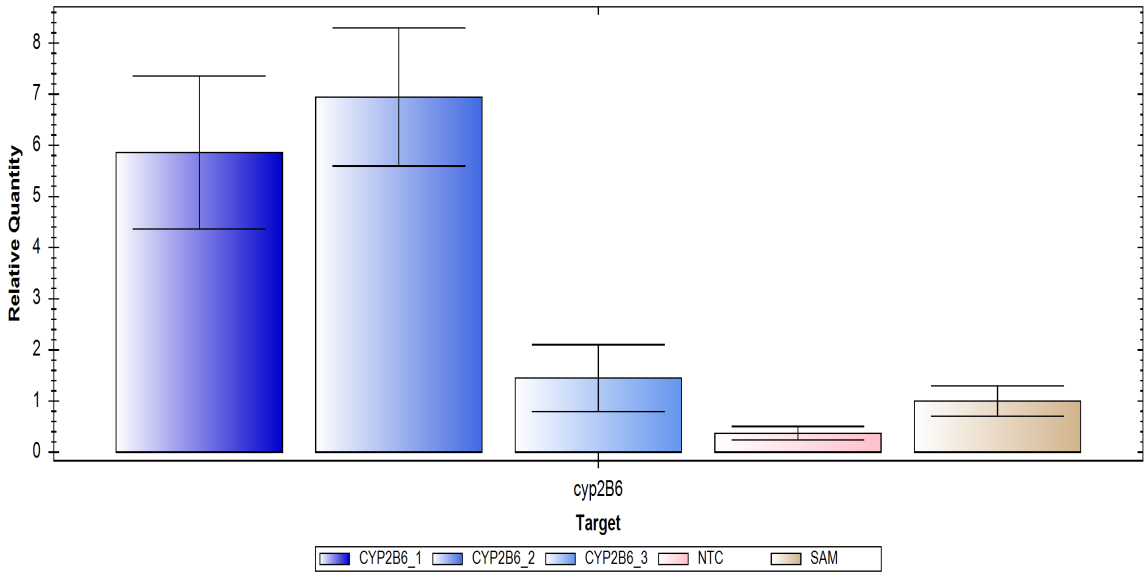


Figure 4.2: Guides Test for CYP1A2, CYP2E1, CYP3A4 and CYP2B6 in HEK293T Cells HEK293T cells were transfected with dCas9, MS2-P65-HSF1, and three different test sgRNA for A) individual CYP1A2, CYP3A4, CYP2E1 guide sequences, or B) CYP2B6 guides, to identify the most effective candidate sgRNA for inducing gene expression. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated HEK-293T cells and DMSO-treated HEK-293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

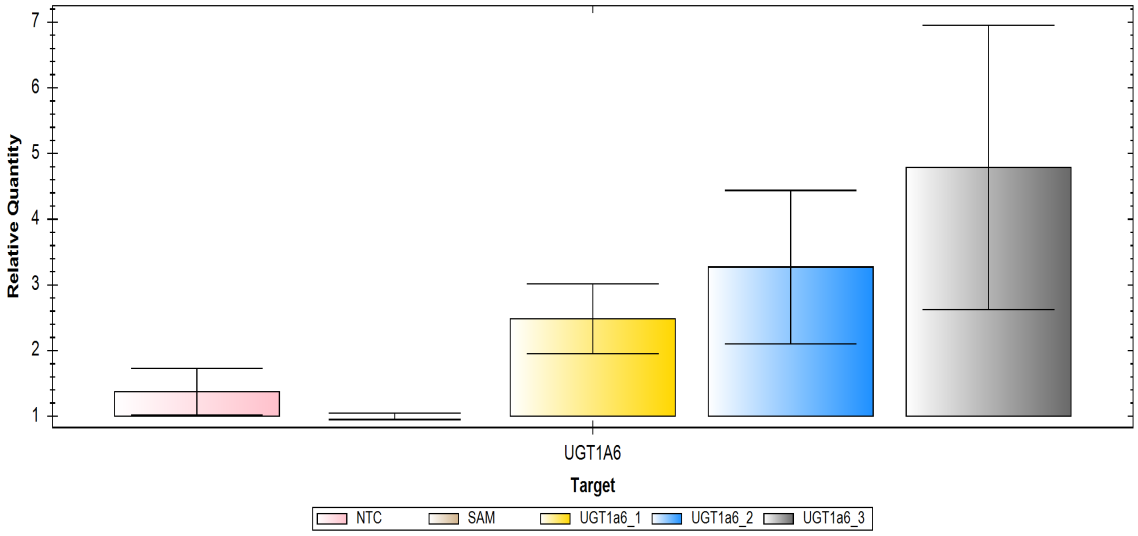


Figure 4.3: Guides Test for UGT1A6 in HEK293T Cells HEK293T cells were transfected with dCas9, MS2-P65-HSF1, and three different test sgRNA for guides, to identify the most effective candidate sgRNA for inducing gene expression. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYP and UGT genes compared to untreated HEK-293T cells and DMSO-treated HEK-293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

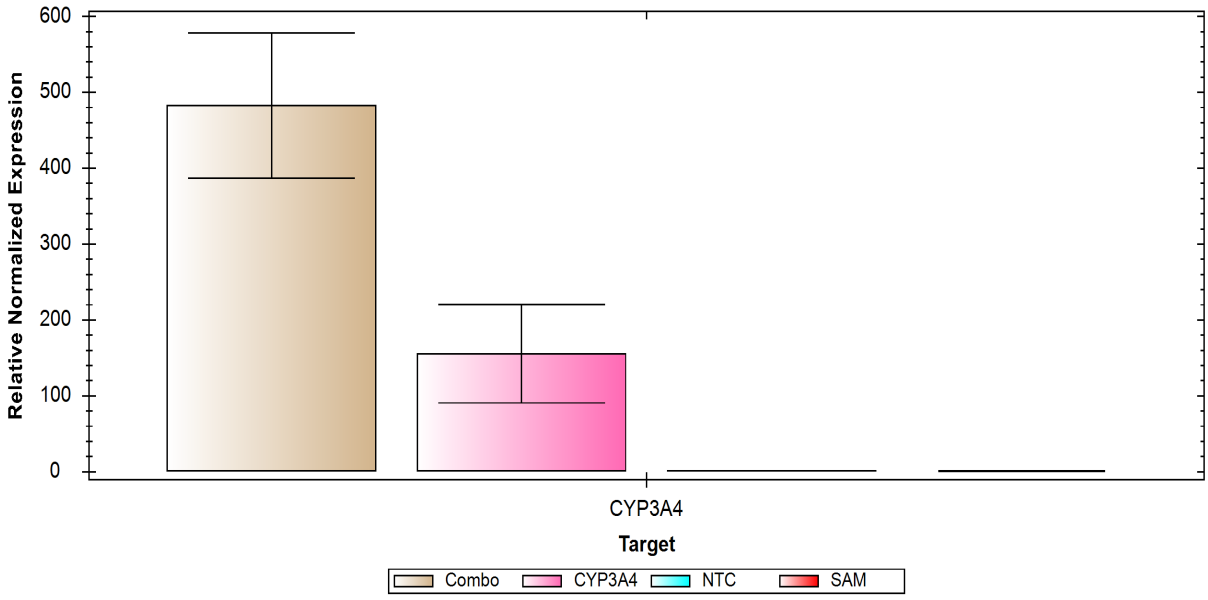


Figure 4.4: **HEK293T Cells Transfected With Multiple CYPs Versus CYP3A4 Alone** HEK293T cell line were transfected with combination of dCas9, MS2-P65-HSF1, and sgRNA for either a "combo" of CYP1A2, CYP2E1, and CYP3A4or only CYP3A4. Presented are data from RT-qPCR experiments (using RPL19 as a standard) comparing induction of CYP3A4 in cells transfected with the multi-CYP combo versus cells transfected with guides for CYP3A4 alone. The "combo" treatments resulted in increased activation for CYP3A4, suggesting a synergistic effect. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

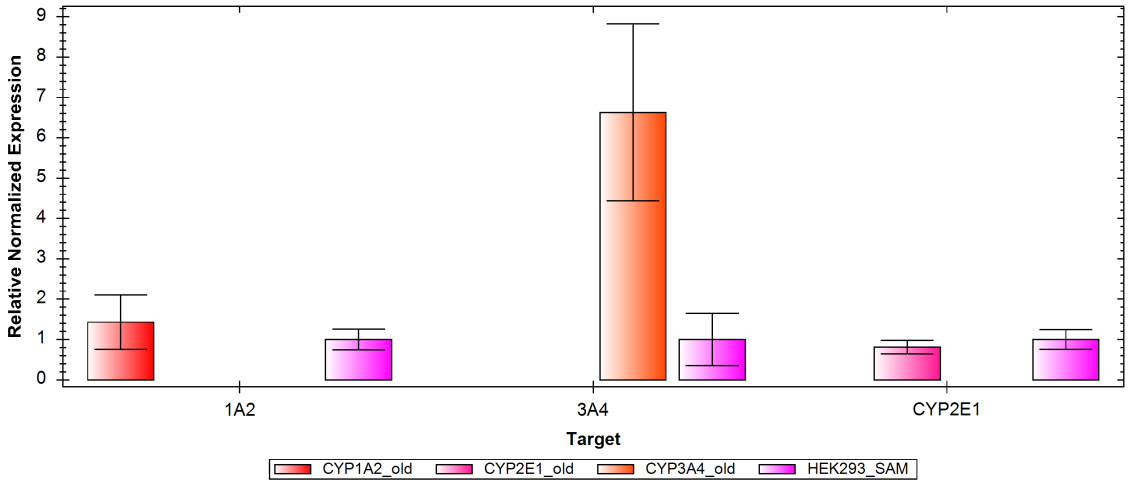


Figure 4.5: Generation of HEK-SAM, a Cell Line Stably Expressing dCas9 and MS2-P65-HSF1 Lentiviral transduction was used to create a stably-transfected cell line that constitutively expressed the dCas9 enzyme and the MS2-P65-HSF1 activator helper complex that we called HEK-SAM. The cells in this experiment were transfected only with the sgRNA for either CYP1A2, CYP2E1, or CYP3A4. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated and DMSO-treated HEK-SAM cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation. The "old" label in the figure legend indicates that the sgRNAs used in this experiment were the same as those used in Figures 2 and 3 to measure activity in HEK293T cells.

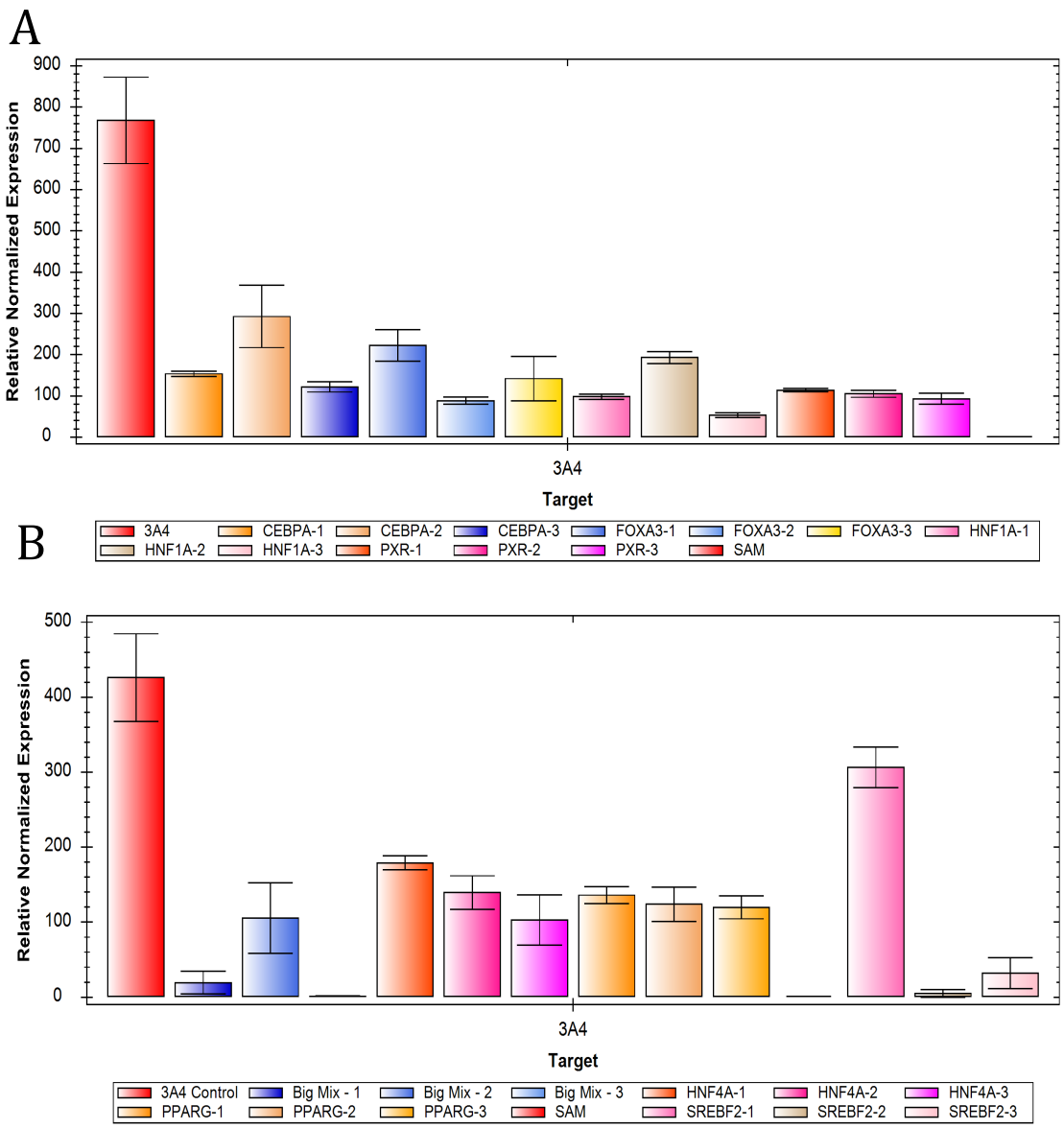


Figure 4.6: Expression of Proteins in the CYP3A4 Regulatory Pathway in the HEK293T Cell Line The HEK293T cell line was transfected with the dCas9 and MS2-P65-HSF1 plus sgRNAs for CYP3A4 and one of seven regulatory proteins as indicated in A) CEBPA, FOXA3, HNF1A, and PXR, and B) HNF4A, SREBF, or PPARγ. As in previous experiments, we tested multiple possible guide sequences per gene. The "Big Mix" condition included sgRNAs for CYP3A4 as well as FOXA3, SREBF, CEBPA, HNF1A, PXR, SREBF, HNF4A, and PPARγ. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated and DMSO-treated HEK293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

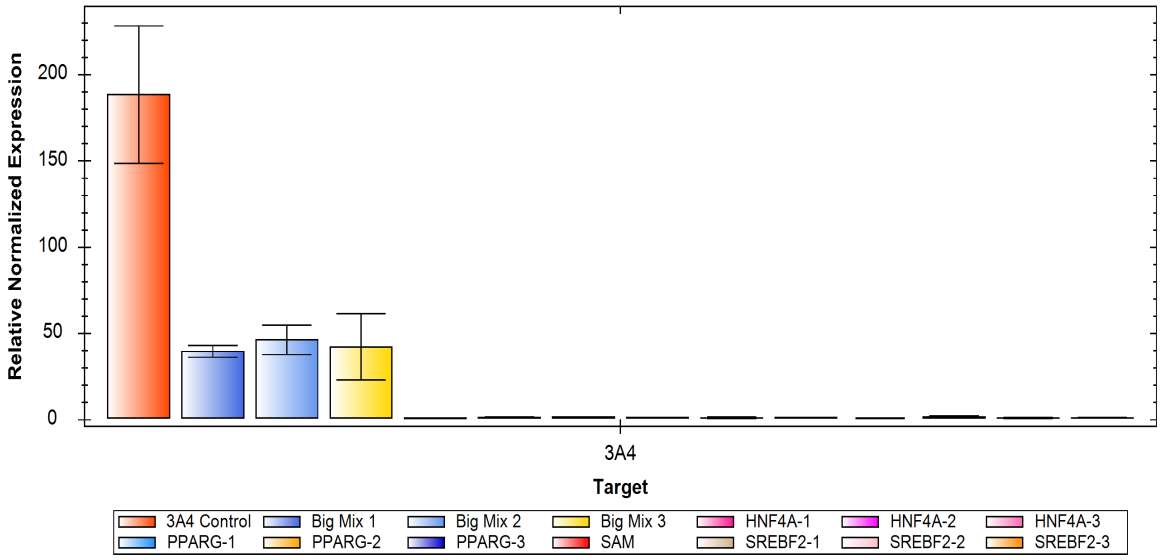


Figure 4.7: Expression of Proteins in the CYP3A4 Regulatory Pathway in the HEK-SAM Cell Line HEK-SAM cells were transfected with the sgRNAs for CYP3A4 and one of seven regulatory proteins as indicated in A) CEBPA, FOXA3, HNF1A, and PXR, and B) HNF4A, SREBF, or PPARγ. As in previous experiments, we tested multiple possible guide sequences per gene. The "Big Mix" condition included sgRNAs for CYP3A4 as well as FOXA3, SREBF, CEBPA, HNF1A, PXR, SREBF, HNF4A, and PPARγ. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated and DMSO-treated HEK293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

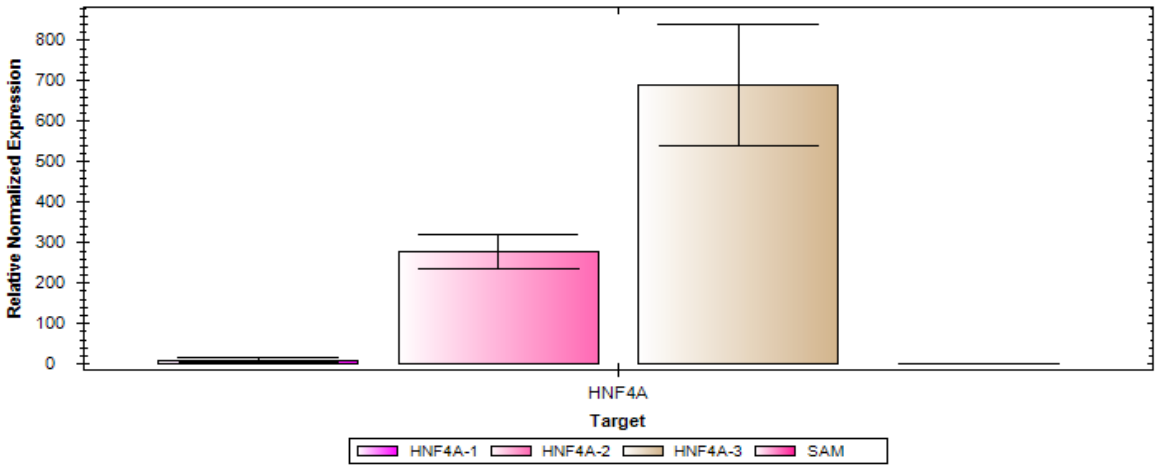
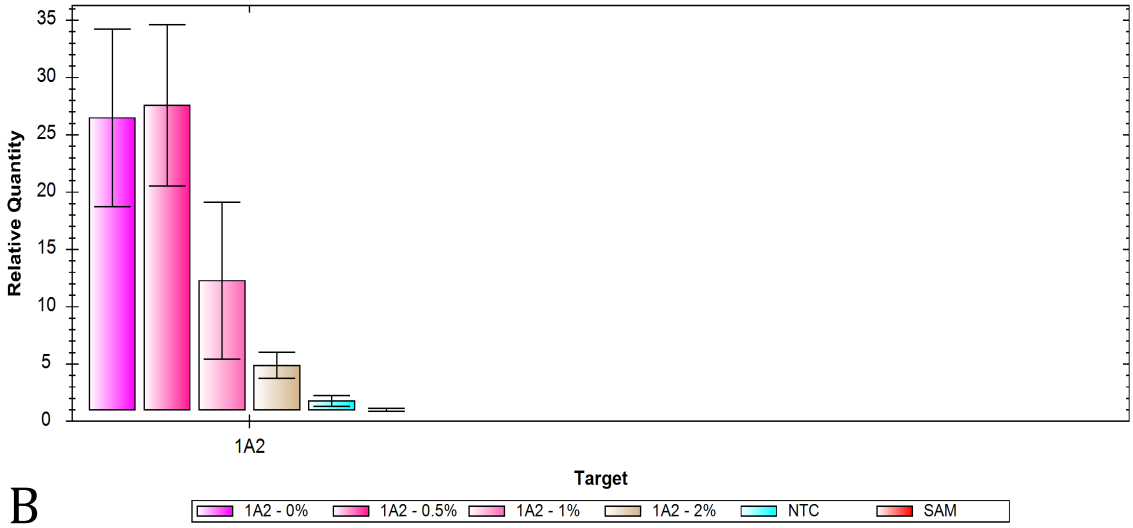


Figure 4.8: HNF4Ain the CYP3A4 Regulatory Pathway in the HEK293T Cell Line HEK293T cells were transfected dCas9 and MS2-P65-HSF1 plus sgRNAs for CYP3A4 and HNF4A. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated and DMSO-treated HEK293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

A



B

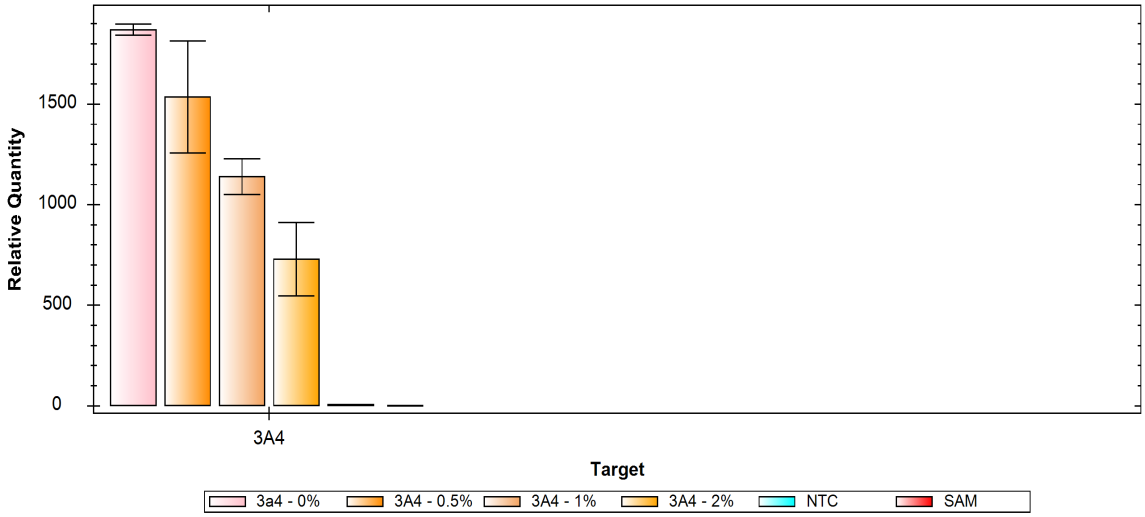


Figure 4.9: DMSO treatment Decreases CYP Expression in CRISPRa-Transfected HEK293T Cells HEK293T cells were co-treated with varying concentrations of DMSO and transfected with dCas9, MS2-P65-HSF1, and sgRNA for either A) CYP1A2 or B) CYP3A4. Percentages denote percent composition of DMSO in media. Notably, there is diminished CYP1A2 and CYP3A4 expression at 1% DMSO concentrations – concentrations not uncommonly found in cell culture studies. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated HEK-293T cells and DMSO-treated HEK-293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

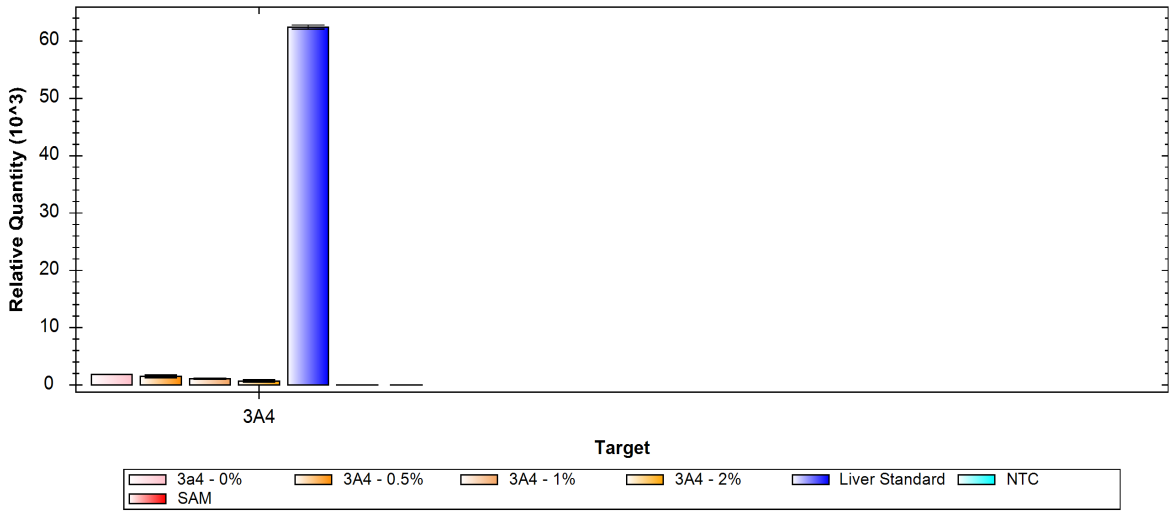


Figure 4.10: CYP Expression in DMSO-Treated CRISPRa-Transfected HEK293T Cells is Much Lower Than In Primary Liver Cells HEK293T cells were co-treated with varying concentrations of DMSO and dCas9, MS2-P65-HSF1, and sgRNA for CYP3A4. Percentages denote percent composition of DMSO in media. Neither CYP1A2 nor CYP3A4 induction was close to levels found in primary liver cells. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated HEK-293T cells and DMSO-treated HEK-293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

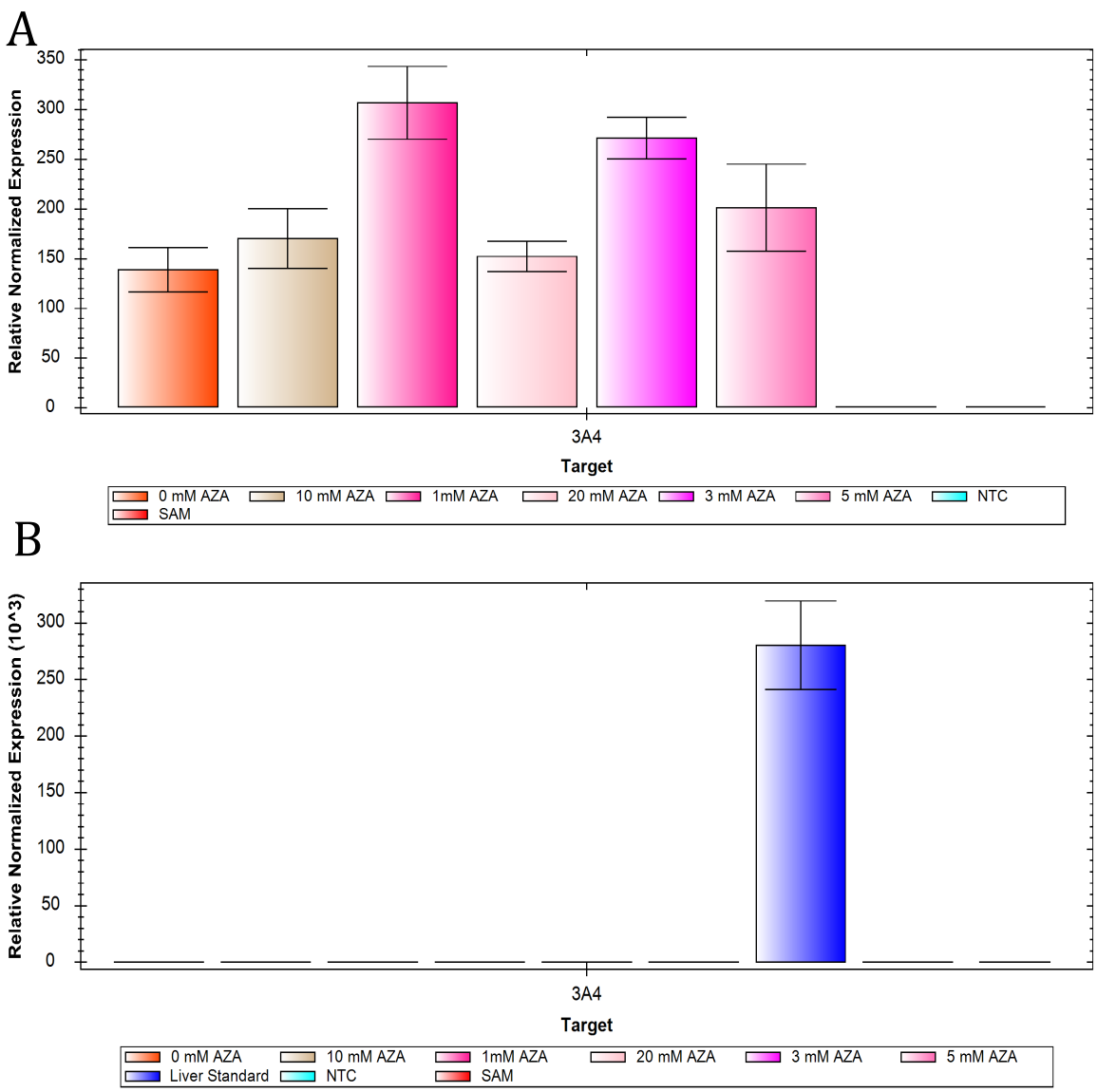


Figure 4.11: AZA Treatment Enchances CYP3A4 expression in HEK293T cells HEK293T cells were co-treated HEK293T cells transfected with dCas9, MS2-P65-HSF1, and CYP3A4, and varying amounts of Azacytidine (AZA). A) Presents a comparison ofCYP3A4 expression among the different doses of AZA, while B) compares CYP3A4 expression at these doses to CYP3A4 in primary liver samples. The optimal dosing seemed to be around the 1-3 mM AZA concentrations. However, these increases in expression still pale in comparison to the expression of CYP3A4 in primary liver cells. Presented are data from RT-qPCR experiments measuring relative quantity of mRNA transcripts for CYPs compared to untreated HEK-293T cells and DMSO-treated HEK-293T cells. RPL19 was used to normalize the mRNA between samples. Columns represent the average of 3 replicates, error bars indicate a single standard deviation. Note that although the axes of the graphs seem similar, that the *y* axis on B is expressed in multiples of 10^3.

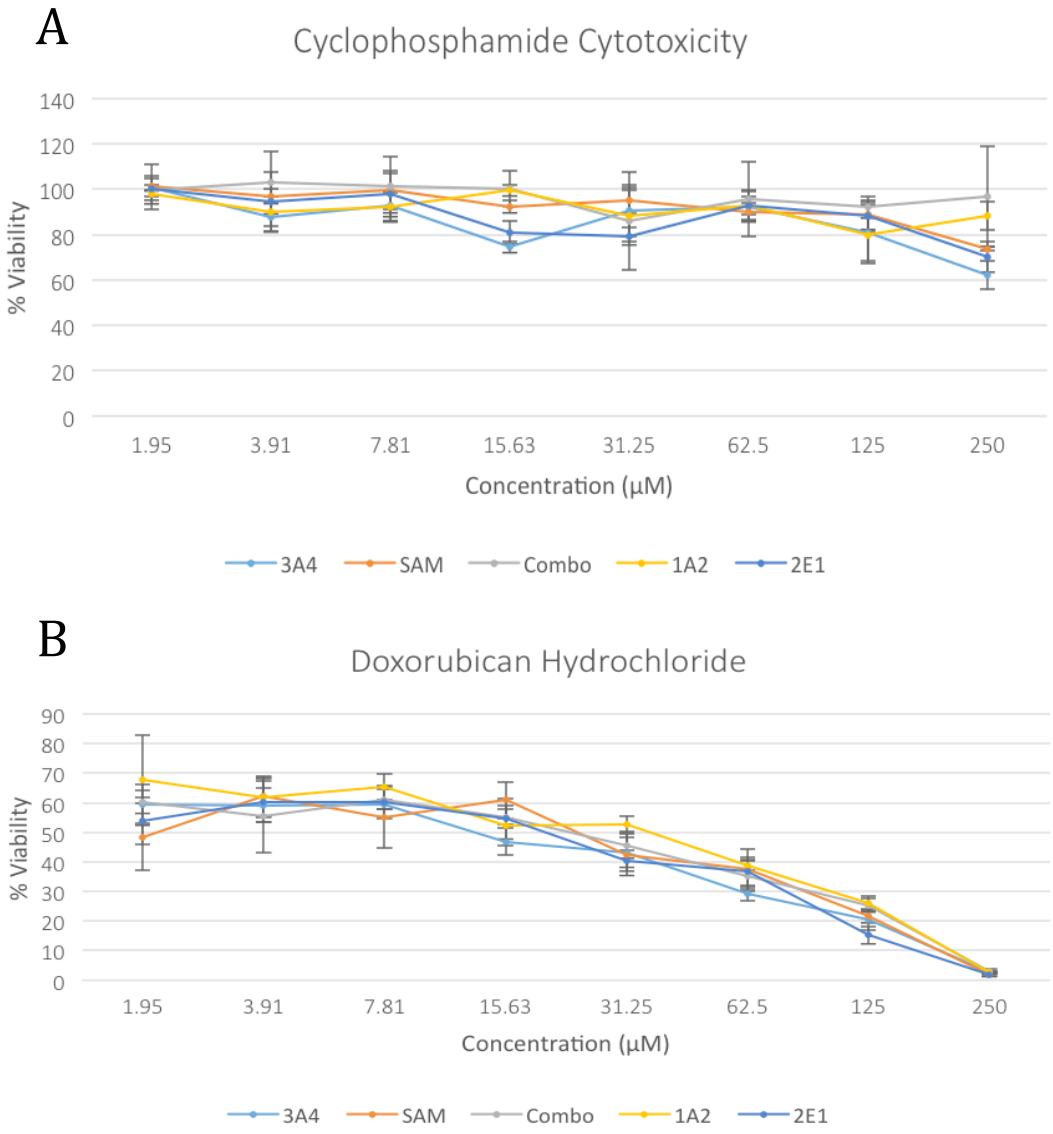


Figure 4.12: Cytotoxicity Testing with Cyclophosphamide and Doxorubican Hydrochloride HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for either CYP1A2, CYP3A4, CYP2E1, or a combination of all three (the "combo" condition). Transfected cells, DMSO-treated, and untreated controls were exposed to A) Cyclophosphamide or B) Doxorubican for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

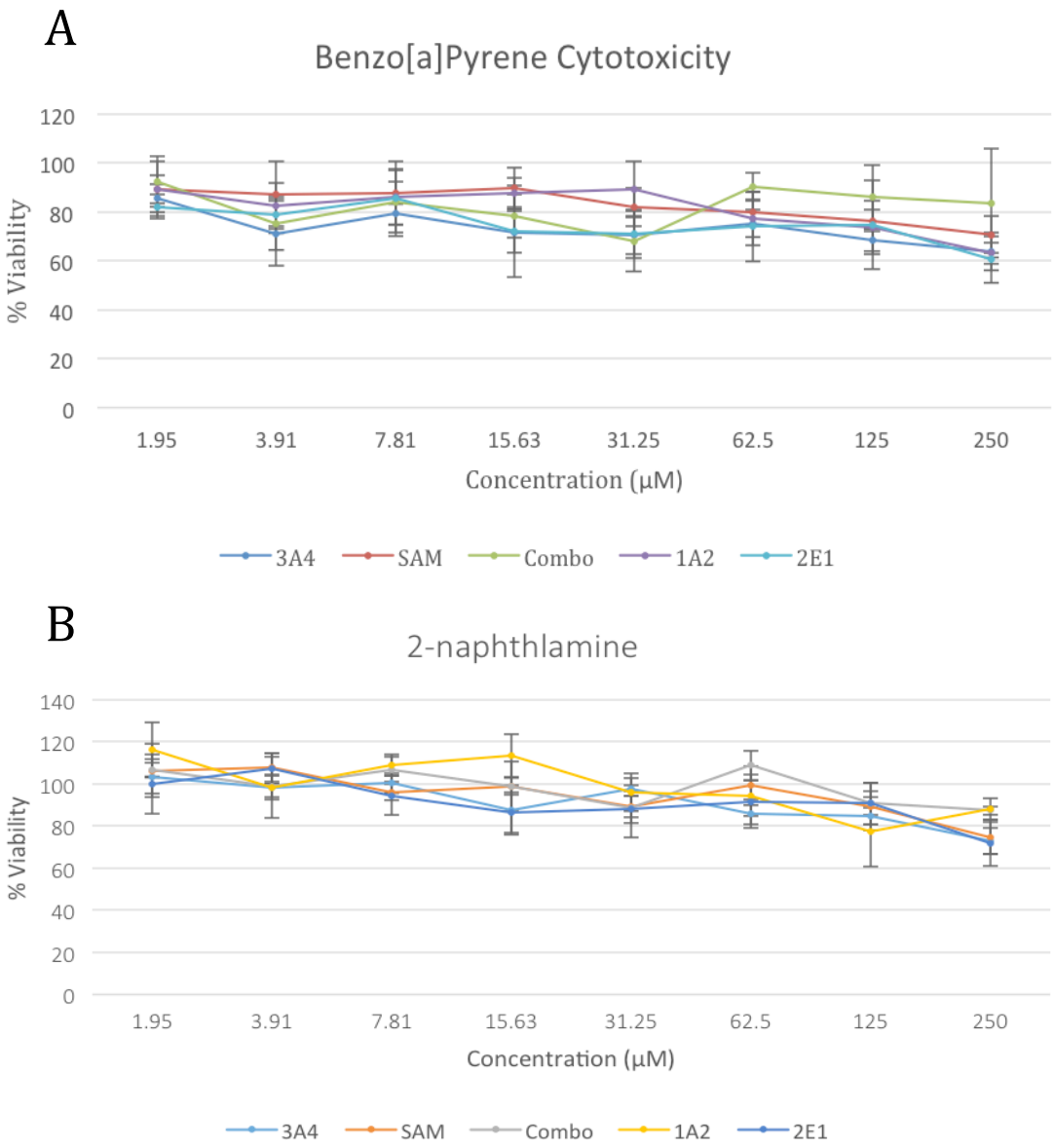


Figure 4.13: Cytotoxicity Testing with Benzo[a]Pyrene and 2-naphthlamine HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for either CYP1A2, CYP3A4, CYP2E1, or a combination of all three (the "combo" condition). Transfected cells, DMSO-treated, and untreated controls were exposed to A) Benzo[a]Pyrene or B) 2-naphthlamine for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

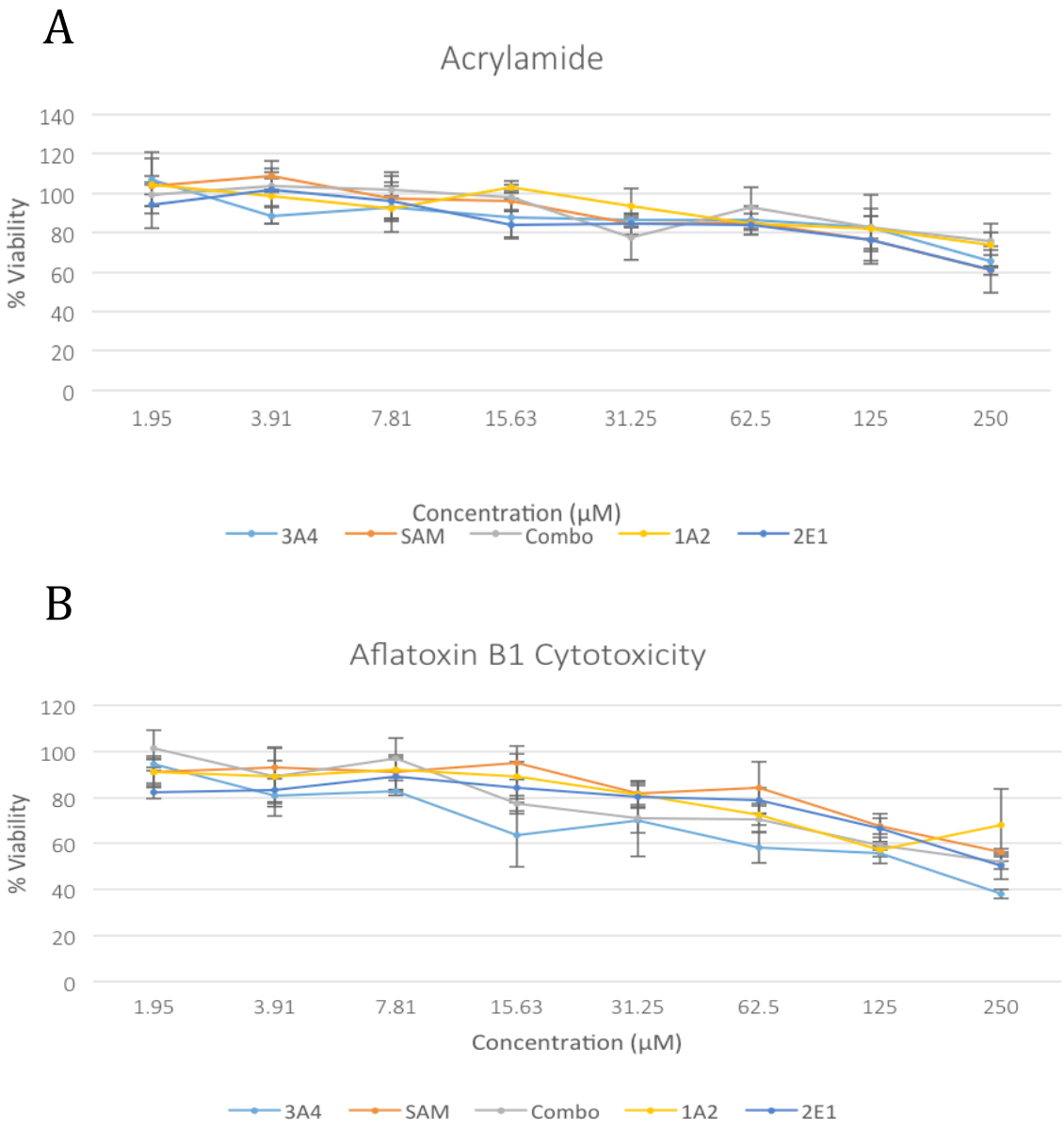


Figure 4.14: Cytotoxicity Testing with Acrylamide and Aflatoxin B1 HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for either CYP1A2, CYP3A4, CYP2E1, or a combination of all three (the "combo" condition). Transfected cells, DMSO-treated, and untreated controls were exposed to A) Acrylamide or B) Aflatoxin B1 for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

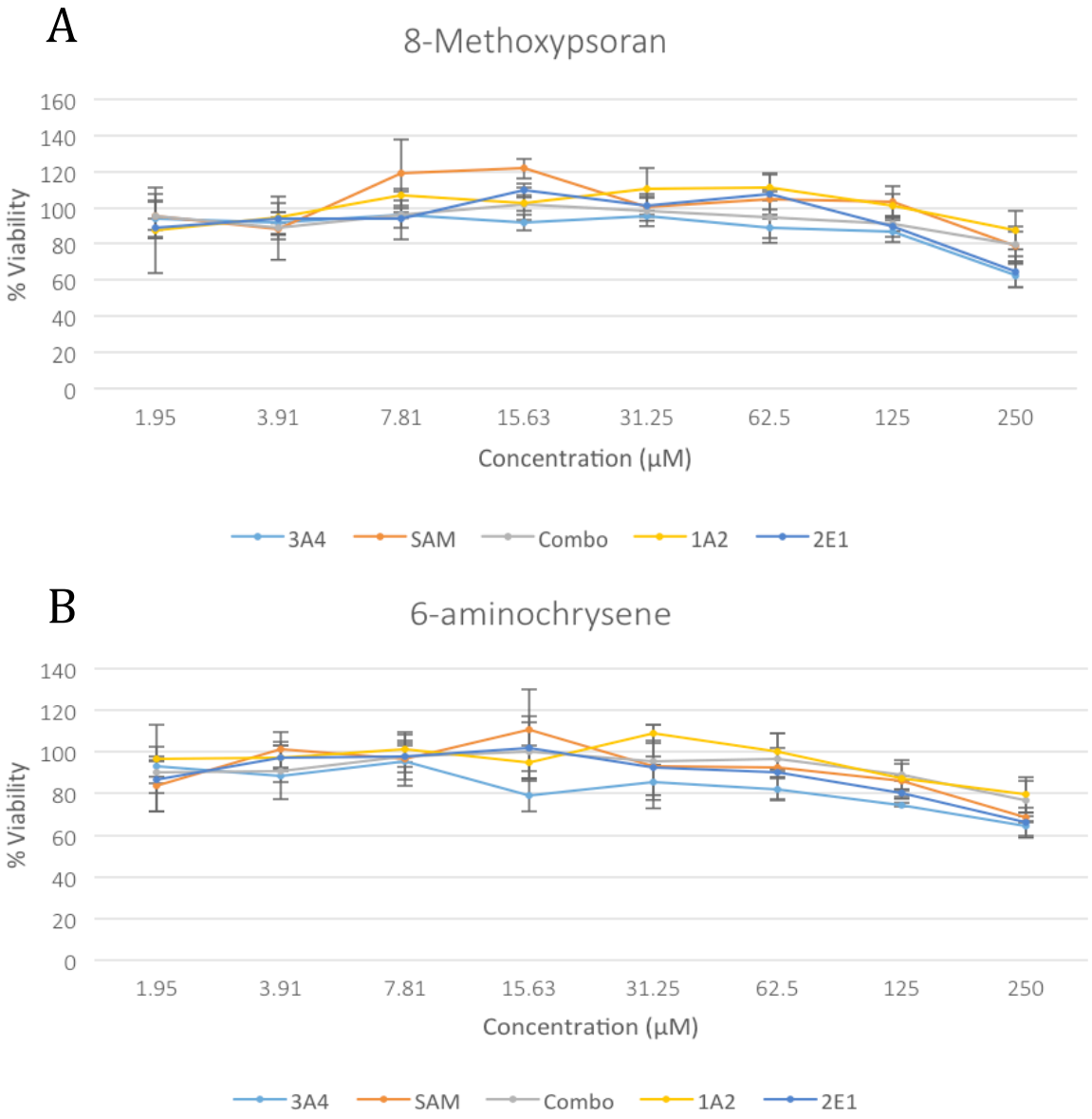


Figure 4.15: Cytotoxicity Testing with 8-Methoxypsoran and 6-aminochrysene HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for either CYP1A2, CYP3A4, CYP2E1, or a combination of all three (the "combo" condition). Transfected cells, DMSO-treated, and untreated controls were exposed to A) 8-Methoxypsoran or B) 6-aminochrysene for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

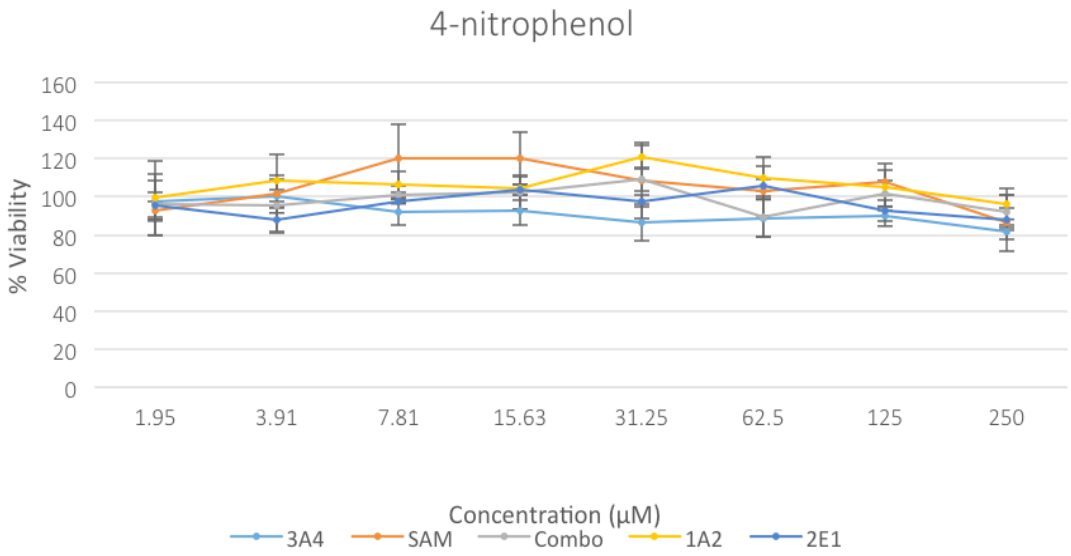


Figure 4.16: Cytotoxicity Testing with 4-nitrophenol HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for either CYP1A2, CYP3A4, CYP2E1, or a combination of all three (the "combo" condition). Transfected cells, DMSO-treated, and untreated controls were exposed to 4-nitrophenol for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

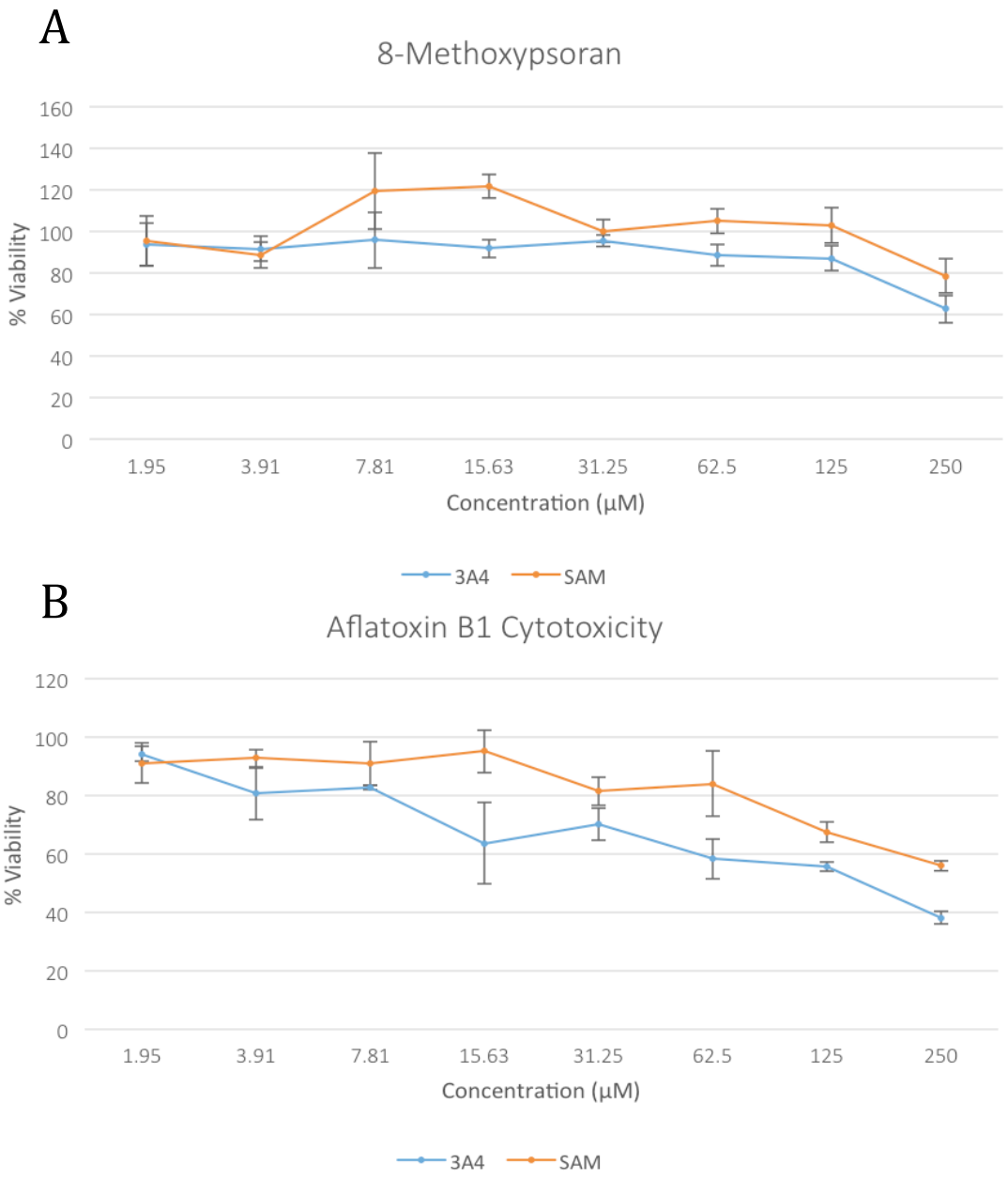


Figure 4.17: CYP3A4 Activation with CRISPRa Increases sensitivity of HEK293T Cells to 8-Methoxypsoran and Aflatoxin B1 HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for CYP3A4. Transfected cells, DMSO-treated, and untreated controls were exposed to A) 8-Methoxypsoran or B) Aflatoxin B1 for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

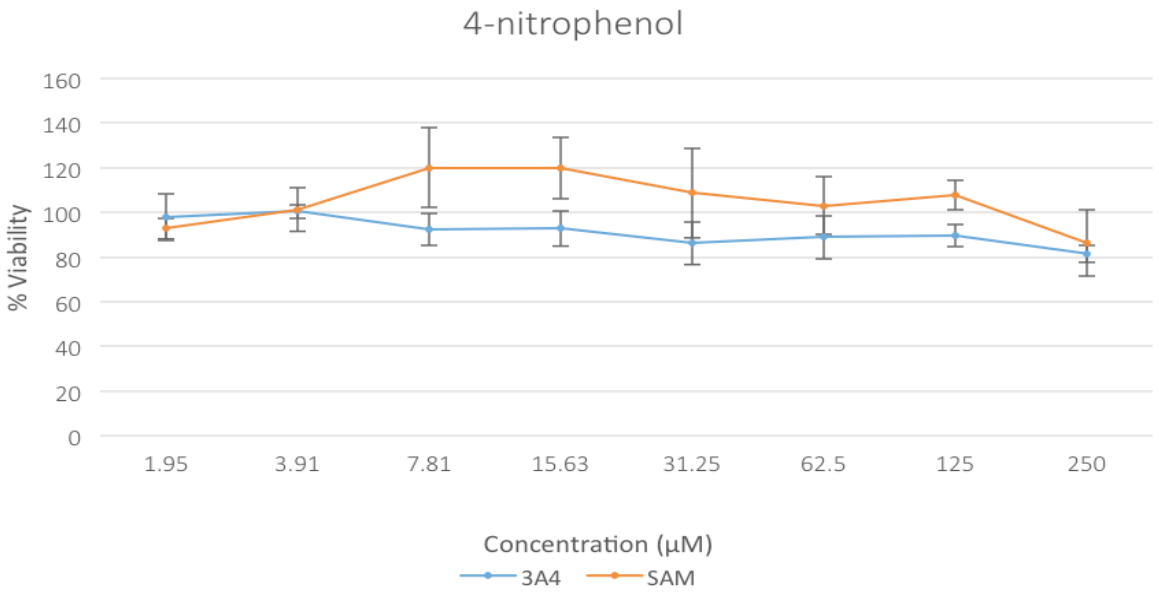


Figure 4.18: CYP3A4 Activation with CRISPRa Increases sensitivity of HEK293T Cells to 4-nitrophenol HEK293 cells were transfected with dCas9, MS2-P65-HSF1, and sgRNA for CYP3A4. Transfected cells, DMSO-treated, and untreated controls were exposed to A) 8-Methoxypsoran or B) Aflatoxin B1 for 24 hours at multiple concentrations from 1.95 µM – 250 µM. Cytotoxicity was measured using CellTiter-Glo (Promega) luminescence assays. Columns represent the average of 3 replicates, error bars indicate a single standard deviation.

Sources cited

1. Andersen, M. E. & Krewski, D. The Vision of Toxicity Testing in the 21st Century: Moving from Discussion to Action. *Toxicol. Sci.* **117,** 17–24 (2010).
2. Mulvihill, M. J., Beach, E. S., Zimmerman, J. B. & Anastas, P. T. Green Chemistry and Green Engineering: A Framework for Sustainable Technology Development. *Annu. Rev. Environ. Resour.* **36,** 271–293 (2011).
3. Guidance Document on Considerations for Waiving or Bridging of Mammalian Acute Toxicity Tests. (2016).
4. DiMasi, J. A., Grabowski, H. G. & Hansen, R. W. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J. Health Econ.* **47,** 20–33 (2016).
5. Kavlock, R. *et al.* Update on EPA's ToxCast Program: Providing High Throughput Decision Support Tools for Chemical Risk Management. *Chem. Res. Toxicol.* **25,** 1287–1302 (2012).
6. Lynch, C. *et al.* Identifying environmental chemicals as agonists of the androgen receptor by using a quantitative high-throughput screening platform. *Toxicology* **385,** 48–58 (2017).
7. European Agency For Safety At Work. REACH - Regulation for Registration, Evaluation, Authorisation and Restriction of Chemicals. (2017). Available at: https://osha.europa.eu/en/themes/dangerous-substances/reach.
8. USEPA. Highlights of Key Provisions in the Frank R. Lautenberg Chemical Safety for the 21st Century Act. *Assessing and Managing Chemicals Under TSCA* Available at: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/highlights-key-provisions-frank-r-lautenberg-chemical.
9. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. (National Academies Press, 2007).
10. Shukla, S. J., Huang, R., Austin, C. P. & Xia, M. The future of toxicity testing: a focus on in vitro methods using a quantitative high-throughput screening platform. *Drug Discov. Today* **15,** 997–1007 (2010).
11. Judson, R. S. *et al.* In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project. *Environ. Health Perspect.* **118,** 485–492 (2009).
12. Kleinstreuer, N. C. *et al.* Development and Validation of a Computational Model for Androgen Receptor Activity. *Chem. Res. Toxicol.* **30,** 946–964 (2017).
13. Rodrigues, A. D. Preclinical drug metabolism in the age of high-throughput screening: an industrial perspective. *Pharm. Res.* **14,** 1504–1510 (1997).
14. NIH Roadmap and Roadmap-affiliated Initiatives. *National Institute of Environmental Health Services* Available at: https://www.niehs.nih.gov/funding/grants/announcements/roadmap/index.cfm. (Accessed: 5th November 2017)
15. Zerhouni, E. A. Clinical research at a crossroads: the NIH roadmap. *J. Investig. Med. Off. Publ. Am. Fed. Clin. Res.* **54,** 171–173 (2006).
16. Dix, D. J. *et al.* The ToxCast Program for Prioritizing Toxicity Testing of Environmental Chemicals. *Toxicol. Sci.* **95,** 5–12 (2007).
17. US EPA, O. About the TSCA Chemical Substance Inventory. *US EPA* (2015). Available at: https://www.epa.gov/tsca-inventory/about-tsca-chemical-substance-inventory. (Accessed: 5th November 2017)

18. Sipes, N. S. *et al.* Profiling 976 ToxCast Chemicals across 331 Enzymatic and Receptor Signaling Assays. *Chem. Res. Toxicol.* **26,** 878–895 (2013).
19. Shah, I. *et al.* Using ToxCast™ Data to Reconstruct Dynamic Cell State Trajectories and Estimate Toxicological Points of Departure. *Environ. Health Perspect.* **124,** (2015).
20. Sipes, N. S. *et al.* An Intuitive Approach for Predicting Potential Human Health Risk with the Tox21 10k Library. *Environ. Sci. Technol.* **51,** 10786–10796 (2017).
21. Richard, A. M. *et al.* ToxCast Chemical Landscape: Paving the Road to 21st Century Toxicology. *Chem. Res. Toxicol.* **29,** 1225–1251 (2016).
22. Bell, C. C. *et al.* Transcriptional, Functional, and Mechanistic Comparisons of Stem Cell–Derived Hepatocytes, HepaRG Cells, and Three-Dimensional Human Hepatocyte Spheroids as Predictive In Vitro Systems for Drug-Induced Liver Injury. *Drug Metab. Dispos.* **45,** 419–429 (2017).
23. Chang, T. T. & Hughes-Fulford, M. Monolayer and Spheroid Culture of Human Liver Hepatocellular Carcinoma Cell Line Cells Demonstrate Distinct Global Gene Expression Patterns and Functional Phenotypes. *Tissue Eng. Part A* **15,** 559–567 (2009).
24. Gómez-Lechón, M. J., Tolosa, L., Conde, I. & Donato, M. T. Competency of different cell models to predict human hepatotoxic drugs. *Expert Opin. Drug Metab. Toxicol.* **10,** 1553–1568 (2014).
25. Soldatow, V. Y., LeCluyse, E. L., Griffith, L. G. & Rusyn, I. In vitro models for liver toxicity testing. *Toxicol Res* **2,** 23–39 (2013).
26. Bale, S. S., Moore, L., Yarmush, M. & Jindal, R. Emerging *In Vitro* Liver Technologies for Drug Metabolism and Inter-Organ Interactions. *Tissue Eng. Part B Rev.* **22,** 383–394 (2016).
27. Guillouzo, A. Liver cell models in in vitro toxicology. *Environ. Health Perspect.* **106,** 511–532 (1998).
28. Meng, Q. Three-dimensional culture of hepatocytes for prediction of drug-induced hepatotoxicity. *Expert Opin. Drug Metab. Toxicol.* **6,** 733–746 (2010).
29. Dunn, J. C. Y., Tompkins, R. G. & Yarmush, M. L. Long-Term in Vitro Function of Adult Hepatocytes in a Collagen Sandwich Configuration. *Biotechnol. Prog.* **7,** 237–245 (1991).
30. Mathijs, K. *et al.* Assessing the Metabolic Competence of Sandwich-Cultured Mouse Primary Hepatocytes. *Drug Metab. Dispos.* **37,** 1305–1311 (2009).
31. Westerink, W. M. A. & Schoonen, W. G. E. J. Cytochrome P450 enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. *Toxicol. In Vitro* **21,** 1581–1591 (2007).
32. Westerink, W. M. A. & Schoonen, W. G. E. J. Phase II enzyme levels in HepG2 cells and cryopreserved primary human hepatocytes and their induction in HepG2 cells. *Toxicol. In Vitro* **21,** 1592–1602 (2007).
33. Aninat, C. EXPRESSION OF CYTOCHROMES P450, CONJUGATING ENZYMES AND NUCLEAR RECEPTORS IN HUMAN HEPATOMA HepaRG CELLS. *Drug Metab. Dispos.* **34,** 75–83 (2005).
34. Tolosa, L., Donato, M. T., Pérez-Cataldo, G., Castell, J. V. & Gómez-Lechón, M. J. Upgrading cytochrome P450 activity in HepG2 cells co-transfected with

adenoviral vectors for drug hepatotoxicity assessment. *Toxicol. In Vitro* **26,** 1272–1277 (2012).
35. Tolosa, L., Gómez-Lechón, M. J., Pérez-Cataldo, G., Castell, J. V. & Donato, M. T. HepG2 cells simultaneously expressing five P450 enzymes for the screening of hepatotoxicity: identification of bioactivable drugs and the potential mechanism of toxicity involved. *Arch. Toxicol.* **87,** 1115–1127 (2013).
36. Dambach, D. M., Andrews, B. A. & Moulin, F. New Technologies and Screening Strategies for Hepatotoxicity: Use of In Vitro Models. *Toxicol. Pathol.* **33,** 17–26 (2005).
37. Gustafsson, F., Foster, A. J., Sarda, S., Bridgland-Taylor, M. H. & Kenna, J. G. A correlation between the in vitro drug toxicity of drugs to cell lines that express human P450s and their propensity to cause liver injury in humans. *Toxicol. Sci. Off. J. Soc. Toxicol.* **137,** 189–211 (2014).
38. Huang, L. *et al.* Development of an optimized cytotoxicity assay system for CYP3A4-mediated metabolic activation via modified piggyBac transposition. *Toxicol. In Vitro* **32,** 132–137 (2016).
39. Yoshitomi, S. *et al.* Establishment of the transformants expressing human cytochrome P450 subtypes in HepG2, and their applications on drug metabolism and toxicology. *Toxicol. In Vitro* **15,** 245–256 (2001).
40. Gerets, H. H. J. *et al.* Characterization of primary human hepatocytes, HepG2 cells, and HepaRG cells at the mRNA level and CYP activity in response to inducers and their predictivity for the detection of human hepatotoxins. *Cell Biol. Toxicol.* **28,** 69–87 (2012).
41. Ma, X. *et al.* Highly Efficient Differentiation of Functional Hepatocytes From Human Induced Pluripotent Stem Cells. *STEM CELLS Transl. Med.* **2,** 409–419 (2013).
42. Mitani, S. *et al.* Human ESC/iPSC-Derived Hepatocyte-like Cells Achieve Zone-Specific Hepatic Properties by Modulation of WNT Signaling. *Mol. Ther.* **25,** 1420–1433 (2017).
43. Jozefczuk, J., Prigione, A., Chavez, L. & Adjaye, J. Comparative Analysis of Human Embryonic Stem Cell and Induced Pluripotent Stem Cell-Derived Hepatocyte-Like Cells Reveals Current Drawbacks and Possible Strategies for Improved Differentiation. *Stem Cells Dev.* **20,** 1259–1275 (2010).
44. Czekaj, P. & Skowronek, R. Transcription Factors Potentially Involved in Regulation of Cytochrome P450 Gene Expression. in *Topics on Drug Metabolism* (ed. Paxton, J.) (InTech, 2012). doi:10.5772/27817
45. Rodríguez-Antona, C. *et al.* Cytochrome P450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells. *Xenobiotica* **32,** 505–520 (2002).
46. Kim, I.-W., Han, N., Burckart, G. J. & Oh, J. M. Epigenetic Changes in Gene Expression for Drug-Metabolizing Enzymes and Transporters. *Pharmacother. J. Hum. Pharmacol. Drug Ther.* **34,** 140–150 (2014).
47. Gailhouste, L. *et al.* Epigenetic Reprogramming of Human Hepatoma Cells: A Low-Cost Option for Drug Metabolism Assessment. *Cell. Mol. Gastroenterol. Hepatol.* (2017). doi:10.1016/j.jcmgh.2017.11.006

48. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517,** 583–588 (2014).
49. Ueda, R. *et al.* A Common Regulatory Region Functions Bidirectionally in Transcriptional Activation of the Human CYP1A1 and CYP1A2 Genes. *Mol. Pharmacol.* **69,** 1924–1930 (2006).
50. Reed, J. R. & Backes, W. L. The functional effects of physical interactions involving cytochromes P450: putative mechanisms of action and the extent of these effects in biological membranes. *Drug Metab. Rev.* **48,** 453–469 (2016).
51. Guengerich, F. P. A malleable catalyst dominates the metabolism of drugs. *Proc. Natl. Acad. Sci.* **103,** 13565–13566 (2006).
52. Guengerich, F. P. *et al.* Heterologous expression of human drug-metabolizing enzymes. *Drug Metab. Dispos. Biol. Fate Chem.* **25,** 1234–1241 (1997).
53. Cas9 Activator Tool. (2016). Available at: http://sam.genome-engineering.org/database.
54. (EN) - QIAGEN Plasmid Purification Handbook — April 2012. Available at: https://www.qiagen.com/us/resources/resourcedetail?id=46205595-0440-459e-9d93-50eb02e5707e&lang=en. (Accessed: 8th November 2017)
55. Thomas, P. & Smart, T. G. HEK293 cell line: A vehicle for the expression of recombinant proteins. *J. Pharmacol. Toxicol. Methods* **51,** 187–200 (2005).
56. Dai, D. P. *et al.* 293FT is a highly suitable mammalian cell line for the in vitro enzymatic activity analysis of typical P450 proteins. 33–37 (2015). doi:10.1691/ph.2015.4067
57. Snykers, S. *et al.* Role of epigenetics in liver-specific gene transcription, hepatocyte differentiation and stem cell reprogrammation. *J. Hepatol.* **51,** 187–211 (2009).
58. Aitken, A. E., Richardson, T. A. & Morgan, E. T. REGULATION OF DRUG-METABOLIZING ENZYMES AND TRANSPORTERS IN INFLAMMATION. *Annu. Rev. Pharmacol. Toxicol.* **46,** 123–149 (2006).
59. Burns, K. E., Shepherd, P., Finlay, G., Tingle, M. D. & Helsby, N. A. Indirect regulation of CYP2C19 gene expression via DNA methylation. *Xenobiotica* **0,** 1–12 (2017).
60. Chen, F. *et al.* Up-Regulating CYP3A4 Expression in C3A Cells by Transfection with a Novel Chimeric Regulator of hPXR-p53-AD. *PLoS ONE* **9,** e95752 (2014).
61. Kumagai, T. *et al.* Indirubin, a component of Ban-Lan-Gen, activates CYP3A4 gene transcription through the human pregnane X receptor. *Drug Metab. Pharmacokinet.* **31,** 139–145 (2016).
62. Horton, J. D., Goldstein, J. L. & Brown, M. S. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.* **109,** 1125–1131 (2002).
63. Tokizane, T. *et al.* Cytochrome P450 1B1 Is Overexpressed and Regulated by Hypomethylation in Prostate Cancer. *Clin. Cancer Res.* **11,** 5793–5801 (2005).
64. Jo, W. J. *et al.* Comparative Functional Genomic Analysis Identifies Distinct and Overlapping Sets of Genes Required for Resistance to Monomethylarsonous Acid (MMAIII) and Arsenite (AsIII) in Yeast. *Toxicol. Sci.* **111,** 424–436 (2009).

65. Habano, W. *et al.* Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes. *Clin. Epigenetics* **7,** 105 (2015).
66. Kacevska, M. *et al.* DNA methylation dynamics in the hepatic CYP3A4 gene promoter. *Biochimie* **94,** 2338–2344 (2012).
67. Nakamura, K., Aizawa, K., Aung, K. H., Yamauchi, J. & Tanoue, A. Zebularine upregulates expression of CYP genes through inhibition of DNMT1 and PKR in HepG2 cells. *Sci. Rep.* **7,** (2017).
68. Garzon, R. *et al.* MicroRNA-29b induces global DNA hypomethylation and tumor suppressor gene reexpression in acute myeloid leukemia by targeting directly DNMT3A and 3B and indirectly DNMT1. *Blood* **113,** 6411–6418 (2009).
69. Xiang, Y. *et al.* MiR-152 and miR-185 co-contribute to ovarian cancer cells cisplatin sensitivity by targeting DNMT1 directly: a novel epigenetic therapy independent of decitabine. *Oncogene* **33,** 378 (2014).

Appendix 1: Mutants with altered growth in 2,5-DMF

2,5-DMF 5G Resistant

2,5-DMF 5G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YKL020C	2.917449	*BUD9*	Protein involved in bud-site selection; mutant has increased aneuploidy tolerance
YFR022W	3.494680	*PRM10*	Pheromone-regulated protein; proposed to be involved in mating
YGR041W	2.608475	*ROG3*	Alpha-arrestin involved in ubiquitin-dependent endocytosis
YCR073W-A	3.416565	*SOL2*	Protein with possible role in tRNA export
YJL108C	1.943484	*SPT23*	ER membrane protein involved in regulation of OLE1 transcription
YCR006C	1.008867	-	Putative protein of unknown function
YNL058C	5.424345	-	Putative protein of unknown function

2,5-DMF 10G Resistant

2,5-DMF 10G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YNR067C	0.625615	*DSE4*	Daughter cell-specific secreted protein with similarity to glucanases; degrades cell wall from the daughter side causing daughter to separate from mother

2,5-DMF 15G Resistant

2,5-DMF 15G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YPL154C	0.857534	*PEP4*	Vacuolar aspartyl protease (proteinase A); required for posttranslational precursor maturation of vacuolar proteinases; important for protein turnover after oxidative damage

2,5-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YDL072C	1.535777	*VPS21*	Endosomal Rab family GTPase; required for endocytic transport and sorting of vacuolar hydrolases; required for endosomal localization of the CORVET complex; required with YPT52 for MVB biogenesis and sorting; involved in autophagy and ionic stress tolerance;
YMR090W	0.944358	*YET3*	Protein of unknown function; YET3 null mutant decreases the level of secreted invertase; homolog of human BAP31 protein
YDL086W	0.716625	-	Putative protein of unknown function
YOR089C	0.975989	-	Putative carboxymethylenebutenolidase
YBR242W	1.124726	-	Putative protein of unknown function

2,5-DMF 5G Sensitive

2,5-DMF 5G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YML007W	-1.280161	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress

2,5-DMF 10G Sensitive

2,5-DMF 10G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YCR086W	-2.273000	*CSM1*	Nucleolar protein that mediates homolog segregation during meiosis I
YIL154C	-1.704455	*IMP2'*	Transcriptional activator involved in maintenance of ion homeostasis; also involved in protection against DNA damage caused by bleomycin and other oxidants
YAL024C	-3.986959	*LTE1*	Protein similar to GDP/GTP exchange factors; without detectable GEF activity
YKL064W	-1.114260	*MNR2*	Vacuolar membrane protein required for magnesium homeostasis; putative magnesium transporter
YHR179W	-1.100554	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN)

2,5-DMF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR463W	-3.289455	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YBR069C	-1.810932	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine; low-affinity tryptophan and histidine transporter
YML007W	-1.223951	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YLR202C	-2.858706	undefined ORF	Dubious open reading frame
YGR035C	-1.696455	undefined ORF	Putative protein of unknown function

2,5-DMF 15G Sensitive

2,5-DMF 15G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YKL114C	-6.262799	*APN1*	Major apurinic/apyrimidinic endonuclease; 3'-repair diesterase; involved in repair of DNA damage by oxidation and alkylating agents; also functions as a 3'-5' exonuclease to repair 7,8-dihydro-8-oxodeoxyguanosine
YCR086W	-3.806423	*CSM1*	Nucleolar protein that mediates homolog segregation during meiosis I
YKL213C	-1.209951	*DOA1*	WD repeat protein required for ubiquitin-mediated protein degradation
YIL065C	-1.413206	*FIS1*	Protein involved in mitochondrial fission and peroxisome abundance
YGL020C	-2.892588	*GET1*	Subunit of the GET complex; involved in insertion of proteins into the ER membrane
YKR019C	-4.003128	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YDR439W	-2.934677	*LRS4*	Nucleolar protein that forms a complex with Csm1p
YGL035C	-0.645629	*MIG1*	Transcription factor involved in glucose repression
YEL007W	-1.205716	*MIT1*	Transcriptional regulator of pseudohyphal growth
YKL064W	-1.515564	*MNR2*	Vacuolar membrane protein required for magnesium homeostasis

2,5-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YHR179W	-1.582784	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN)
YDR004W	-1.718706	*RAD57*	Protein that stimulates strand exchange; stimulates strand exchange by stabilizing the binding of Rad51p to single-stranded DNA
YJL004C	-2.258376	*SYS1*	Integral membrane protein of the Golgi
YBR069C	-2.170562	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YCR053W	-1.661451	*THR4*	Threonine synthase
YOL018C	-2.169236	*TLG2*	Syntaxin-like t-SNARE
YEL012W	-5.053212	*UBC8*	Ubiquitin-conjugating enzyme that regulates gluconeogenesis
YNL054W	-2.484237	*VAC7*	Integral vacuolar membrane protein
YIL017C	-2.583924	*VID28*	GID Complex subunit, serves as adaptor for regulatory subunit Vid24p
YER072W	-1.039239	*VTC1*	Regulatory subunit of the vacuolar transporter chaperone (VTC) complexvacuolar fusion
YOR043W	-1.189910	*WHI2*	Protein required for full activation of the general stress response
YHR134W	-4.384623	*WSS1*	Metalloprotease involved in DNA repair, removes DNA-protein crosslinks at stalled replication forks during replication of damaged DNA
YML007W	-1.510256	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YOR364W	-1.755933	-	Dubious open reading frame
YIL077C	-1.659320	-	Putative protein of unknown function

2,5-DMF Sensitive across all time points

2,5-DMF Sensitive Strains Across All Time Points		
Deleted ORF Name	**Deleted Gene Name**	**Deleted Gene Function**
YML007W	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress

Appendix 2: Mutants with altered growth in 2,3-DMF

2,3-DMF 5G Resistant

2,3-DMF 5G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YGR037C	0.838921	*ACB1*	Acyl-CoA-binding protein
YLR131C	1.665009	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YDL073W	0.998191	*AHK1*	Scaffold protein in the HKR1 sub-branch of the Hog1p-signaling pathway
YHR126C	1.393973	*ANS1*	Putative GPI protein; SWAT-GFP and mCherry fusion proteins localize to the vacuole; transcription dependent upon Azf1p
YPR128C	1.193388	*ANT1*	Peroxisomal adenine nucleotide transporter; involved in beta-oxidation of medium-chain fatty acid; required for peroxisome proliferation
YDR275W	1.322754	*BSC2*	Protein of unknown function
YML042W	1.021996	*CAT2*	Carnitine acetyl-CoA transferase; present in both mitochondria and peroxisomes
YPR013C	0.857573	*CMR3*	Putative zinc finger protein; YPR013C is not an essential gene
YIR004W	0.977763	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YKL213C	0.966467	*DOA1*	WD repeat protein required for ubiquitin-mediated protein degradation
YDR519W	1.466857	*FPR2*	Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase)
YLL043W	1.801699	*FPS1*	Aquaglyceroporin, plasma membrane channel
YER145C	1.058331	*FTR1*	High affinity iron permease; involved in the transport of iron across the plasma membrane
YDR507C	2.010428	*GIN4*	Protein kinase involved in bud growth and assembly of the septin ring
YER020W	0.604779	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YPR179C	0.892709	*HDA3*	Subunit of the HDA1 histone deacetylase complex
YHR158C	0.764921	*KEL1*	Protein required for proper cell fusion and cell morphology
YLR239C	1.309906	*LIP2*	Lipoyl ligase; involved in the modification of mitochondrial enzymes by the attachment of lipoic acid groups

2,3-DMF 5G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YLL007C	1.138447	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility); upstream component for regulation through the small GTPase Rho5p
YGL087C	2.522190	*MMS2*	Ubiquitin-conjugating enzyme variant; involved in error-free postreplication repair; forms a heteromeric complex with Ubc13p, an active ubiquitin-conjugating enzyme
YPL013C	0.909683	*MRPS16*	Mitochondrial ribosomal protein of the small subunit
YHR195W	0.886093	*NVJ1*	Nuclear envelope protein; anchored to the nuclear inner membrane
YBR129C	0.855971	*OPY1*	Protein of unknown function; overproduction blocks cell cycle arrest in the presence of mating pheromone
YJL023C	0.878661	*PET130*	Protein required for respiratory growth; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies
YBL068W	0.806861	*PRS4*	5-phospho-ribosyl-1(alpha)-pyrophosphate synthetase, synthesizes PRPP
YJL078C	0.595059	*PRY3*	Cell wall-associated protein involved in export of acetylated sterols
YDR419W	0.988646	*RAD30*	DNA polymerase eta; involved in translesion synthesis during post-replication repair
YDR279W	1.105300	*RNH202*	Ribonuclease H2 subunit; required for RNase H2 activity
YLR325C	1.178127	*RPL38*	Ribosomal 60S subunit protein L38; homologous to mammalian ribosomal protein L38, no bacterial homolog
YDR389W	1.282906	*SAC7*	GTPase activating protein (GAP) for Rho1p
YMR140W	0.617454	*SIP5*	Protein of unknown function; interacts with both the Reg1p/Glc7p phosphatase and the Snf1p kinase
YNL086W	0.848602	*SNN1*	Subunit of the BLOC-1 complex involved in endosomal maturation
YER115C	1.027114	*SPR6*	Protein of unknown function
YCL037C	1.392603	*SRO9*	Cytoplasmic RNA-binding protein; shuttles between nucleus and cytoplasm and is exported from the nucleus in an mRNA export-dependent manner
YDR297W	0.903066	*SUR2*	Sphinganine C4-hydroxylase; catalyses the conversion of sphinganine to phytosphingosine in sphingolipid biosyntheis
YDR334W	1.759848	*SWR1*	Swi2/Snf2-related ATPase; structural component of the SWR1 complex, which exchanges histone variant H2AZ (Htz1p) for chromatin-bound histone H2A
YPR156C	1.024701	*TPO3*	Polyamine transporter of the major facilitator superfamily; member of the 12-spanner drug:H(+) antiporter DHA1 family

YDR092W	0.988425	*UBC13*	E2 ubiquitin-conjugating enzyme; involved in the error-free DNA postreplication repair pathway
YBR058C	0.701915	*UBP14*	Ubiquitin-specific protease
YDL091C	0.999012	*UBX3*	Clathrin-coated vesicle component, regulator of endocytosis
YDL034W	0.650114	-	Dubious open reading frame
YLL020C	0.811845	-	Dubious open reading frame
YKR018C	0.967791	-	Protein of unknown function
YCR022C	1.393580	-	Protein of unknown function

2,3-DMF 10G Resistant

2,3-DMF 10G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YCR088W	0.705663	*ABP1*	Actin-binding protein of the cortical actin cytoskeleton; important for activation of the Arp2/3 complex that plays a key role actin in cytoskeleton organization
YNR033W	1.125465	*ABZ1*	Para-aminobenzoate (PABA) synthase
YGR037C	1.036195	*ACB1*	Acyl-CoA-binding protein; transports newly synthesized acyl-CoA esters from fatty acid synthetase (Fas1p-Fas2p) to acyl-CoA-consuming processes
YGL180W	1.516110	*ATG1*	Protein serine/threonine kinase; required for vesicle formation in autophagy and the cytoplasm-to-vacuole targeting (Cvt) pathway
YOR152C	0.623584	*ATG40*	Autophagy receptor with a role in endoplasmic reticulum degradation; involved specifically in autophagy of cortical and cytoplasmic ER in response to nitrogen starvation or rapamycin treatment
YCR032W	1.155791	*BPH1*	Protein homologous to Chediak-Higashi syndrome and Beige proteins
YDR275W	1.520321	*BSC2*	Protein of unknown function
YMR055C	0.877054	*BUB2*	Mitotic exit network regulator
YML042W	1.078828	*CAT2*	Carnitine acetyl-CoA transferase; present in both mitochondria and peroxisomes
YPR013C	0.906810	*CMR3*	Putative zinc finger protein
YER130C	0.989395	*COM2*	Transcription factor that binds IME1 Upstream Activation Signal (UAS)ru
YKR034W	1.193822	*DAL80*	Negative regulator of genes in multiple nitrogen degradation pathways
YIR004W	0.981755	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YDL174C	0.546207	*DLD1*	Major mitochondrial D-lactate dehydrogenase

2,3-DMF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YKL213C	1.243369	*DOA1*	WD repeat protein required for ubiquitin-mediated protein degradation
YIL064W	0.956157	*EFM4*	Lysine methyltransferase
YCL045C	1.181235	*EMC1*	Member of conserved endoplasmic reticulum membrane complex
YPR037C	0.824014	*ERV2*	Flavin-linked sulfhydryl oxidase localized to the ER lumen; involved in disulfide bond formation within the endoplasmic reticulum (ER)
YIL065C	0.901891	*FIS1*	Protein involved in mitochondrial fission and peroxisome abundance
YDR519W	1.727919	*FPR2*	Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase); binds to the drugs FK506 and rapamycin
YAL022C	1.231956	*FUN26*	High affinity, broad selectivity, nucleoside/nucleobase transporter
YER020W	1.076537	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YDL022W	1.325735	*GPD1*	NAD-dependent glycerol-3-phosphate dehydrogenase; key enzyme of glycerol synthesis, essential for growth under osmotic stress
YKL109W	1.281778	*HAP4*	Transcription factor; subunit of the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex
YPL001W	0.856329	*HAT1*	Catalytic subunit of the Hat1p-Hat2p histone acetyltransferase complex
YDR305C	1.133255	*HNT2*	Dinucleoside triphosphate hydrolase
YDR258C	0.954653	*HSP78*	Oligomeric mitochondrial matrix chaperone
YJL051W	0.741940	*IRC8*	Bud tip localized protein of unknown function
YIL085C	0.739010	*KTR7*	Putative mannosyltransferase involved in protein glycosylation
YLL007C	1.025227	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility)
YOR142W	0.744547	*LSC1*	Alpha subunit of succinyl-CoA ligase; succinyl-CoA ligase is a mitochondrial enzyme of the TCA cycle that catalyzes the nucleotide-dependent conversion of succinyl-CoA to succinate
YGL154C	1.076493	*LYS5*	Phosphopantetheinyl transferase involved in lysine biosynthesis
YER001W	1.176455	*MNN1*	Alpha-1,3-mannosyltransferase
YPL013C	1.184731	*MRPS16*	Mitochondrial ribosomal protein of the small subunit
YDL027C	1.466174	*MRX9*	Protein that associates with mitochondrial ribosome

2,3-DMF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YJL116C	0.836433	*NCA3*	Protein involved in mitochondrion organization
YKL151C	1.209968	*NNR2*	Widely-conserved NADHX dehydratase; converts (S)-NADHX to NADH in ATP-dependent manner
YHR195W	0.876481	*NVJ1*	Nuclear envelope protein
YGL094C	0.734275	*PAN2*	Catalytic subunit of the Pan2p-Pan3p poly(A)-ribonuclease complex
YIL050W	0.918297	*PCL7*	Pho85p cyclin of the Pho80p subfamily; forms a functional kinase complex with Pho85p which phosphorylates Mmr1p and is regulated by Pho81p
YOL100W	1.136155	*PKH2*	Serine/threonine protein kinase; involved in sphingolipid-mediated signaling pathway that controls endocytosis
YOR161C	0.803839	*PNS1*	Protein of unknown function
YML047C	1.112164	*PRM6*	Potassium transporter that mediates K+ influx; activates high-affinity Ca2+ influx system (HACS) during mating pheromone response
YGL053W	0.900690	*PRM8*	Pheromone-regulated protein; contains with 2 predicted transmembrane segments and an FF sequence, a motif involved in COPII binding
YJL078C	0.738385	*PRY3*	Cell wall-associated protein involved in export of acetylated sterols
YDR257C	1.282331	*RKM4*	Ribosomal lysine methyltransferase
YDR465C	1.111731	*RMT2*	Arginine N5 methyltransferase; methylates ribosomal protein Rpl12 (L12) on Arg67
YCL028W	1.171583	*RNQ1*	[PIN(+)] prion; an infectious protein conformation that is generally an ordered protein aggregate
YIL066C	0.569618	*RNR3*	Minor isoform of large subunit of ribonucleotide-diphosphate reductase
YLR325C	1.310320	*RPL38*	Ribosomal 60S subunit protein L38
YMR074C	0.789708	*SDD2*	Protein with homology to human PDCD5; PDCD5 is involved in programmed cell death
YMR140W	0.904082	*SIP5*	Protein of unknown function
YNL047C	1.097793	*SLM2*	Phosphoinositide PI4,5P(2) binding protein, forms a complex with Slm1p
YNL086W	0.819439	*SNN1*	Subunit of the BLOC-1 complex involved in endosomal maturation
YGL131C	0.965289	*SNT2*	Subunit of Snt2C complex, RING finger ubiquitin ligase (E3)
YKL184W	1.053011	*SPE1*	Ornithine decarboxylase; catalyzes the first step in polyamine biosynthesis

2,3-DMF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YIL073C	0.925329	*SPO22*	Meiosis-specific protein essential for chromosome synapsis
YGR059W	0.413457	*SPR3*	Sporulation-specific homolog of the CDC3/10/11/12 family of genes
YPR151C	0.856022	*SUE1*	Protein required for degradation of unstable forms of cytochrome c
YPR156C	1.186842	*TPO3*	Polyamine transporter of the major facilitator superfamily
YFR010W	1.147452	*UBP6*	Ubiquitin-specific protease; situated in the base subcomplex of the 26S proteasome, releases free ubiquitin from branched polyubiquitin chains en bloc, rather than from the distal tip of the chain
YLR386W	2.001068	*VAC14*	Enzyme regulator; involved in synthesis of phosphatidylinositol 3,5-bisphosphate, in control of trafficking of some proteins to the vacuole lumen via the MVB, and in maintenance of vacuole size and acidity
YCL069W	1.078316	*VBA3*	Permease of basic amino acids in the vacuolar membrane
YER128W	0.841194	*VFA1*	Protein that interacts with Vps4p and has a role in vacuolar sorting
YOR083W	0.730495	*WHI5*	Repressor of G1 transcription; binds to SCB binding factor (SBF) at SCB target promoters in early G1
YDR259C	1.072535	*YAP6*	Basic leucine zipper (bZIP) transcription factor; physically interacts with the Tup1-Cyc8 complex and recruits Tup1p to its targets
YDL072C	1.199146	*YET3*	Protein of unknown function; YET3 null mutant decreases the level of secreted invertase
YGR054W	0.576241	-	Eukaryotic initiation factor eIF2A; associates specifically with both 40S subunits and 80 S ribosomes, and interacts genetically with both eIF5b and eIF4E
YOR325W	0.697116	-	Dubious open reading frame; unlikely to encode a functional protein, based on available experimental and comparative sequence data
YEL068C	0.771487	-	Protein of unknown function; expressed at both mRNA and protein levels
YDL121C	0.907863	-	Putative protein of unknown function
YDL034W	0.930035	-	Dubious open reading frame
YLL020C	0.960599	-	Dubious open reading frame
YDL023C	0.980155	-	Dubious open reading frame
YIL077C	1.010735	-	Putative protein of unknown function
YKL066W	1.019045	-	Dubious open reading frame
YGL109W	1.120340	-	Dubious open reading frame
YCR022C	1.308430	-	Putative protein of unknown function

2,3-DMF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOR139C	1.349639	-	Dubious open reading frame
YGL081W	1.381750	-	Putative protein of unknown function

2,3-DMF 15G Resistant

2,3-DMF 15G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YCR088W	0.752527	*ABP1*	Actin-binding protein of the cortical actin cytoskeleton
YGR037C	1.129270	*ACB1*	Acyl-CoA-binding protein; transports newly synthesized acyl-CoA esters from fatty acid synthetase (Fas1p-Fas2p) to acyl-CoA-consuming processes
YGL032C	1.121650	*AGA2*	Adhesion subunit of a-agglutinin of a-cells
YPR021C	0.902986	*AGC1*	Mitochondrial amino acid transporter; acts both as a glutamate uniporter and as an aspartate-glutamate exchanger
YHR199C	0.743028	*AIM46*	Protein of unknown function; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies
YJL084C	0.660857	*ALY2*	Alpha arrestin; controls nutrient-mediated intracellular sorting of permease Gap1p
YHR126C	1.301665	*ANS1*	Putative GPI protein; SWAT-GFP and mCherry fusion proteins localize to the vacuole
YPR128C	1.292918	*ANT1*	Peroxisomal adenine nucleotide transporter; involved in beta-oxidation of medium-chain fatty acid; required for peroxisome proliferation
YDR530C	1.205663	*APA2*	Diadenosine 5',5'''-P1,P4-tetraphosphate phosphorylase II
YCR048W	0.841336	*ARE1*	Acyl-CoA:sterol acyltransferase; endoplasmic reticulum enzyme that contributes the major sterol esterification activity in the absence of oxygen
YHL047C	1.354465	*ARN2*	Transporter; member of the ARN family of transporters that specifically recognize siderophore-iron chelates
YHR137W	0.890741	*ARO9*	Aromatic aminotransferase II; catalyzes the first step of tryptophan, phenylalanine, and tyrosine catabolism
YGL180W	1.218934	*ATG1*	Protein serine/threonine kinase; required for vesicle formation in autophagy and the cytoplasm-to-vacuole targeting (Cvt) pathway
YOR152C	0.438283	*ATG40*	Autophagy receptor with a role in endoplasmic reticulum degradation
YLR412W	1.888104	*BER1*	Protein involved in microtubule-related processes

2,3-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR275W	1.178372	*BSC2*	Protein of unknown function; ORF exhibits genomic organization compatible with a translational readthrough-dependent mode of expression
YGR041W	0.447530	*BUD9*	Protein involved in bud-site selection; mutant has increased aneuploidy tolerance
YIL083C	0.930752	*CAB2*	Subunit of the CoA-Synthesizing Protein Complex (CoA-SPC)
YPR013C	1.034299	*CMR3*	Putative zinc finger protein; YPR013C is not an essential gene
YMR244C-A	2.188208	*COA6*	Protein involved in cytochrome c oxidase (Complex IV) assembly; involved in delivery of copper to Complex IV
YIL111W	1.004974	*COX5B*	Subunit Vb of cytochrome c oxidase; cytochrome c oxidase is the terminal member of the mitochondrial inner membrane electron transport chain
YOR303W	0.614698	*CPA1*	Small subunit of carbamoyl phosphate synthetase
YHR109W	0.962441	*CTM1*	Cytochrome c lysine methyltransferase
YER143W	0.958496	*DDI1*	DNA damage-inducible v-SNARE binding protein; role in suppression of protein secretion
YDR403W	1.438155	*DIT1*	Sporulation-specific enzyme required for spore wall maturation
YIR004W	1.062476	*DJP1*	Cytosolic J-domain-containing protein
YDL174C	0.489053	*DLD1*	Major mitochondrial D-lactate dehydrogenase
YKL213C	0.833262	*DOA1*	WD repeat protein required for ubiquitin-mediated protein degradation; ubiquitin binding cofactor that complexes with Cdc48p; required for ribophagy
YDR446W	1.195093	*ECM11*	Meiosis-specific protein; component of the Synaptonemal Complex (SC) along with Gmc2p
YIL064W	1.115680	*EFM4*	Lysine methyltransferase
YNL024C	1.086366	*EFM6*	Putative S-adenosylmethionine-dependent lysine methyltransferase
YLR206W	1.118795	*ENT2*	Epsin-like protein required for endocytosis and actin patch assembly
YMR058W	1.018959	*FET3*	Ferro-O2-oxidoreductase; multicopper oxidase that oxidizes ferrous (Fe2+) to ferric iron (Fe3+) for subsequent cellular uptake by transmembrane permease Ftr1p
YDR519W	1.587633	*FPR2*	Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase)
YER145C	0.952464	*FTR1*	High affinity iron permease; involved in the transport of iron across the plasma membrane

2,3-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR506C	1.121662	*GMC1*	Protein involved in meiotic progression
YER020W	1.394784	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YPL189W	1.168801	*GUP2*	Probable membrane protein
YKL109W	1.310380	*HAP4*	Transcription factor; subunit of the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex
YML075C	0.908144	*HMG1*	HMG-CoA reductase; catalyzes conversion of HMG-CoA to mevalonate, which is a rate-limiting step in sterol biosynthesis
YOL155C	0.637568	*HPF1*	Haze-protective mannoprotein
YIL110W	1.315206	*HPM1*	AdoMet-dependent methyltransferase
YBR072W	0.868595	*HSP26*	Small heat shock protein (sHSP) with chaperone activity
YDR258C	0.798028	*HSP78*	Oligomeric mitochondrial matrix chaperone
YJL051W	0.743465	*IRC8*	Bud tip localized protein of unknown function
YOR155C	0.681009	*ISN1*	Inosine 5'-monophosphate (IMP)-specific 5'-nucleotidase
YOL002C	1.440972	*IZH2*	Plasma membrane receptor for plant antifungal osmotin
YNL227C	2.101595	*JJJ1*	Co-chaperone that stimulates the ATPase activity of Ssa1p
YDR148C	1.069452	*KGD2*	Dihydrolipoyl transsuccinylase
YPL155C	0.615888	*KIP2*	Kinesin-related motor protein involved in mitotic spindle positioning
YKR061W	0.915108	*KTR2*	Mannosyltransferase involved in N-linked protein glycosylation
YLL007C	1.181845	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility
YOR142W	0.815504	*LSC1*	Alpha subunit of succinyl-CoA ligase -dependent conversion of succinyl-CoA to succinate
YIL094C	1.091424	*LYS12*	Homo-isocitrate dehydrogenase; an NAD-linked mitochondrial enzyme
YGL154C	1.119934	*LYS5*	Phosphopantetheinyl transferase involved in lysine biosynthesis
YDL027C	0.977450	*MRX9*	Protein that associates with mitochondrial ribosome
YCR092C	1.203733	*MSH3*	Mismatch repair protein; forms dimers with Msh2p that mediate repair of insertion or deletion mutations and removal of nonhomologous DNA ends
YBR255W	0.884745	*MTC4*	Protein of unknown function
YKL151C	0.993247	*NNR2*	Widely-conserved NADHX dehydratase
YGL151W	1.568784	*NUT1*	Component of the RNA polymerase II mediator complex
YHR195W	0.848331	*NVJ1*	Nuclear envelope protein

2,3-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YPR091C	0.950980	*NVJ2*	Lipid-binding ER protein, enriched at nucleus-vacuolar junctions (NVJ)
YGL094C	0.538516	*PAN2*	Catalytic subunit of the Pan2p-Pan3p poly(A)-ribonuclease complex
YIL071C	0.817978	*PCI8*	Possible shared subunit of Cop9 signalosome (CSN) and eIF3
YGR087C	0.918940	*PDC6*	Minor isoform of pyruvate decarboxylase; decarboxylates pyruvate to acetaldehyde, involved in amino acid catabolism; transcription is glucose- and ethanol-dependent, and is strongly induced during sulfur limitation
YOR161C	0.851129	*PNS1*	Protein of unknown function; has similarity to Torpedo californica tCTL1p, which is postulated to be a choline transporter
YML047C	1.155697	*PRM6*	Potassium transporter that mediates K+ influx
YBL068W	1.350950	*PRS4*	5-phospho-ribosyl-1(alpha)-pyrophosphate synthetase, synthesizes PRPP
YJL078C	0.655987	*PRY3*	Cell wall-associated protein involved in export of acetylated sterols
YNL201C	0.782810	*PSY2*	Subunit of protein phosphatase PP4 complex
YER089C	0.746334	*PTC2*	Type 2C protein phosphatase (PP2C); dephosphorylates Hog1p to limit maximal osmostress induced kinase activity
YDR419W	1.036992	*RAD30*	DNA polymerase eta; involved in translesion synthesis during post-replication repair; catalyzes the synthesis of DNA opposite cyclobutane pyrimidine dimers and other lesions; involved in formation of post-replicative damage-induced genome-wide cohesion; may also have a role in protection against mitochondrial mutagenesis; mutations in human pol eta are responsible for XPV
YDL059C	1.184554	*RAD59*	Protein involved DNA double-strand break repair; repairs breaks in DNA during vegetative growth via recombination and single-strand annealing; anneals complementary single-stranded DNA; forms nuclear foci upon DNA replication stress; required for loading of Rad52p to DSBs; regulates replication fork progression in DNA ligase I-deficient cells; paralog of Rad52p
YOR265W	1.038245	*RBL2*	Protein involved in microtubule morphogenesis
YDR379W	0.921755	*RGA2*	GTPase-activating protein for polarity-establishment protein Cdc42p
YIL057C	0.723975	*RGI2*	Protein of unknown function

2,3-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR257C	0.960625	*RKM4*	Ribosomal lysine methyltransferase
YDR465C	0.802613	*RMT2*	Arginine N5 methyltransferase
YDR279W	1.133031	*RNH202*	Ribonuclease H2 subunit; required for RNase H2 activity
YIL066C	0.484629	*RNR3*	Minor isoform of large subunit of ribonucleotide-diphosphate reductase
YFL034C-A	0.977646	*RPL22B*	Ribosomal 60S subunit protein L22A
YDL136W	0.632150	*RPL35B*	Ribosomal 60S subunit protein L35B; homologous to mammalian ribosomal protein L35
YLR325C	1.260014	*RPL38*	Ribosomal 60S subunit protein L38; homologous to mammalian ribosomal protein L38
YBL066C	0.976522	*SEF1*	Putative transcription factor; has homolog in Kluyveromyces lactis
YMR140W	0.720635	*SIP5*	Protein of unknown function
YEL065W	0.719452	*SIT1*	Ferrioxamine B transporter; member of the ARN family of transporters that specifically recognize siderophore-iron chelates
YNL047C	0.948829	*SLM2*	Phosphoinositide PI4,5P(2) binding protein, forms a complex with Slm1p
YHR030C	0.980646	*SLT2*	Serine/threonine MAP kinase; coordinates expression of all 19S regulatory particle assembly-chaperones (RACs) to control proteasome abundance
YNL086W	1.410124	*SNN1*	Subunit of the BLOC-1 complex involved in endosomal maturation
YGL131C	1.110632	*SNT2*	Subunit of Snt2C complex, RING finger ubiquitin ligase (E3)
YIL073C	0.753199	*SPO22*	Meiosis-specific protein essential for chromosome synapsis
YGR059W	0.347663	*SPR3*	Sporulation-specific homolog of the CDC3/10/11/12 family of genes
YPR151C	0.829205	*SUE1*	Protein required for degradation of unstable forms of cytochrome c
YDR297W	0.886634	*SUR2*	Sphinganine C4-hydroxylase; catalyses the conversion of sphinganine to phytosphingosine in sphingolipid biosyntheis
YIL137C	0.729602	*TMA108*	Ribosome-associated, nascent chain binding factor; binds N-terminal region of nascent peptides during translation
YJL093C	0.593966	*TOK1*	Outward-rectifier potassium channel of the plasma membrane
YGL096W	0.808417	*TOS8*	Homeodomain-containing protein and putative transcription factor; found associated with chromatin
YKL166C	1.375181	*TPK3*	cAMP-dependent protein kinase catalytic subunit

2,3-DMF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR092W	0.864283	*UBC13*	E2 ubiquitin-conjugating enzyme; involved in the error-free DNA postreplication repair pathway
YLL039C	0.930752	*UBI4*	Ubiquitin; becomes conjugated to proteins, marking them for selective degradation via the ubiquitin-26S proteasome system; essential for the cellular stress response
YDL091C	0.882922	*UBX3*	Clathrin-coated vesicle component, regulator of endocytosis; copurifies with the DSC ubiquitin ligase complex
YIL017C	1.099921	*VID28*	GID Complex subunit, serves as adaptor for regulatory subunit Vid24p
YLR410W	0.971720	*VIP1*	Inositol hexakisphosphate and inositol heptakisphosphate kinase
YDL072C	0.825793	*YET3*	Protein of unknown function
YDR319C	1.111615	*YFT2*	Protein required for normal ER membrane biosynthesis
YKR014C	1.065498	*YPT52*	Endosomal Rab family GTPase; required for vacuolar protein sorting, endocytosis and multivesicular body (MVB) biogenesis and sorting
YNL058C	0.657045	-	Putative protein of unknown function
YPR064W	0.685975	-	Putative protein of unknown function
YDL034W	0.757243	-	Dubious open reading frame
YBR225W	0.793997	-	Putative protein of unknown function
YDL121C	0.807392	-	Putative protein of unknown function
YGR291C	0.810272	-	Dubious open reading frame
YDR467C	0.945695	-	Dubious open reading frame
YGR054W	0.950759	-	Eukaryotic initiation factor eIF2A
YHR210C	0.983847	-	Putative aldose 1-epimerase superfamily protein
YPR117W	1.008556	-	Putative protein of unknown function
YDL023C	1.053316	-	Dubious open reading frame
YLL020C	1.059517	-	Dubious open reading frame
YGL140C	1.127625	-	Putative protein of unknown function
YCR102C	1.171009	-	Putative protein of unknown function
YCR101C	1.215706	-	Putative protein of unknown function
YGL081W	1.249041	-	Putative protein of unknown function
YCR022C	1.298300	-	Putative protein of unknown function
YDL086W	1.352890	-	Putative carboxymethylenebutenolidase
YOR139C	1.366406	-	Dubious open reading frame
YPR114W	1.401512	-	Putative protein of unknown function
YDL062W	1.796456	-	Dubious open reading frame
YKR078W	3.604044	-	Cytoplasmic protein of unknown function

2,3- DMF Resistant across all time points

2,3-DMF Resistant Strains Across All Time Points		
Deleted ORF Name	**Deleted Gene Name**	**Deleted Gene Function**
YGR037C	*ACB1*	Acyl-CoA-binding protein
YDR275W	*BSC2*	Protein of unknown function
YPR013C	*CMR3*	Putative zinc finger protein
YIR004W	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YKL213C	*DOA1*	WD repeat protein required for ubiquitin-mediated protein degradation
YDR519W	*FPR2*	Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase)
YER020W	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YLL007C	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility)
YHR195W	*NVJ1*	Nuclear envelope protein
YJL078C	*PRY3*	Cell wall-associated protein involved in export of acetylated sterols
YLR325C	*RPL38*	Ribosomal 60S subunit protein L38
YMR140W	*SIP5*	Protein of unknown function
YNL086W	*SNN1*	Subunit of the BLOC-1 complex involved in endosomal maturation
YDL034W	-	Dubious ORF; unlikely to encode a functional protein
YLL020C	-	Dubious ORF; unlikely to encode a functional protein
YCR022C	-	Putative protein of unknown function

2,3-DMF 5G Sensitive

2,3-DMF 5G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YMR303C	-0.971537	*ADH2*	Glucose-repressible alcohol dehydrogenase II
YBR286W	-0.833081	*APE3*	Vacuolar aminopeptidase Y; processed to mature form by Prb1p
YNL077W	-1.535892	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family; conserved across eukaryotes
YBL099W	-1.179454	*ATP1*	Alpha subunit of the F1 sector of mitochondrial F1F0 ATP synthase; which is a large, evolutionarily conserved enzyme complex required for ATP synthesis
YLR015W	-0.950723	*BRE2*	Subunit of COMPASS (Set1C) complex; COMPASS methylates Lys4 of histone H3 and functions in silencing at telomeres

2,3-DMF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YNR051C	-1.278724	*BRE5*	Ubiquitin protease cofactor; forms deubiquitination complex with Ubp3p that coregulates anterograde and retrograde transport between the endoplasmic reticulum and Golgi compartments
YOR125C	-1.522010	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YNL051W	-2.910711	*COG5*	Component of the conserved oligomeric Golgi complex
YNL041C	-2.929606	*COG6*	Component of the conserved oligomeric Golgi complex
YJL005W	-1.599749	*CYR1*	Adenylate cyclase; required for cAMP production and cAMP-dependent protein kinase signaling
YGR092W	-2.806373	*DBF2*	Ser/Thr kinase involved in transcription and stress response; functions as part of a network of genes in exit from mitosis
YDL219W	-0.612812	*DTD1*	D-Tyr-tRNA(Tyr) deacylase; functions in protein translation, may affect nonsense suppression via alteration of the protein synthesis machinery; ubiquitous among eukaryotes
YMR299C	-1.539484	*DYN3*	Dynein light intermediate chain (LIC); localizes with dynein, null mutant is defective in nuclear migration
YLR047C	-0.835818	*FRE8*	Protein with sequence similarity to iron/copper reductases; involved in iron homeostasis; deletion mutant has iron deficiency/accumulation growth defects
YGL020C	-1.961348	*GET1*	Subunit of the GET complex; involved in insertion of proteins into the ER membrane; required for the retrieval of HDEL proteins from the Golgi to the ER in an ERD2 dependent fashion and for normal mitochondrial morphology and inheritance
YER083C	-2.765680	*GET2*	Subunit of the GET complex; involved in insertion of proteins into the ER membrane; required for the retrieval of HDEL proteins from the Golgi to the ER in an ERD2 dependent fashion and for meiotic nuclear division
YIR037W	-1.020394	*HYR1*	Thiol peroxidase; functions as a hydroperoxide receptor to sense intracellular hydroperoxide levels and transduce a redox signal to the Yap1p transcription factor
YIR024C	-0.720088	*INA22*	F1F0 ATP synthase peripheral stalk assembly factor
YKR019C	-3.971788	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YJR097W	-1.419853	*JJJ3*	Protein of unknown function; contains a CSL Zn finger and a DnaJ-domain; involved in diphthamide biosynthesis; ortholog human Dph4

2,3-DMF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR005C	-1.495639	*MAF1*	Highly conserved negative regulator of RNA polymerase III
YEL053C	-0.874518	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YPR051W	-1.355698	*MAK3*	Catalytic subunit of the NatC type N-terminal acetyltransferase (NAT)
YGL219C	-1.316541	*MDM34*	Mitochondrial component of the ERMES complex; links the ER to mitochondria
YIR033W	-1.962927	*MGA2*	ER membrane protein involved in regulation of OLE1 transcription
YJR074W	-2.376334	*MOG1*	Conserved nuclear protein that interacts with GTP-Gsp1p
YGR028W	-0.863340	*MSP1*	Highly-conserved N-terminally anchored AAA-ATPase
YGR089W	-0.708472	*NNF2*	Protein that exhibits physical and genetic interactions with Rpb8p
YNL099C	-0.936782	*OCA1*	Putative protein tyrosine phosphatase; required for cell cycle arrest in response to oxidative damage of DNA
YOR017W	-0.729197	*PET127*	Protein with a role in 5'-end processing of mitochondrial RNAs; located in the mitochondrial membrane
YNL329C	-1.122094	*PEX6*	AAA-peroxin
YJR153W	-0.720847	*PGU1*	Endo-polygalacturonase; pectolytic enzyme that hydrolyzes the alpha-1,4-glycosidic bonds in the rhamnogalacturonan chains in pectins
YFR034C	-1.054225	*PHO4*	Basic helix-loop-helix (bHLH) transcription factor of the myc-family; activates transcription cooperatively with Pho2p in response to phosphate limitation
YML123C	-1.161148	*PHO84*	High-affinity inorganic phosphate (Pi) transporter
YNL098C	-1.387257	*RAS2*	GTP-binding protein; regulates nitrogen starvation response, sporulation, and filamentous growth
YDR202C	-1.119030	*RAV2*	Subunit of RAVE complex (Rav1p, Rav2p, Skp1p)
YMR283C	-0.904050	*RIT1*	Initiator methionine 2'-O-ribosyl phosphate transferase
YDL020C	-1.968369	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YDL216C	-0.708152	*RRI1*	Catalytic subunit of the COP9 signalosome (CSN) complex
YMR263W	-1.230442	*SAP30*	Component of Rpd3L histone deacetylase complex
YDR077W	-0.943983	*SED1*	Major stress-induced structural GPI-cell wall glycoprotein
YLR079W	-2.719011	*SIC1*	Cyclin-dependent kinase inhibitor (CKI)
YHR206W	-1.145970	*SKN7*	Nuclear response regulator and transcription factor
YLR055C	-1.442292	*SPT8*	Subunit of the SAGA transcriptional regulatory complex

2,3-DMF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR463W	-2.473489	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YJL004C	-3.121976	*SYS1*	Integral membrane protein of the Golgi; required for targeting of the Arf-like GTPase Arl3p to the Golgi
YPR074C	-1.640505	*TKL1*	Transketolase
YOL018C	-5.186796	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi
YDR108W	-2.457089	*TRS85*	Component of transport protein particle (TRAPP) complex III
YKR088C	-0.675722	*TVP38*	Integral membrane protein; localized to late Golgi vesicles along with the v-SNARE Tlg2p; required for asymmetric localization of Kar9p during mitosis
YKR098C	-0.965617	*UBP11*	Ubiquitin-specific protease; cleaves ubiquitin from ubiquitinated proteins
YDR484W	-4.556583	*VPS52*	Component of the GARP (Golgi-associated retrograde protein) complex; GARP is required for the recycling of proteins from endosomes to the late Golgi, and for mitosis after DNA damage induced checkpoint arrest
YFR007W	-0.976674	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YGR281W	-1.155185	*YOR1*	Plasma membrane ATP-binding cassette (ABC) transporter
YML122C	-2.689283	-	Putative protein of unknown function
YJL120W	-1.869174	-	Dubious open reading frame
YNR042W	-1.463566	-	Dubious open reading frame
YPL150W	-1.427868	-	Protein kinase of unknown cellular role
YDR048C	-1.414127	-	Dubious open reading frame
YOR008C-A	-1.162266	-	Putative protein of unknown function
YOL075C	-1.042813	-	Putative ABC transporter
YGL042C	-0.997619	-	Dubious open reading frame
YDR338C	-0.981416	-	Putative protein of unknown function
YBR292C	-0.963285	-	Dubious open reading frame
YGL214W	-0.877412	-	Dubious open reading frame
YDR102C	-0.796059	-	Putative protein of unknown function
YNL122C	-0.767971	-	Mitochondrial ribosomal protein of the large subunit
YIR016W	-0.615145	-	Putative protein of unknown function

2,3-DMF 10G Sensitive

2,3-DMF 10G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YGL071W	-2.735213	*AFT1*	Transcription factor involved in iron utilization and homeostasis
YBR286W	-1.288483	*APE3*	Vacuolar aminopeptidase Y; processed to mature form by Prb1p
YNL077W	-1.491120	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family; conserved across eukaryotes forms nuclear foci upon DNA replication stress
YNL315C	-2.222232	*ATP11*	Molecular chaperone; required for the assembly of alpha and beta subunits into the F1 sector of mitochondrial F1F0 ATP synthase; N-terminally propionylated in vivo
YGR286C	-0.943762	*BIO2*	Biotin synthase; catalyzes the conversion of dethiobiotin to biotin, which is the last step of the biotin biosynthesis pathway; complements E. coli bioB mutant
YNR051C	-1.919521	*BRE5*	Ubiquitin protease cofactor
YOR125C	-1.290148	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YKL208W	-1.257861	*CBT1*	Protein involved in 5' RNA end processing
YOR039W	-2.267557	*CKB2*	Beta' regulatory subunit of casein kinase 2 (CK2)
YNL051W	-6.892976	*COG5*	Component of the conserved oligomeric Golgi complex
YNL041C	-7.300942	*COG6*	Component of the conserved oligomeric Golgi complex
YGL005C	-7.321791	*COG7*	Component of the conserved oligomeric Golgi complex
YML071C	-7.832654	*COG8*	Component of the conserved oligomeric Golgi complex
YCR086W	-1.824835	*CSM1*	Nucleolar protein that mediates homolog segregation during meiosis I
YBR291C	-1.484676	*CTP1*	Mitochondrial inner membrane citrate transporter
YJL005W	-1.799016	*CYR1*	Adenylate cyclase; required for cAMP production and cAMP-dependent protein kinase signaling
YGR092W	-4.787526	*DBF2*	Ser/Thr kinase involved in transcription and stress response
YDL160C	-2.337756	*DHH1*	Cytoplasmic DEAD-box helicase, stimulates mRNA decapping
YBR278W	-1.582657	*DPB3*	Third-largest subunit of DNA polymerase II (DNA polymerase epsilon); required to maintain fidelity of chromosomal replication and also for inheritance of telomeric silencing
YBR281C	-1.267645	*DUG2*	Component of glutamine amidotransferase (GATase II)
YMR299C	-1.337928	*DYN3*	Dynein light intermediate chain (LIC); localizes with dynein, null mutant is defective in nuclear migration

2,3-DMF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGL020C	-6.358921	*GET1*	Subunit of the GET complex; involved in insertion of proteins into the ER membrane
YER083C	-5.764015	*GET2*	Subunit of the GET complex; involved in insertion of proteins into the ER membrane
YDL100C	-1.944567	*GET3*	Guanine nucleotide exchange factor for Gpa1p; amplifies G protein signaling; functions as a chaperone under ATP-depleted oxidative stress conditions; subunit of GET complex, involved in ATP dependent Golgi to ER trafficking and insertion of tail-anchored (TA) proteins into ER membrane under non-stress conditions
YOR070C	-2.382902	*GYP1*	Cis-golgi GTPase-activating protein (GAP) for yeast Rabs
YNL281W	-0.790311	*HCH1*	Heat shock protein regulator; binds to Hsp90p and may stimulate ATPase activity
YKR019C	-7.686402	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YJR097W	-1.485647	*JJJ3*	Protein of unknown function
YAL024C	-4.607979	*LTE1*	Protein similar to GDP/GTP exchange factors
YEL053C	-1.821274	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase; required for replication of dsRNA virus; expression is glucose-repressible
YPR051W	-2.271533	*MAK3*	Catalytic subunit of the NatC type N-terminal acetyltransferase (NAT
YOL027C	-1.317384	*MDM38*	Mitochondrial protein; forms a complex with Mba1p to facilitate recruitment of mRNA-specific translational activators to ribosomes
YML062C	-1.648295	*MFT1*	Subunit of the THO complex
YIR033W	-4.325982	*MGA2*	ER membrane protein involved in regulation of OLE1 transcription
YMR115W	-1.403710	*MGR3*	Subunit of the mitochondrial (mt) i-AAA protease supercomplex
YFR011C	-0.784141	*MIC19*	Component of the MICOS complex
YDR347W	-3.109173	*MRP1*	Mitochondrial ribosomal protein of the small subunit
YNL005C	-1.453042	*MRP7*	Mitochondrial ribosomal protein of the large subunit
YMR193W	-4.393036	*MRPL24*	Mitochondrial ribosomal protein of the large subunit
YML009C	-1.212511	*MRPL39*	Mitochondrial ribosomal protein of the large subunit
YPL174C	-1.430655	*NIP100*	Large subunit of the dynactin complex
YNL056W	-1.185107	*OCA2*	Protein of unknown function
YJR073C	-1.439331	*OPI3*	Methylene-fatty-acyl-phospholipid synthase
YFR034C	-1.376279	*PHO4*	Basic helix-loop-helix (bHLH) transcription factor of the myc-family

2,3-DMF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR435C	-1.929866	*PPM1*	Carboxyl methyltransferase; methylates the C terminus of the protein phosphatase 2A catalytic subunit (Pph21p or Pph22p), which is important for complex formation with regulatory subunits
YLR204W	-1.693130	*QRI5*	Mitochondrial inner membrane protein; required for accumulation of spliced COX1 mRNA; may have an additional role in translation of COX1 mRNA
YNL098C	-2.107583	*RAS2*	GTP-binding protein; regulates nitrogen starvation response, sporulation, and filamentous growth
YCR036W	-1.839380	*RBK1*	Putative ribokinase
YMR274C	-0.968692	*RCE1*	Type II CAAX prenyl protease
YJL217W	-0.930888	*REE1*	Cytoplasmic protein involved in the regulation of enolase (ENO1)
YMR283C	-0.818155	*RIT1*	Initiator methionine 2'-O-ribosyl phosphate transferase
YBR030W	-1.197373	*RKM3*	Ribosomal lysine methyltransferase
YNL248C	-2.805783	*RPA49*	RNA polymerase I subunit A49
YDL133C-A	-1.262419	*RPL41B*	Ribosomal 60S subunit protein L41B
YDL020C	-4.664259	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YIL153W	-1.222704	*RRD1*	Peptidyl-prolyl cis/trans-isomerase; activator of the phosphotyrosyl phosphatase activity of PP2A
YMR263W	-1.358035	*SAP30*	Component of Rpd3L histone deacetylase complex
YLR268W	-7.261507	*SEC22*	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p; cycles between the ER and Golgi complex
YHR206W	-1.783726	*SKN7*	Nuclear response regulator and transcription factor; physically interacts with the Tup1-Cyc8 complex and recruits Tup1p to its targets
YOR327C	-1.382896	*SNC2*	Vesicle membrane receptor protein (v-SNARE); involved in the fusion between Golgi-derived secretory vesicles with the plasma membrane
YLR055C	-1.527220	*SPT8*	Subunit of the SAGA transcriptional regulatory complex
YBR283C	-4.080281	*SSH1*	Subunit of the Ssh1 translocon complex
YMR183C	-1.240148	*SSO2*	Plasma membrane t-SNARE; involved in fusion of secretory vesicles at the plasma membrane
YLR452C	-1.064612	*SST2*	GTPase-activating protein for Gpa1p; regulates desensitization to alpha factor pheromone

2,3-DMF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR463W	-6.338989	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YJL004C	-6.715130	*SYS1*	Integral membrane protein of the Golgi; required for targeting of the Arf-like GTPase Arl3p to the Golgi
YBR069C	-2.211228	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YPR074C	-3.455886	*TKL1*	Transketolase; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway
YOL018C	-8.282081	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi; required along with VPS45 for an early step of the constitutive CVT pathway
YDR120C	-1.915219	*TRM1*	tRNA methyltransferase
YDR108W	-4.368972	*TRS85*	Component of transport protein particle (TRAPP) complex III
YML028W	-1.603424	*TSA1*	Thioredoxin peroxidase; acts as both ribosome-associated and free cytoplasmic antioxidant
YOR006C	-1.761923	*TSR3*	Protein required for 20S pre-rRNA processing
YBR006W	-1.246184	*UGA2*	Succinate semialdehyde dehydrogenase
YDR484W	-6.356575	*VPS52*	Component of the GARP (Golgi-associated retrograde protein) complex
YML007W	-1.583690	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YFR007W	-1.418923	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YGR281W	-1.045704	*YOR1*	Plasma membrane ATP-binding cassette (ABC) transporter
YNL093W	-0.813317	*YPT53*	Stress-induced Rab family GTPase; required for vacuolar protein sorting and endocytosis; involved in ionic stress tolerance
YML122C	-4.851465	-	Putative protein of unknown function
YDR149C	-2.027454	-	Dubious open reading frame
YNL120C	-1.804663	-	Dubious open reading frame
YGR035C	-1.746188	-	Putative protein of unknown function
YPR050C	-1.474772	-	Dubious open reading frame
YCR087W	-1.398220	-	Dubious open reading frame

2,3-DMF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGL214W	-1.379099	-	Dubious open reading frame
YCR050C	-1.311811	-	Non-essential protein of unknown function
YBR292C	-1.265182	-	Dubious open reading frame
YHL044W	-1.184607	-	Putative integral membrane protein
YOR008C-A	-1.120812	-	Putative protein of unknown function
YNL115C	-1.032246	-	Putative protein of unknown function
YJR026W	-0.980077	-	Putative protein of unknown function
YGL042C	-0.979831	-	Dubious open reading frame
YIR016W	-0.752606	-	Putative protein of unknown function

2,3-DMF 15G Sensitive

2,3-DMF 15G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YLR131C	-5.474887	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YMR120C	-1.029601	*ADE17*	Enzyme of 'de novo' purine biosynthesis
YMR282C	-3.615445	*AEP2*	Mitochondrial protein; likely involved in translation of the mitochondrial OLI1 mRNA
YER017C	-4.115650	*AFG3*	Mitochondrial inner membrane m-AAA protease component
YHR093W	-1.549456	*AHT1*	Dubious open reading frame
YOR067C	-1.070793	*ALG8*	Glucosyl transferase; involved in N-linked glycosylation
YBR286W	-0.946290	*APE3*	Vacuolar aminopeptidase Y; processed to mature form by Prb1p
YNL077W	-1.237276	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family; conserved across eukaryotes
YBR288C	-1.399158	*APM3*	Mu3-like subunit of the clathrin associated protein complex (AP-3); functions in transport of alkaline phosphatase to the vacuole via the alternate pathway
YJL024C	-1.447442	*APS3*	Small subunit of the clathrin-associated adaptor complex AP-3; involved in vacuolar protein sorting
YDL192W	-4.828171	*ARF1*	ADP-ribosylation factor; GTPase of the Ras superfamily involved in regulation of coated vesicle formation in intracellular trafficking within the Golgi

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YCR068W	-2.132727	*ATG15*	Phospholipase; preferentially hydrolyses phosphatidylserine, with minor activity against cardiolipin and phosphatidylethanolamine
YIL088C	-0.991925	*AVT7*	Putative transporter; member of a family of seven S. cerevisiae genes (AVT1-7) related to vesicular GABA-glycine transporters
YJL020C	-0.853276	*BBC1*	Protein possibly involved in assembly of actin patches
YGR286C	-0.819173	*BIO2*	Biotin synthase; catalyzes the conversion of dethiobiotin to biotin, which is the last step of the biotin biosynthesis pathway; complements E. coli bioB mutant
YLR015W	-1.344640	*BRE2*	Subunit of COMPASS (Set1C) complex; COMPASS methylates Lys4 of histone H3 and functions in silencing at telomeres
YNR051C	-2.048330	*BRE5*	Ubiquitin protease cofactor; forms deubiquitination complex with Ubp3p that coregulates anterograde and retrograde transport between the endoplasmic reticulum and Golgi compartments
YFL025C	-4.079415	*BST1*	GPI inositol deacylase of the endoplasmic reticulum (ER); negatively regulates COPII vesicle formation; prevents production of vesicles with defective subunits
YLR319C	-1.437797	*BUD6*	Actin- and formin-interacting protein; participates in actin cable assembly and organization as a nucleation-promoting factor (NPF) for formins Bni1p and Bnr1p
YMR275C	-1.676364	*BUL1*	Ubiquitin-binding component of the Rsp5p E3-ubiquitin ligase complex
YOR125C	-1.495594	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YCR005C	-1.302360	*CIT2*	Citrate synthase, peroxisomal isozyme involved in glyoxylate cycle; catalyzes condensation of acetyl coenzyme A and oxaloacetate to form citrate
YNR041C	-1.283001	*COQ2*	Para hydroxybenzoate polyprenyl transferase; catalyzes the second step in ubiquinone (coenzyme Q) biosynthesis
YLR087C	-2.463395	*CSF1*	Protein required for fermentation at low temperature; plays a role in the maturation of secretory proteins
YDR179C	-1.040380	*CSN9*	Subunit of the Cop9 signalosome
YBR291C	-1.972605	*CTP1*	Mitochondrial inner membrane citrate transporter
YLR286C	-0.892555	*CTS1*	Endochitinase; required for cell separation after mitosis
YJL005W	-1.908016	*CYR1*	Adenylate cyclase; required for cAMP production and cAMP-dependent protein kinase signaling
YML113W	-1.116362	*DAT1*	DNA binding protein that recognizes oligo(dA).oligo(dT) tracts

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGR092W	-6.868667	*DBF2*	Ser/Thr kinase involved in transcription and stress response
YPL265W	-2.073181	*DIP5*	Dicarboxylic amino acid permease; mediates high-affinity and high-capacity transport of L-glutamate and L-aspartate
YBR278W	-1.793029	*DPB3*	Third-largest subunit of DNA polymerase II (DNA polymerase epsilon); required to maintain fidelity of chromosomal replication and also for inheritance of telomeric silencing
YBR281C	-1.253103	*DUG2*	Component of glutamine amidotransferase (GATase II)
YMR299C	-1.428060	*DYN3*	Dynein light intermediate chain (LIC)
YCR034W	-1.130975	*ELO2*	Fatty acid elongase, involved in sphingolipid biosynthesis; acts on fatty acids of up to 24 carbons in length
YFL048C	-0.986428	*EMP47*	Integral membrane component of ER-derived COPII-coated vesicles; functionS in ER to Golgi transport
YJR125C	-0.846133	*ENT3*	Protein containing an N-terminal epsin-like domain; involved in clathrin recruitment and traffic between the Golgi and endosomes
YMR015C	-1.441653	*ERG5*	C-22 sterol desaturase; a cytochrome P450 enzyme that catalyzes the formation of the C-22(23) double bond in the sterol side chain in ergosterol biosynthesis
YGL054C	-2.335261	*ERV14*	COPII-coated vesicle protein; involved in vesicle formation and incorporation of specific secretory cargo
YFR008W	-1.780069	*FAR7*	Protein involved in recovery from pheromone-induced cell cycle arrest
YPL248C	-1.297858	*GAL4*	DNA-binding transcription factor required for activating GAL genes
YGL020C	-7.201491	*GET1*	Subunit of the GET complex; insertion of proteins into the ER membrane
YER083C	-5.725994	*GET2*	Subunit of the GET complex; insertion of proteins into the ER membrane
YDL100C	-1.974779	*GET3*	Guanine nucleotide exchange factor for Gpa1p; amplifies G protein signaling
YMR135C	-1.192951	*GID8*	Subunit of GID Complex, binds strongly to central component Vid30p
YOR070C	-2.845722	*GYP1*	Cis-golgi GTPase-activating protein (GAP) for yeast Rabs
YNL281W	-0.884462	*HCH1*	Heat shock protein regulator; binds to Hsp90p and may stimulate ATPase activity
YGR187C	-2.556753	*HGH1*	Nonessential protein of unknown function

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YLR113W	-1.309326	*HOG1*	Mitogen-activated protein kinase involved in osmoregulation; controls global reallocation of RNAPII during osmotic shock
YJR036C	-0.843981	*HUL4*	Protein with similarity to hect domain E3 ubiquitin-protein ligases
YPR006C	-1.169679	*ICL2*	2-methylisocitrate lyase of the mitochondrial matrix
YIR024C	-0.824788	*INA22*	F1F0 ATP synthase peripheral stalk assembly factor
YKR019C	-8.243924	*IRS4*	EH domain-containing protein
YJR097W	-1.752606	*JJJ3*	Protein of unknown function; contains a CSL Zn finger and a DnaJ-domain
YNL322C	-1.176025	*KRE1*	Cell wall glycoprotein involved in beta-glucan assembly
YNL323W	-1.665462	*LEM3*	Membrane protein of the plasma membrane and ER
YPL004C	-1.131886	*LSP1*	Eisosome core component
YEL053C	-2.484753	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YPR051W	-3.336516	*MAK3*	Catalytic subunit of the NatC type N-terminal acetyltransferase (NAT
YBR298C	-1.294671	*MAL31*	Maltose permease; high-affinity maltose transporter (alpha-glucoside transporter)
YBR297W	-1.183519	*MAL33*	MAL-activator protein; part of complex locus MAL3
YNL307C	-1.291941	*MCK1*	Dual-specificity ser/thr and tyrosine protein kinase
YJR137C	-0.746921	*MET5*	Sulfite reductase beta subunit; involved in amino acid biosynthesis, transcription repressed by methionine
YML062C	-1.970496	*MFT1*	Subunit of the THO complex; THO is a nuclear complex comprised of Hpr1p, Mft1p, Rlr1p, and Thp2p, that is involved in transcription elongation and mitotic recombination; involved in telomere maintenance
YIR033W	-5.567892	*MGA2*	ER membrane protein involved in regulation of OLE1 transcription
YFR011C	-0.635537	*MIC19*	Component of the MICOS complex
YNL005C	-1.337731	*MRP7*	Mitochondrial ribosomal protein of the large subunit
YML009C	-0.886252	*MRPL39*	Mitochondrial ribosomal protein of the large subunit
YOR066W	-1.007462	*MSA1*	Activator of G1-specific transcription factors MBF and SBF
YGR028W	-0.939018	*MSP1*	Highly-conserved N-terminally anchored AAA-ATPase
YKR048C	-2.105289	*NAP1*	Histone chaperone; involved in histone exchange by removing and replacing histone H2A-H2B dimers or histone variant dimers from assembled nucleosomes
YMR285C	-0.762354	*NGL2*	Protein involved in 5.8S rRNA processing; Ccr4p-like RNase required for correct 3'-end formation of 5.8S rRNA at site E

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YNR009W	-0.800896	*NRM1*	Transcriptional co-repressor of MBF-regulated gene expression
YBL079W	-2.011578	*NUP170*	Subunit of inner ring of nuclear pore complex (NPC);
YNL099C	-0.869513	*OCA1*	Putative protein tyrosine phosphatase; required for cell cycle arrest in response to oxidative damage of DNA
YNL056W	-1.071552	*OCA2*	Protein of unknown function
YJR073C	-2.937260	*OPI3*	Methylene-fatty-acyl-phospholipid synthase
YPL272C	-1.685850	*PBI1*	Putative protein of unknown function
YJL128C	-1.080331	*PBS2*	MAP kinase kinase of the HOG signaling pathway
YJL053W	-2.273779	*PEP8*	Vacuolar protein component of the retromer
YLR064W	-1.092421	*PER33*	Protein that localizes to the endoplasmic reticulum
YOR017W	-0.953320	*PET127*	Protein with a role in 5'-end processing of mitochondrial RNAs
YDL106C	-6.411557	*PHO2*	Homeobox transcription factor; regulatory targets include genes involved in phosphate metabolism
YNL097C	-2.897726	*PHO23*	Component of the Rpd3L histone deacetylase complex
YCR024C-A	-1.405870	*PMP1*	Regulatory subunit for the plasma membrane H(+)-ATPase Pma1p
YMR278W	-0.755640	*PRM15*	Phosphoribomutase; catalyzes interconversion of ribose-1-phosphate and ribose-5-phosphate
YCR079W	-0.948109	*PTC6*	Mitochondrial type 2C protein phosphatase (PP2C)
YDR004W	-1.502502	*RAD57*	Protein that stimulates strand exchange; stimulates strand exchange by stabilizing the binding of Rad51p to single-stranded DNA; involved in the recombinational repair of double-strand breaks in DNA during vegetative growth and meiosis
YNL098C	-2.202599	*RAS2*	GTP-binding protein; regulates nitrogen starvation response, sporulation, and filamentous growth
YCR036W	-1.358981	*RBK1*	Putative ribokinase
YMR274C	-1.918059	*RCE1*	Type II CAAX prenyl protease
YDR028C	-4.667586	*REG1*	Regulatory subunit of type 1 protein phosphatase Glc7p
YLR059C	-0.676246	*REX2*	3'-5' RNA exonuclease; involved in 3'-end processing of U4 and U5 snRNAs, 5S and 5.8S rRNAs, and RNase P and RNase MRP RNA
YOL080C	-0.965479	*REX4*	Putative RNA exonuclease
YMR283C	-1.088724	*RIT1*	Initiator methionine 2'-O-ribosyl phosphate transferase
YDL020C	-7.652667	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YGR215W	-5.946229	*RSM27*	Mitochondrial ribosomal protein of the small subunit
YOR216C	-1.899614	*RUD3*	Golgi matrix protein; involved in the structural organization of the cis-Golgi

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YMR263W	-1.762057	*SAP30*	Component of Rpd3L histone deacetylase complex
YLR268W	-5.598313	*SEC22*	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p
YDR077W	-0.894212	*SED1*	Major stress-induced structural GPI-cell wall glycoprotein
YOR021C	-0.741516	*SFM1*	SPOUT methyltransferase; catalyzes omega-monomethylation of Rps3p on Arg-146; not an essential gene
YDR078C	-1.156823	*SHU2*	Component of Shu complex (aka PCSS complex
YNL032W	-1.404055	*SIW14*	Tyrosine phosphatase involved in actin organization and endocytosis
YKR100C	-1.007535	*SKG1*	Transmembrane protein with a role in cell wall polymer composition
YHR206W	-2.048764	*SKN7*	Nuclear response regulator and transcription factor
YOR327C	-1.752104	*SNC2*	Vesicle membrane receptor protein (v-SNARE
YMR107W	-1.230816	*SPG4*	Protein required for high temperature survival during stationary phase
YDR218C	-0.859372	*SPR28*	Sporulation-specific homolog of the CDC3/10/11/12 family of genes
YBR283C	-6.271840	*SSH1*	Subunit of the Ssh1 translocon complex
YMR183C	-2.156262	*SSO2*	Plasma membrane t-SNARE; involved in fusion of secretory vesicles at the plasma membrane
YJL004C	-5.691171	*SYS1*	Integral membrane protein of the Golgi
YBR069C	-2.550501	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine; low-affinity tryptophan and histidine transporter
YEL048C	-0.843165	*TCA17*	Component of transport protein particle (TRAPP) complex II
YJR116W	-0.778804	*TDA4*	Putative protein of unknown function
YOL018C	-7.793015	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi
YNL273W	-1.512969	*TOF1*	Subunit of a replication-pausing checkpoint complex
YNL300W	-1.015088	*TOS6*	Glycosylphosphatidylinositol-dependent cell wall protein
YDR120C	-2.246689	*TRM1*	tRNA methyltransferase; two forms of protein are made by alternative translation starts
YGR166W	-1.243164	*TRS65*	Component of transport protein particle (TRAPP) complex II
YDR108W	-4.886154	*TRS85*	Component of transport protein particle (TRAPP) complex III

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDL190C	-0.541939	*UFD2*	Ubiquitin chain assembly factor (E4); cooperates with a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin protein ligase (E3) to conjugate ubiquitin to substrates; also functions as an E3
YOR068C	-0.953911	*VAM10*	Protein involved in vacuole morphogenesis; acts at an early step of homotypic vacuole fusion that is required for vacuole tethering
YJL154C	-1.717544	*VPS35*	Endosomal subunit of membrane-associated retromer complex
YNL283C	-3.128723	*WSC2*	Sensor-transducer of the stress-activated PKC1-MPK1 signaling pathway
YJR133W	-0.855397	*XPT1*	Xanthine-guanine phosphoribosyl transferase; required for xanthine utilization and for optimal utilization of guanine
YML007W	-2.151902	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YHL009C	-1.151434	*YAP3*	Basic leucine zipper (bZIP) transcription factor
YLR020C	-1.250598	*YEH2*	Steryl ester hydrolase; catalyzes steryl ester hydrolysis at the plasma membrane
YFR007W	-0.975740	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YCR059C	-1.304332	*YIH1*	Negative regulator of eIF2 kinase Gcn2p
YGR281W	-1.077599	*YOR1*	Plasma membrane ATP-binding cassette (ABC) transporter
YML122C	-8.102209	-	Putative protein of unknown function
YPR050C	-2.026034	-	Dubious open reading frame
YCR050C	-1.983362	-	Non-essential protein of unknown function
YCL001W-A	-1.974990	-	Putative protein of unknown function
YCR061W	-1.899767	-	Protein of unknown function
YOR008C-A	-1.780504	-	Putative protein of unknown function
YJL193W	-1.753162	-	Putative protein of unknown function
YKL023W	-1.744617	-	Putative protein of unknown function
YDR149C	-1.718523	-	Dubious open reading frame
YCR085W	-1.547682	-	Putative protein of unknown function
YHR079C-B	-1.342525	-	Dubious open reading frame
YPR053C	-1.289874	-	Putative protein of unknown function
YJR142W	-1.286881	*YJR142W*	8-oxo-dGTP diphosphatase of the Nudix hydrolase family

2,3-DMF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOL075C	-1.230936	-	Putative ABC transporter
YGL042C	-1.196844	-	Dubious open reading frame
YAL058C-A	-1.184683	-	Dubious open reading frame
YGL214W	-1.135707	-	Dubious open reading frame
YBR292C	-1.078380	-	Dubious open reading frame
YLR046C	-1.056413	-	Putative membrane protein
YJR026W	-1.031693	-	Transposable element gene
YDL050C	-0.973306	-	Dubious open reading frame
YNL035C	-0.927139	-	Nuclear protein of unknown function
YNR021W	-0.899226	-	Putative protein of unknown function
YBR284W	-0.888252	-	Putative metallo-dependent hydrolase superfamily protein
YGR018C	-0.828899	-	Protein of unknown function
YLR112W	-0.788112	-	Dubious open reading frame
YGR022C	-0.761221	-	Dubious open reading frame
YDR102C	-0.754136	-	Putative protein of unknown function
YMR153C-A	-0.726078	-	Dubious open reading frame
YIR016W	-0.665132	-	Putative protein of unknown function

2,3-DMF Sensitive across all time points

2,3-DMF Sensitive Strains Across All Time Points		
Deleted ORF Name	**Deleted Gene Name**	**Deleted Gene Function**
YBR286W	*APE3*	Vacuolar aminopeptidase Y
YNL077W	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation
YNR051C	*BRE5*	Ubiquitin protease cofactor
YOR125C	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YJL005W	*CYR1*	Adenylate cyclase
YGR092W	*DBF2*	Ser/Thr kinase involved in transcription and stress response
YMR299C	*DYN3*	Dynein light intermediate chain (LIC)
YGL020C	*GET1*	Subunit of the GET complex
YER083C	*GET2*	Subunit of the GET complex
YKR019C	*IRS4*	EH domain-containing protein
YJR097W	*JJJ3*	Protein of unknown function
YEL053C	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YPR051W	*MAK3*	Catalytic subunit of the NatC type N-terminal acetyltransferase (NAT
YIR033W	*MGA2*	ER membrane protein involved in regulation of OLE1 transcription

2,3-DMF Sensitive Strains Across All Time Points (continued)		
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**
YNL098C	*RAS2*	GTP-binding protein
YMR283C	*RIT1*	Initiator methionine 2'-O-ribosyl phosphate transferase
YDL020C	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YMR263W	*SAP30*	Component of Rpd3L histone deacetylase complex
YHR206W	*SKN7*	Nuclear response regulator and transcription factor
YJL004C	*SYS1*	Integral membrane protein of the Golgi
YOL018C	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi
YDR108W	*TRS85*	Component of transport protein particle (TRAPP) complex III
YFR007W	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YGR281W	*YOR1*	Plasma membrane ATP-binding cassette (ABC) transporter
YML122C	-	Putative protein of unknown function
YOR008C-A	-	Putative protein of unknown function
YGL042C	-	Dubious open reading frame
YBR292C	-	Dubious open reading frame
YGL214W	-	Dubious open reading frame
YIR016W	-	Putative protein of unknown function

Appendix 3: Mutants with altered growth in 2-MF

2-MF 5G Resistant

2-MF 5G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YGR037C	0.861899	*ACB1*	Acyl-CoA-binding protein; transports newly synthesized acyl-CoA esters from fatty acid synthetase (Fas1p-Fas2p) to acyl-CoA-consuming processes
YHR126C	1.332337	*ANS1*	Putative GPI protein
YCR048W	0.915686	*ARE1*	Acyl-CoA:sterol acyltransferase; endoplasmic reticulum enzyme
YBR068C	0.758662	*BAP2*	High-affinity leucine permease
YKR027W	1.113822	*BCH2*	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family
YCR032W	1.090285	*BPH1*	Protein homologous to Chediak-Higashi syndrome and Beige proteins
YKR036C	1.287047	*CAF4*	WD40 repeat-containing protein associated with the CCR4-NOT complex
YPR013C	0.776348	*CMR3*	Putative zinc finger protein; YPR013C is not an essential gene
YOR303W	0.632355	*CPA1*	Small subunit of carbamoyl phosphate synthetase
YGL078C	1.354602	*DBP3*	RNA-Dependent ATPase, member of DExD/H-box family
YIR004W	0.966997	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YDR294C	1.006600	*DPL1*	Dihydrosphingosine phosphate lyase
YLR372W	2.122839	*ELO3*	Elongase; involved in fatty acid and sphingolipid biosynthesis
YGL057C	0.788809	*GEP7*	Protein of unknown function
YER020W	0.963924	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YBR244W	0.424950	*GPX2*	Phospholipid hydroperoxide glutathione peroxidase; protects cells from phospholipid hydroperoxides and nonphospholipid peroxides during oxidative stress; induced by glucose starvation; protein abundance increases in response to DNA replication stress
YJL165C	1.095023	*HAL5*	Putative protein kinase
YKL109W	1.263131	*HAP4*	Transcription factor; subunit of the heme-activated, glucose-repressed Hap2p/3p/4p/5p CCAAT-binding complex
YGL033W	0.504466	*HOP2*	Meiosis-specific protein that localizes to chromosomes

2-MF 5G Resistant Strains (continue)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YCR021C	0.936182	*HSP30*	Negative regulator of the H(+)-ATPase Pma1p; stress-responsive protein
YJL082W	0.895780	*IML2*	Protein required for clearance of inclusion bodies; localizes to the inclusion bodies formed under protein misfolding stress
YER019W	1.890192	*ISC1*	Inositol phosphosphingolipid phospholipase C; mitochondrial membrane localized
YLR239C	0.971188	*LIP2*	Lipoyl ligase; involved in the modification of mitochondrial enzymes by the attachment of lipoic acid groups
YOR142W	0.560119	*LSC1*	Alpha subunit of succinyl-CoA ligase
YJL124C	1.096419	*LSM1*	Lsm (Like Sm) protein
YIL070C	0.746344	*MAM33*	Specific translational activator for the mitochondrial COX1 mRNA
YBL091C	1.164116	*MAP2*	Methionine aminopeptidase; catalyzes the cotranslational removal of N-terminal methionine from nascent polypeptides
YER001W	0.944140	*MNN1*	Alpha-1,3-mannosyltransferase; integral membrane glycoprotein of the Golgi complex, required for addition of alpha1,3-mannose linkages to N-linked and O-linked oligosaccharides, one of five S. cerevisiae proteins of the MNN1 family
YOL090W	1.053158	*MRX10*	Mitochondrial inner membrane protein of unknown function
YOL090W	0.876705	*MSH2*	Protein that binds to DNA mismatches
YIL007C	0.582793	*NAS2*	Evolutionarily conserved 19S regulatory particle assembly-chaperone
YOL041C	1.156315	*NOP12*	Nucleolar protein involved in pre-25S rRNA processing
YIL038C	0.832525	*NOT3*	Component of the CCR4-NOT core complex, involved in mRNA decapping
YDR001C	0.742001	*NTH1*	Neutral trehalase, degrades trehalose; required for thermotolerance and may mediate resistance to other cellular stresses
YDL053C	0.935003	*PBP4*	Pbp1p binding protein; interacts strongly with Pab1p-binding protein 1 (Pbp1p) in the yeast two-hybrid system
YIL050W	0.751064	*PCL7*	Pho85p cyclin of the Pho80p subfamily; forms a functional kinase complex with Pho85p which phosphorylates Mmr1p and is regulated by Pho81p
YIL037C	1.587038	*PRM2*	Pheromone-regulated protein

2-MF 5G Resistant Strains (continue)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YJL078C	0.562116	*PRY3*	Cell wall-associated protein involved in export of acetylated sterols
YKL038W	1.013758	*RGT1*	Glucose-responsive transcription factor
YHL027W	0.773971	*RIM101*	Cys2His2 zinc-finger transcriptional repressor
YMR154C	0.863577	*RIM13*	Calpain-like cysteine protease; involved in proteolytic activation of Rim101p in response to alkaline pH
YGL045W	1.294868	*RIM8*	Protein involved in proteolytic activation of Rim101p; part of response to alkaline pH; interacts with ESCRT-1 subunits Stp22p and Vps28p
YIL133C	1.041550	*RPL16A*	Ribosomal 60S subunit protein L16A; N-terminally acetylated, binds 5.8 S rRNA
YGR034W	0.976025	*RPL26B*	Ribosomal 60S subunit protein L26B; binds to 5.8S rRNA
YER056C-A	1.419492	*RPL34A*	Ribosomal 60S subunit protein L34A; homologous to mammalian ribosomal protein L34, no bacterial homolog
YDL136W	0.705646	*RPL35B*	Ribosomal 60S subunit protein L35B; homologous to mammalian ribosomal protein L35 and bacterial L29
YKR094C	1.064640	*RPL40B*	Ubiquitin-ribosomal 60S subunit protein L40B fusion protein; cleaved to yield ubiquitin and ribosomal protein L40B
YIL069C	0.899697	*RPS24B*	Protein component of the small (40S) ribosomal subunit; homologous to mammalian ribosomal protein S24, no bacterial homolog
YBR147W	0.904173	*RTC2*	Putative vacuolar membrane transporter for cationic amino acids
YMR140W	0.644718	*SIP5*	Protein of unknown function; interacts with both the Reg1p/Glc7p phosphatase and the Snf1p kinase; forms cytoplasmic foci upon DNA replication stress
YLR025W	1.344478	*SNF7*	One of four subunits of the ESCRT-III complex; involved in the sorting of transmembrane proteins into the multivesicular body (MVB) pathway
YPL002C	1.222533	*SNF8*	Component of the ESCRT-II complex; ESCRT-II is involved in ubiquitin-dependent sorting of proteins into the endosome
YNL202W	1.183504	*SPS19*	Peroxisomal 2,4-dienoyl-CoA reductase
YDR463W	1.643277	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus

2-MF 5G Resistant Strains (continue)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YCL008C	1.406259	*STP22*	Component of the ESCRT-I complex; ESCRT-I is involved in ubiquitin-dependent sorting of proteins into the endosome
YLR372W	1.528722	*SYO1*	Transport adaptor or symportin; facilitates synchronized nuclear coimport of the two 5S-rRNA binding proteins Rpl5p and Rpl11p
YBR069C	1.347894	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine; low-affinity tryptophan and histidine transporter
YHR003C	0.806800	*TCD1*	tRNA threonylcarbamoyladenosine dehydratase
YMR313C	1.176192	*TGL3*	Bifunctional triacylglycerol lipase and LPE acyltransferase
YGR033C	0.531786	*TIM21*	Nonessential component of the TIM23 complex
YER175C	1.300788	*TMT1*	Trans-aconitate methyltransferase
YCR084C	2.470599	*TUP1*	General repressor of transcription
YDR092W	1.069309	*UBC13*	E2 ubiquitin-conjugating enzyme; involved in the error-free DNA postreplication repair pathway; interacts with Mms2p to assemble ubiquitin chains at the Ub Lys-63 residue; DNA damage triggers redistribution from the cytoplasm to the nucleus
YDL091C	1.014029	*UBX3*	Clathrin-coated vesicle component, regulator of endocytosis; copurifies with the DSC ubiquitin ligase complex
YMR077C	1.572156	*VPS20*	Myristoylated subunit of the ESCRT-III complex
YJR102C	1.236798	*VPS25*	Component of the ESCRT-II complex; ESCRT-II is involved in ubiquitin-dependent sorting of proteins into the endosome
YNR006W	1.700811	*VPS27*	Endosomal protein that forms a complex with Hse1p; required for recycling Golgi proteins
YLR417W	1.147501	*VPS36*	Component of the ESCRT-II complex; contains the GLUE (GRAM Like Ubiquitin binding in EAP45) domain which is involved in interactions with ESCRT-I and ubiquitin-dependent sorting of proteins into the endosome
YPR087W	1.150295	*VPS69*	Dubious open reading frame; unlikely to encode a functional protein, based on available experimental and comparative sequence data
YER123W	2.157933	*YCK3*	Palmitoylated vacuolar membrane-localized casein kinase I isoform
YDL072C	0.969937	*YET3*	Protein of unknown function
YBR090C	0.478962	-	Dubious open reading frame
YCR022C	0.573108	-	Dubious open reading frame

2-MF 5G Resistant Strains (continue)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDL063C	0.633821	-	Protein of unknown function
YDL187C	0.773123	-	Putative protein of unknown function
YDR282C	0.890066	-	Uncharacterized protein of unknown function
YDR286C	0.930466	-	Putative protein of unknown function
YDR417C	0.944626	-	Putative protein of unknown function
YGL088W	1.020849	-	Putative protein of unknown function
YGR122W	1.121476	-	Putative protein of unknown function
YHR003C	1.222179	-	Dubious open reading frame
YIL054W	1.229369	-	Dubious open reading frame
YIL067C	1.393537	-	Protein that may be involved in pH regulation
YJL132W	1.408255	-	Dubious open reading frame
YKR047W	1.437846	-	Dubious open reading frame
YLR050C	1.482371	-	Putative protein of unknown function
YNL109W	1.583849	-	Dubious open reading frame
YNL226W	1.664121	-	Dubious open reading frame

2-MF 10G Resistant

2-MF 10G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YCR107W	0.823396	*AAD3*	Putative aryl-alcohol dehydrogenase
YCR088W	0.696176	*ABP1*	Actin-binding protein of the cortical actin cytoskeleton; important for activation of the Arp2/3 complex that plays a key role actin in cytoskeleton organization
YNR033W	0.940598	*ABZ1*	Para-aminobenzoate (PABA) synthase; protein abundance increases in response to DNA replication stress
YLR040C	0.977940	*AFB1*	MATalpha-specific a-factor blocker
YCL050C	0.895886	*APA1*	AP4A phosphorylase; bifunctional diadenosine 5',5'''-P1,P4-tetraphosphate phosphorylase and ADP sulfurylase involved in catabolism of bis(5'-nucleosidyl) tetraphosphates
YCR048W	1.018289	*ARE1*	Acyl-CoA:sterol acyltransferase
YBL069W	0.881994	*AST1*	Lipid raft associated protein
YFL010W-A	0.891854	*AUA1*	Protein required for the negative regulation by ammonia of Gap1p; Gap1p is a general amino acid permease

2-MF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YBR068C	0.806415	*BAP2*	High-affinity leucine permease; functions as a branched-chain amino acid permease involved in uptake of leucine, isoleucine and valine
YKR099W	0.958811	*BAS1*	Myb-related transcription factor; involved in regulating basal and induced expression of genes of the purine and histidine biosynthesis pathways
YKR027W	1.000190	*BCH2*	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family
YFL007W	1.091339	*BLM10*	Proteasome activator
YCR032W	1.547986	*BPH1*	Protein homologous to Chediak-Higashi syndrome and Beige protein
YDR275W	1.342521	*BSC2*	Protein of unknown function
YDL151C	2.314531	*BUD30*	Dubious open reading frame
YKR036C	1.164941	*CAF4*	WD40 repeat-containing protein associated with the CCR4-NOT complex
YOR061W	1.002724	*CKA2*	Alpha' catalytic subunit of casein kinase 2 (CK2); CK2 is a Ser/Thr protein kinase with roles in cell growth and proliferation
YOR303W	0.694967	*CPA1*	Small subunit of carbamoyl phosphate synthetase; carbamoyl phosphate synthetase catalyzes a step in the synthesis of citrulline, an arginine precursor
YHR146W	0.861997	*CRP1*	Protein that binds to cruciform DNA structures
YIL132C	0.790177	*CSM2*	Component of Shu complex (aka PCSS complex); Shu complex also includes Psy3, Shu1, Shu2, and promotes error-free DNA repair
YGL078C	1.328866	*DBP3*	RNA-Dependent ATPase, member of DExD/H-box family
YLR422W	1.015083	*DCK1*	Dock family protein (Dedicator Of CytoKinesis), homolog of human DOCK1
YOR030W	1.742333	*DFG16*	Probable multiple transmembrane protein
YIR004W	1.028459	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YDL178W	0.489152	*DLD2*	D-2-hydroxyglutarate dehydrogenase, and minor D-lactate dehydrogenase
YMR126C	1.037315	*DLT1*	Protein of unknown function
YNL024C	1.678058	*EFM6*	Putative S-adenosylmethionine-dependent lysine methyltransferase
YLR372W	2.185969	*ELO3*	Elongase; involved in fatty acid and sphingolipid biosynthesis;
YKR096W	0.952516	*ESL2*	hEST1A/B (SMG5/6)-like protein; contributes to environment-sensing adaptive gene expression responses

2-MF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YML051W	0.668564	*GAL80*	Transcriptional regulator involved in the repression of GAL genes
YAL048C	1.041528	*GEM1*	Outer mitochondrial membrane GTPase, subunit of the ERMES complex
YGL057C	0.870683	*GEP7*	Protein of unknown function
YER020W	1.093352	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YDL022W	1.132977	*GPD1*	NAD-dependent glycerol-3-phosphate dehydrogenase; key enzyme of glycerol synthesis, essential for growth under osmotic stress
YGR163W	1.974387	*GTR2*	Subunit of a TORC1-stimulating GTPase complex
YPL001W	0.905596	*HAT1*	Catalytic subunit of the Hat1p-Hat2p histone acetyltransferase complex
YGL033W	0.731266	*HOP2*	Meiosis-specific protein that localizes to chromosomes; prevents synapsis between nonhomologous chromosomes and ensures synapsis between homologs
YJL082W	1.093158	*IML2*	Protein required for clearance of inclusion bodies; localizes to the inclusion bodies formed under protein misfolding stress
YJL051W	0.811875	*IRC8*	Bud tip localized protein of unknown function
YGL133W	1.317653	*ITC1*	Subunit of ATP-dependent Isw2p-Itc1p chromatin remodeling complex;
YLR451W	1.219118	*LEU3*	Zinc-knuckle transcription factor, repressor and activator
YLL007C	1.301549	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility)
YOR142W	0.897092	*LSC1*	Alpha subunit of succinyl-CoA ligase
YDL182W	0.646976	*LYS20*	Homocitrate synthase isozyme and functions in DNA repair
YIL070C	0.934593	*MAM33*	Specific translational activator for the mitochondrial COX1 mRNA
YBL091C	1.451405	*MAP2*	Methionine aminopeptidase
YPR153W	1.381125	*MAY24*	Protein of unknown function
YER001W	1.320033	*MNN1*	Alpha-1,3-mannosyltransferase; integral membrane glycoprotein of the Golgi complex, required for addition of alpha1,3-mannose linkages to N-linked and O-linked oligosaccharides
YOL090W	0.876784	*MSH2*	Protein that binds to DNA mismatches; forms heterodimers with Msh3p and Msh6p that bind to DNA mismatches to initiate the mismatch repair process
YBR255W	0.707760	*MTC4*	Protein of unknown function; required for normal growth rate at 15 degrees C

2-MF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YHR151C	1.020187	*MTC6*	Protein of unknown function
YIL007C	0.653050	*NAS2*	Evolutionarily conserved 19S regulatory particle assembly-chaperone
YKL151C	1.348943	*NNR2*	Widely-conserved NADHX dehydratase
YOL041C	1.168498	*NOP12*	Nucleolar protein involved in pre-25S rRNA processing
YDR001C	0.929250	*NTH1*	Neutral trehalase, degrades trehalose; required for thermotolerance and may mediate resistance to other cellular stresses
YGL151W	1.417570	*NUT1*	Component of the RNA polymerase II mediator complex
YER037W	0.706209	*PHM8*	Lysophosphatidic acid (LPA) phosphatase, nucleotidase
YGL023C	0.969040	*PIB2*	Protein of unknown function; contains FYVE domain; similar to Fab1 and Vps27
YGL037C	0.568543	*PNC1*	Nicotinamidase that converts nicotinamide to nicotinic acid; part of the NAD(+) salvage pathway
YOR161C	0.787003	*PNS1*	Protein of unknown function
YBL068W	0.932497	*PRS4*	5-phospho-ribosyl-1(alpha)-pyrophosphate synthetase, synthesizes PRPP; which is required for nucleotide, histidine, and tryptophan biosynthesis
YBR125C	1.512033	*PTC4*	Cytoplasmic type 2C protein phosphatase (PP2C)
YDL036C	1.351622	*PUS9*	Mitochondrial tRNA:pseudouridine synthase
YKL038W	1.079869	*RGT1*	Glucose-responsive transcription factor
YHL027W	1.931559	*RIM101*	Cys2His2 zinc-finger transcriptional repressor
YMR154C	2.158045	*RIM13*	Calpain-like cysteine protease
YOR275C	2.012190	*RIM20*	Protein involved in proteolytic activation of Rim101p
YNL294C	1.967248	*RIM21*	pH sensor molecule, component of the RIM101 pathway
YGL045W	2.382535	*RIM8*	Protein involved in proteolytic activation of Rim101p
YIL133C	0.810479	*RPL16A*	Ribosomal 60S subunit protein L16A; N-terminally acetylated, binds 5.8 S rRNA
YNL069C	1.651146	*RPL16B*	Ribosomal 60S subunit protein L16B; N-terminally acetylated, binds 5.8 S rRNA
YJL177W	1.535398	*RPL17B*	Ribosomal 60S subunit protein L17B; required for processing of 27SB pre-rRNA and formation of stable 66S assembly intermediates
YGL031C	1.814550	*RPL24A*	Ribosomal 60S subunit protein L24A
YDR471W	1.381398	*RPL27B*	Ribosomal 60S subunit protein L27B
YOR234C	1.485380	*RPL33B*	Ribosomal 60S subunit protein L33B
YER056C-A	1.319150	*RPL34A*	Ribosomal 60S subunit protein L34A
YDL136W	1.211105	*RPL35B*	Ribosomal 60S subunit protein L35B

2-MF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YKR094C	1.455254	*RPL40B*	Ubiquitin-ribosomal 60S subunit protein L40B fusion protein
YGL147C	1.198569	*RPL9A*	Ribosomal 60S subunit protein L9A
YNL206C	0.879954	*RTT106*	Histone chaperone
YPR129W	1.011860	*SCD6*	Repressor of translation initiation
YPR022C	0.623750	*SDD4*	Putative transcription factor, as suggested by computational analysis
YNL047C	1.003923	*SLM2*	Phosphoinositide PI4,5P(2) binding protein, forms a complex with Slm1p
YPL027W	0.638867	*SMA1*	Protein of unknown function
YLR025W	2.525267	*SNF7*	One of four subunits of the ESCRT-III complex
YPL002C	2.609805	*SNF8*	Component of the ESCRT-II complex
YDR463W	2.017219	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YCL008C	2.985180	*STP22*	Component of the ESCRT-I complex
YPR151C	0.730487	*SUE1*	Protein required for degradation of unstable forms of cytochrome c
YPL057C	1.534464	*SUR1*	Mannosylinositol phosphorylceramide (MIPC) synthase catalytic subunit
YDL063C	1.608369	*SYO1*	Transport adaptor or symportin
YBR069C	1.117648	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YBR083W	1.103002	*TEC1*	Transcription factor targeting filamentation genes and Ty1 expression
YMR313C	1.329278	*TGL3*	Bifunctional triacylglycerol lipase and LPE acyltransferase; major lipid particle-localized triacylglycerol (TAG) lipase
YGL179C	1.083416	*TOS3*	Protein kinase; related to and functionally redundant with Elm1p and Sak1p for the phosphorylation and activation of Snf1p
YGL096W	0.615299	*TOS8*	Homeodomain-containing protein and putative transcription factor
YPR156C	0.969790	*TPO3*	Polyamine transporter of the major facilitator superfamily
YNL299W	1.092556	*TRF5*	Non-canonical poly(A) polymerase
YCR084C	3.486019	*TUP1*	General repressor of transcription; forms complex with Cyc8p, involved in the establishment of repressive chromatin structure through interactions with histones H3 and H4, appears to enhance expression of some genes

2-MF 10G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGR080W	0.807046	*TWF1*	Twinfilin; highly conserved actin monomer-sequestering protein involved in regulation of the cortical actin cytoskeleton
YIL156W	0.671430	*UBP7*	Ubiquitin-specific protease that cleaves ubiquitin-protein fusions
YIL056W	0.785656	*VHR1*	Transcriptional activator; required for the vitamin H-responsive element (VHRE) mediated induction of VHT1 (Vitamin H transporter) and BIO5 (biotin biosynthesis intermediate transporter) in response to low biotin concentrations
YMR077C	3.159319	*VPS20*	Myristoylated subunit of the ESCRT-III complex
YJR102C	2.875490	*VPS25*	Component of the ESCRT-II complex
YNR006W	1.700430	*VPS27*	Endosomal protein that forms a complex with Hse1p
YPL065W	2.346278	*VPS28*	Component of the ESCRT-I complex
YLR417W	2.781907	*VPS36*	Component of the ESCRT-II complex
YDL072C	1.450249	*YET3*	Protein of unknown function
YOR003W	0.846431	*YSP3*	Putative precursor of the subtilisin-like protease III
YDL180W	0.483389	-	Putative protein of unknown function
YDL187C	0.637092	-	Dubious open reading frame
YOR325W	0.677892	-	Dubious open reading frame
YIL089W	0.716185	-	Protein of unknown function
YJL132W	0.839923	-	Putative protein of unknown function
YJL064W	0.898393	-	Dubious open reading frame
YLR108C	0.934379	-	Protein of unknown function
YGL081W	0.969411	-	Putative protein of unknown function
YDL034W	1.074698	-	Dubious open reading frame
YPR064W	1.182479	-	Putative protein of unknown function
YPR114W	1.197710	-	Putative protein of unknown function
YCR022C	1.229342	-	Putative protein of unknown function
YOR139C	1.406996	-	Dubious open reading frame
YDL062W	1.967881	-	Dubious open reading frame
YNL226W	2.035067	-	Dubious open reading frame
YGR122W	2.260126	-	Putative protein of unknown function

2-MF 15G Resistant

2-MF 15G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YCR088W	0.803360	*ABP1*	Actin-binding protein of the cortical actin cytoskeleton
YBR059C	1.146950	*AKL1*	Ser-Thr protein kinase; member (with Ark1p and Prk1p) of the Ark kinase family
YCR048W	1.097544	*ARE1*	Acyl-CoA:sterol acyltransferase
YDR101C	1.593369	*ARX1*	Nuclear export factor for the ribosomal pre-60S subunit
YPR026W	1.073300	*ATH1*	Acid trehalase required for utilization of extracellular trehalose
YKR099W	1.096564	*BAS1*	Myb-related transcription factor
YKR027W	1.474396	*BCH2*	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family
YCR032W	1.289379	*BPH1*	Protein homologous to Chediak-Higashi syndrome and Beige proteins
YDL151C	3.537701	*BUD30*	Dubious open reading frame
YKR036C	1.263340	*CAF4*	WD40 repeat-containing protein associated with the CCR4-NOT complex
YLR330W	1.568143	*CHS5*	Component of the exomer complex
YPR013C	0.943588	*CMR3*	Putative zinc finger protein
YOR303W	0.802635	*CPA1*	Small subunit of carbamoyl phosphate synthetase
YCR069W	1.040735	*CPR4*	Peptidyl-prolyl cis-trans isomerase (cyclophilin)
YBR036C	1.464484	*CSG2*	Endoplasmic reticulum membrane protein; protein abundance increases in response to DNA replication stress
YGL078C	1.923856	*DBP3*	RNA-Dependent ATPase, member of DExD/H-box family
YOR030W	3.031828	*DFG16*	Probable multiple transmembrane protein
YDR480W	1.200336	*DIG2*	MAP kinase-responsive inhibitor of the Ste12p transcription factor
YIR004W	1.309252	*DJP1*	Cytosolic J-domain-containing protein
YJL065C	0.955465	*DLS1*	Subunit of ISW2/yCHRAC chromatin accessibility complex
YMR126C	1.059443	*DLT1*	Protein of unknown function
YDR121W	1.579941	*DPB4*	Subunit of DNA pol epsilon and of ISW2 chromatin accessibility complex
YGR054W	3.510106	*ELF2A*	Eukaryotic initiation factor eIF2A
YLR372W	1.305546	*ELO3*	Elongase; involved in fatty acid and sphingolipid biosynthesis
YOR246C	1.364227	*ENV9*	Protein proposed to be involved in vacuolar functions

2-MF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YKR096W	1.542110	*ESL2*	hEST1A/B (SMG5/6)-like protein; contributes to environment-sensing adaptive gene expression responses; Esl2p and Esl1p contain a 14-3-3-like domain and a putative PilT N-terminus ribonuclease domain
YAL048C	0.841839	*GEM1*	Outer mitochondrial membrane GTPase, subunit of the ERMES complex
YGL057C	1.053351	*GEP7*	Protein of unknown function
YDR506C	1.299463	*GMC1*	Protein involved in meiotic progression
YER020W	1.163441	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YDL022W	1.067153	*GPD1*	NAD-dependent glycerol-3-phosphate dehydrogenase
YJL101C	1.135304	*GSH1*	Gamma glutamylcysteine synthetase
YOL089C	0.805418	*HAL9*	Putative transcription factor containing a zinc finger
YPL001W	1.357867	*HAT1*	Catalytic subunit of the Hat1p-Hat2p histone acetyltransferase complex
YOR038C	0.661511	*HIR2*	Subunit of HIR nucleosome assembly complex
YGL033W	0.824843	*HOP2*	Meiosis-specific protein that localizes to chromosomes
YJL082W	0.915485	*IML2*	Protein required for clearance of inclusion bodies
YJL051W	0.599250	*IRC8*	Bud tip localized protein of unknown function
YOR155C	1.325880	*ISN1*	Inosine 5'-monophosphate (IMP)-specific 5'-nucleotidase
YOR304W	1.612346	*ISW2*	ATP-dependent DNA translocase involved in chromatin remodeling
YGL133W	0.890151	*ITC1*	Subunit of ATP-dependent Isw2p-Itc1p chromatin remodeling complex
YER051W	0.938402	*JHD1*	JmjC domain family histone demethylase specific for H3-K36
YJL134W	1.358630	*LCB3*	Long-chain base-1-phosphate phosphatase
YPL056C	1.288599	*LCL1*	Putative protein of unknown function
YLL007C	0.817264	*LMO1*	Homolog of mammalian ELMO (Engulfment and celL MOtility)
YOR142W	1.168585	*LSC1*	Alpha subunit of succinyl-CoA ligase
YGR057C	1.189185	*LST7*	Subunit of the Lst4p-Lst7p GTPase activating protein complex for Gtr2p
YIL094C	0.751935	*LYS12*	Homo-isocitrate dehydrogenase
YIL070C	1.219028	*MAM33*	Specific translational activator for the mitochondrial COX1 mRNA
YBL091C	1.367231	*MAP2*	Methionine aminopeptidase
YPR153W	1.045356	*MAY24*	Protein of unknown function
YOL090W	0.428841	*MSH2*	Protein that binds to DNA mismatches

2-MF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDL154W	1.201669	*MSH5*	Protein of the MutS family; forms a dimer with Msh4p that facilitates crossovers between homologs during meiosis
YKL098W	1.002371	*MTC2*	Protein of unknown function
YBR255W	0.916276	*MTC4*	Protein of unknown function
YDR128W	1.208440	*MTC5*	Subunit of SEACAT, a subcomplex of the SEA complex
YHR151C	1.569802	*MTC6*	Protein of unknown function
YMR100W	0.762446	*MUB1*	MYND domain-containing protein
YDR493W	2.327581	*MZM1*	Protein required for assembly of the cytochrome bc(1) complex
YNL119W	0.828344	*NCS2*	Protein required for uridine thiolation of Lys(UUU) and Glu(UUC) tRNAs
YIL164C	1.200851	*NIT1*	Nitrilase; member of the nitrilase branch of the nitrilase superfamily
YKL151C	0.945846	*NNR2*	Widely-conserved NADHX dehydratase; converts (S)-NADHX to NADH in ATP-dependent manner; YKL151C promoter contains STREs (stress response elements) and expression is induced by heat shock or methyl methanesulfonate
YDR001C	1.103327	*NTH1*	Neutral trehalase, degrades trehalose; required for thermotolerance and may mediate resistance to other cellular stresses
YPR031W	2.003914	*NTO1*	Subunit of the NuA3 histone acetyltransferase complex
YGL151W	1.946072	*NUT1*	Component of the RNA polymerase II mediator complex
YIL136W	1.001251	*OM45*	Mitochondrial outer membrane protein of unknown function; major constituent of the outer membrane, extending into the intermembrane space
YBR125C	1.047572	*PTC4*	Cytoplasmic type 2C protein phosphatase (PP2C)
YJR059W	1.109999	*PTK2*	Serine/threonine protein kinase; involved in regulation of ion transport across plasma membrane
YKL038W	2.831222	*RGT1*	Glucose-responsive transcription factor
YHL027W	3.015097	*RIM101*	Cys2His2 zinc-finger transcriptional repressor
YMR154C	2.919111	*RIM13*	Calpain-like cysteine protease
YOR275C	2.784632	*RIM20*	Protein involved in proteolytic activation of Rim101p
YNL294C	3.463998	*RIM21*	pH sensor molecule, component of the RIM101 pathway
YGL045W	1.368262	*RIM8*	Protein involved in proteolytic activation of Rim101p
YPR065W	2.125339	*ROX1*	Heme-dependent repressor of hypoxic genes
YNL069C	2.798573	*RPL16B*	Ribosomal 60S subunit protein L16B
YBR191W	1.241494	*RPL21A*	Ribosomal 60S subunit protein L21A
YGR148C	1.545440	*RPL24B*	Ribosomal 60S subunit protein L24B

2-MF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YLR344W	1.073808	*RPL26A*	Ribosomal 60S subunit protein L26A
YGR034W	1.817811	*RPL26B*	Ribosomal 60S subunit protein L26B
YDR471W	1.721161	*RPL27B*	Ribosomal 60S subunit protein L27B
YFR032C-A	1.593129	*RPL29*	Ribosomal 60S subunit protein L29
YER056C-A	1.288523	*RPL34A*	Ribosomal 60S subunit protein L34A
YDL136W	1.393456	*RPL35B*	Ribosomal 60S subunit protein L35B
YKR094C	1.144230	*RPL40B*	Ubiquitin-ribosomal 60S subunit protein L40B fusion protein
YGL147C	0.869991	*RPL9A*	Ribosomal 60S subunit protein L9A
YBR147W	1.420287	*RTC2*	Putative vacuolar membrane transporter for cationic amino acids
YNL206C	0.972681	*RTT106*	Histone chaperone; involved in regulation of chromatin structure in both transcribed and silenced chromosomal regions
YHR206W	2.816944	*SKN7*	Nuclear response regulator and transcription factor
YBR266C	3.563277	*SLM6*	Protein with a potential role in actin cytoskeleton organization
YLR025W	3.584327	*SNF7*	One of four subunits of the ESCRT-III complex
YPL002C	1.007838	*SNF8*	Component of the ESCRT-II complex
YHR136C	1.964533	*SPL2*	Protein with similarity to cyclin-dependent kinase inhibitors
YHR066W	4.240095	*SSF1*	Constituent of 66S pre-ribosomal particles
YDR463W	2.802924	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YCL008C	1.926004	*STP22*	Component of the ESCRT-I complex; ESCRT-I is involved in ubiquitin-dependent sorting of proteins into the endosome
YPL057C	2.142227	*SUR1*	Mannosylinositol phosphorylceramide (MIPC) synthase catalytic subunit;
YDL063C	1.468225	*SYO1*	Transport adaptor or symportin
YBR069C	2.351491	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YBR150C	1.886792	*TBS1*	Protein of unknown function
YJL052W	1.296469	*TDH1*	Glyceraldehyde-3-phosphate dehydrogenase (GAPDH), isozyme 1
YMR313C	0.669662	*TGL3*	Bifunctional triacylglycerol lipase and LPE acyltransferase

2-MF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGL096W	1.716590	*TOS8*	Homeodomain-containing protein and putative transcription factor
YNL299W	4.435195	*TRF5*	Non-canonical poly(A) polymerase
YCR084C	1.026937	*TUP1*	General repressor of transcription
YGR080W	1.039559	*TWF1*	Twinfilin; highly conserved actin monomer-sequestering protein involved in regulation of the cortical actin cytoskeleton
YDL091C	0.761976	*UBX3*	Clathrin-coated vesicle component, regulator of endocytosis
YIL056W	1.143755	*VHR1*	Transcriptional activator; required for the vitamin H-responsive element (VHRE) mediated induction of VHT1 (Vitamin H transporter) and BIO5 (biotin biosynthesis intermediate transporter) in response to low biotin concentrations
YLR410W	4.372946	*VIP1*	Inositol hexakisphosphate and inositol heptakisphosphate kinase
YMR077C	3.243955	*VPS20*	Myristoylated subunit of the ESCRT-III complex
YJR102C	2.520390	*VPS25*	Component of the ESCRT-II complex
YNR006W	3.270868	*VPS27*	Endosomal protein that forms a complex with Hse1p
YPL065W	3.534697	*VPS28*	Component of the ESCRT-I complex
YLR417W	0.856668	*VPS36*	Component of the ESCRT-II complex
YDR089W	1.383326	*VTC5*	Novel subunit of the vacuolar transporter chaperone complex
YDL072C	0.541111	*YET3*	Protein of unknown function
YDL180W	0.677139	-	Putative protein of unknown function
YDL187C	0.739914	-	Dubious open reading frame
YJL132W	0.830627	-	Putative protein of unknown function
YIL089W	0.920996	-	Protein of unknown function
YJL064W	1.029887	-	Dubious open reading frame
YLR108C	1.048183	-	Protein of unknown function
YGL081W	1.067312	-	Putative protein of unknown function
YDR491C	1.192663	-	Dubious open reading frame
YLR434C	1.203335	-	Dubious open reading frame
YOR364W	1.286407	-	Dubious open reading frame
YCR022C	1.325904	-	Putative protein of unknown function
YPR114W	1.557578	-	Putative protein of unknown function
YOR139C	1.566757	-	Dubious open reading frame
YPR064W	1.607261	-	Putative protein of unknown function
YNL109W	1.653033	-	Dubious open reading frame

2-MF 15G Resistant Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDL062W	2.143351	-	Dubious open reading frame
YNL226W	2.943941	-	Dubious open reading frame
YGR122W	3.020169	-	Putative protein of unknown function

2-MF Resistant across all time points

2-MF Resistant Across All Time Points		
ORF Name	**Gene Name**	**Deleted Gene Function**
YCR048W	*ARE1*	Acyl-CoA:sterol acyltransferase
YKR027W	*BCH2*	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family
YCR032W	*BPH1*	Protein homologous to Chediak-Higashi syndrome and Beige proteins
YKR036C	*CAF4*	WD40 repeat-containing protein associated with the CCR4-NOT complex
YOR303W	*CPA1*	Small subunit of carbamoyl phosphate synthetase
YGL078C	*DBP3*	RNA-Dependent ATPase, member of DExD/H-box family
YIR004W	*DJP1*	Cytosolic J-domain-containing protein; required for peroxisomal protein import and involved in peroxisome assembly
YLR372W	*ELO3*	Elongase; involved in fatty acid and sphingolipid biosynthesis
YGL057C	*GEP7*	Protein of unknown function
YER020W	*GPA2*	Nucleotide binding alpha subunit of the heterotrimeric G protein
YGL033W	*HOP2*	Meiosis-specific protein that localizes to chromosomes; complexes with Mnd1p to promote homolog pairing and meiotic double-strand break repair
YJL082W	*IML2*	Protein required for clearance of inclusion bodies; localizes to the inclusion bodies formed under protein misfolding stress
YOR142W	*LSC1*	Alpha subunit of succinyl-CoA ligase
YIL070C	*MAM33*	Specific translational activator for the mitochondrial COX1 mRNA
YBL091C	*MAP2*	Methionine aminopeptidase; catalyzes the cotranslational removal of N-terminal methionine from nascent polypeptides
YOL090W	*MSH2*	Protein that binds to DNA mismatches; forms heterodimers with Msh3p and Msh6p that bind to DNA mismatches to initiate the mismatch repair process
YDR001C	*NTH1*	Neutral trehalase, degrades trehalose; required for thermotolerance and may mediate resistance to other cellular stresses
YKL038W	*RGT1*	Glucose-responsive transcription factor
YHL027W	*RIM101*	Cys2His2 zinc-finger transcriptional repressor
YMR154C	*RIM13*	Calpain-like cysteine protease; involved in proteolytic activation of Rim101p in response to alkaline pH

2-MF Resistant Across All Time Points		
ORF Name	**ORF Name**	**ORF Name**
YGL045W	*RIM8*	Protein involved in proteolytic activation of Rim101p
YER056C-A	*RPL34A*	Ribosomal 60S subunit protein L34A
YDL136W	*RPL35B*	Ribosomal 60S subunit protein L35B
YKR094C	*RPL40B*	Ubiquitin-ribosomal 60S subunit protein L40B fusion protein
YLR025W	*SNF7*	One of four subunits of the ESCRT-III complex
YPL002C	*SNF8*	Component of the ESCRT-II complex
YDR463W	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YCL008C	*STP22*	Component of the ESCRT-I complex
YDL063C	*SYO1*	Transport adaptor or symportin; facilitates synchronized nuclear coimport of the two 5S-rRNA binding proteins Rpl5p and Rpl11p
YBR069C	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YMR313C	*TGL3*	Bifunctional triacylglycerol lipase and LPE acyltransferase
YCR084C	*TUP1*	General repressor of transcription
YMR077C	*VPS20*	Myristoylated subunit of the ESCRT-III complex
YJR102C	*VPS25*	Component of the ESCRT-II complex
YNR006W	*VPS27*	Endosomal protein that forms a complex with Hse1p
YLR417W	*VPS36*	Component of the ESCRT-II complex
YDL072C	*YET3*	Protein of unknown function
YCR022C	-	Putative protein of unknown function
YGR122W	-	Protein that may be involved in pH regulation
YNL226W	-	Dubious open reading frame
YJL132W	-	Putative protein of unknown function
YDL187C	-	Dubious open reading frame

2-MF 5G Sensitive

2-MF 5G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YNL077W	-1.579608	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family; conserved across eukaryotes
YDR035W	-1.602356	*ARO3*	3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase
YMR159C	-0.781731	*ATG16*	Conserved protein involved in autophagy
YBR290W	-1.188367	*BSD2*	Heavy metal ion homeostasis protein
YKL005C	-1.213721	*BYE1*	Negative regulator of transcription elongation

2-MF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOR125C	-0.759814	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YNL051W	-1.595877	*COG5*	Component of the conserved oligomeric Golgi complex
YNL041C	-2.209737	*COG6*	Component of the conserved oligomeric Golgi complex
YPL200W	-3.032029	*CSM4*	Protein required for accurate chromosome segregation during meiosis
YJL005W	-1.425142	*CYR1*	Adenylate cyclase
YGR092W	-1.188071	*DBF2*	Ser/Thr kinase involved in transcription and stress response; functions as part of a network of genes in exit from mitosis; localization is cell cycle regulated
YPL265W	-1.504607	*DIP5*	Dicarboxylic amino acid permease
YMR299C	-1.640758	*DYN3*	Dynein light intermediate chain (LIC)
YGL020C	-1.308598	*GET1*	Subunit of the GET complex
YDR096W	-0.924255	*GIS1*	Histone demethylase and transcription factor
YMR251W	-1.653073	*GTO3*	Omega class glutathione transferase; putative cytosolic localization
YLL060C	-0.969112	*GTT2*	Glutathione S-transferase capable of homodimerization
YKR019C	-1.028624	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YFR024C-A	-0.949540	*LSB3*	Protein containing a C-terminal SH3 domain
YBR298C	-1.435891	*MAL31*	Maltose permease; high-affinity maltose transporter (alpha-glucoside transporter)
YNL307C	-1.079643	*MCK1*	Dual-specificity ser/thr and tyrosine protein kinase
YML062C	-1.260572	*MFT1*	Subunit of the THO complex
YMR115W	-1.142667	*MGR3*	Subunit of the mitochondrial (mt) i-AAA protease supercomplex
YGL209W	-1.207690	*MIG2*	Zinc finger transcriptional repressor
YDR144C	-1.360318	*MKC7*	GPI-anchored aspartyl protease
YJR074W	-1.157446	*MOG1*	Conserved nuclear protein that interacts with GTP-Gsp1p
YKL138C	-1.186758	*MRPL31*	Mitochondrial ribosomal protein of the large subunit
YCR026C	-2.697337	*NPP1*	Nucleotide pyrophosphatase/phosphodiesterase
YHR179W	-2.125392	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN); may be involved in sterol metabolism, oxidative stress response, and programmed cell death; protein abundance increases in response to DNA replication stress
YGR222W	-1.117719	*PET54*	Mitochondrial inner membrane protein

2-MF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGR077C	-1.132406	*PEX8*	Intraperoxisomal organizer of the peroxisomal import machinery
YNL097C	-1.071814	*PHO23*	Component of the Rpd3L histone deacetylase complex
YML123C	-1.231401	*PHO84*	High-affinity inorganic phosphate (Pi) transporter
YGR086C	-2.096025	*PIL1*	Eisosome core component
YNL098C	-0.992745	*RAS2*	GTP-binding protein
YDR202C	-1.118695	*RAV2*	Subunit of RAVE complex (Rav1p, Rav2p, Skp1p)
YMR247C	-1.299218	*RKR1*	RING domain E3 ubiquitin ligase
YMR242C	-1.072786	*RPL20A*	Ribosomal 60S subunit protein L20A
YDL020C	-0.877670	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YER074W	-2.141142	*RPS24A*	Protein component of the small (40S) ribosomal subunit
YDL204W	-3.010831	*RTN2*	Reticulon protein; involved in nuclear pore assembly and maintenance of tubular ER morphology
YMR263W	-2.684506	*SAP30*	Component of Rpd3L histone deacetylase complex; involved in silencing at telomeres, rDNA, and silent mating-type loci; involved in telomere maintenance
YLR268W	-0.770286	*SEC22*	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p
YNL012W	-2.353499	*SPO1*	Meiosis-specific prospore protein
YLR055C	-2.926154	*SPT8*	Subunit of the SAGA transcriptional regulatory complex
YHL007C	-1.000644	*STE20*	Cdc42p-activated signal transducing kinase
YDR457W	-1.250520	*TOM1*	E3 ubiquitin ligase of the hect-domain class
YGR194C	-1.461274	*XKS1*	Xylulokinase; converts D-xylulose and ATP to xylulose 5-phosphate and ADP
YML007W	-1.855474	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; activated by H2O2 through the multistep formation of disulfide bonds and transit from the cytoplasm to the nucleus; Yap1p is degraded in the nucleus after the oxidative stress has passed; mediates resistance to cadmium; relative distribution to the nucleus increases upon DNA replication stress; YAP1 has a paralog, CAD1, that arose from the whole genome duplication
YCR059C	-0.883926	*YIH1*	Negative regulator of eIF2 kinase Gcn2p

2-MF 5G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YLR262C	-1.531928	*YPT6*	Rab family GTPase; required for endosome-to-Golgii, intra-Golgi retrograde, and retrograde Golgi-to-ER transport; temporarily at the Golgi, dissociating into the cytosol on arrival of the late Golgi GTPase Ypt32p; Golgi-localized form is GTP bound, while cytosolic form is GDP-bound; required for delivery of Atg9p to the phagophore assembly site during autophagy under heat stress, with Ypt6p for starvation induced autophagy and for the CVT pathway; homolog of mammalian Rab6
YMR099C	-1.164334	*YMR099C*	Dubious open reading frame; unlikely to encode a functional protein, based on available experimental and comparative sequence data; completely overlaps verified gene COQ2
YCR061W	-3.662968	-	Glucose-6-phosphate 1-epimerase (hexose-6-phosphate mutarotase)
YDR048C	-2.033354	-	Protein of unknown function
YJL185C	-1.449976	-	Dubious open reading frame
YJR128W	-1.397281	-	Dubious open reading frame
YJR129C	-1.389779	-	Dubious open reading frame
YMR244W	-1.012968	-	Putative protein of unknown function
YNR042W	-0.996287	-	Putative protein of unknown function
YNR061C	-0.949444	-	Putative protein of unknown function
YPR039W	-0.848692	-	Putative protein of unknown function

2-MF 10G Sensitive

2-MF 10G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YBR194W	-6.320242	*AIM4*	Protein proposed to be associated with the nuclear pore complex
YLR131C	-2.119451	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YIL087C	-1.024901	*AIM19*	Protein of unknown function; mitochondrial protein that physically interacts with Tim23p
YML035C	-4.041456	*AMD1*	AMP deaminase; tetrameric enzyme that catalyzes the deamination of AMP to form IMP and ammonia; thought to be involved in regulation of intracellular purine (adenine, guanine, and inosine) nucleotide pools

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YBR286W	-1.049249	*APE3*	Vacuolar aminopeptidase Y; processed to mature form by Prb1p
YNL077W	-2.502452	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation; member of the DnaJ-like family
YCR068W	-2.400323	*ATG15*	Phospholipase; preferentially hydrolyses phosphatidylserine, with minor activity against cardiolipin and phosphatidylethanolamine
YPL078C	-1.647820	*ATP4*	Subunit b of the stator stalk of mitochondrial F1F0 ATP synthase
YBR290W	-2.851730	*BSD2*	Heavy metal ion homeostasis protein; facilitates trafficking of Smf1p and Smf2p metal transporters to vacuole where they are degraded
YNL288W	-1.049398	*CAF40*	Component of the CCR4-NOT transcriptional complex
YOR125C	-1.598811	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YDR357C	-1.172977	*CNL1*	Subunit of the BLOC-1 complex involved in endosomal maturation
YML110C	-1.172318	*COQ5*	2-hexaprenyl-6-methoxy-1,4-benzoquinone methyltransferase; involved in ubiquinone (Coenzyme Q) biosynthesis
YML078W	-0.749314	*CPR3*	Mitochondrial peptidyl-prolyl cis-trans isomerase (cyclophilin)
YOL159C	-0.767503	*CSS3*	Protein of unknown function
YBR291C	-2.456760	*CTP1*	Mitochondrial inner membrane citrate transporter; member of the mitochondrial carrier family
YDL117W	-2.859789	*CYK3*	SH3-domain protein located in the bud neck and cytokinetic actin ring
YJL005W	-1.300689	*CYR1*	Adenylate cyclase
YGR092W	-3.206381	*DBF2*	Ser/Thr kinase involved in transcription and stress response
YPL265W	-2.468472	*DIP5*	Dicarboxylic amino acid permease; mediates high-affinity and high-capacity transport of L-glutamate and L-aspartate
YBR278W	-1.369949	*DPB3*	Third-largest subunit of DNA polymerase II (DNA polymerase epsilon); required to maintain fidelity of chromosomal replication and also for inheritance of telomeric silencing
YMR276W	-0.696092	*DSK2*	Nuclear-enriched ubiquitin-like polyubiquitin-binding protein
YBR281C	-1.433802	*DUG2*	Component of glutamine amidotransferase (GATase II)
YMR299C	-1.643736	*DYN3*	Dynein light intermediate chain (LIC)

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YNL136W	-2.170030	*EAF7*	Subunit of the NuA4 histone acetyltransferase complex
YLR436C	-0.898310	*ECM30*	Protein of unknown function
YJR129C	-1.523496	*EFM3*	S-adenosylmethionine-dependent methyltransferase
YCR034W	-1.259304	*ELO2*	Fatty acid elongase, involved in sphingolipid biosynthesis
YLR318W	-1.420056	*EST2*	Reverse transcriptase subunit of the telomerase holoenzyme
YJL155C	-1.672388	*FBP26*	Fructose-2,6-bisphosphatase, required for glucose metabolism
YIL065C	-1.785777	*FIS1*	Protein involved in mitochondrial fission and peroxisome abundance
YOR070C	-1.819124	*GYP1*	Cis-golgi GTPase-activating protein (GAP) for yeast Rabs
YPL244C	-1.487653	*HUT1*	Protein with a role in UDP-galactose transport to the Golgi lumen
YIL154C	-2.485199	*IMP2'*	Transcriptional activator involved in maintenance of ion homeostasis
YMR044W	-1.843268	*IOC4*	Member of a complex (Isw1b) with Isw1p and Ioc2p
YOL081W	-3.276626	*IRA2*	GTPase-activating protein; negatively regulates RAS by converting it from the GTP- to the GDP-bound inactive form
YKR019C	-2.443888	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YJL162C	-1.460199	*JJJ2*	Protein of unknown function
YGL203C	-1.106996	*KEX1*	Cell death protease essential for hypochlorite-induced apoptosis
YGL216W	-1.233412	*KIP3*	Kinesin-related antiparallel sliding motor protein
YNL322C	-1.339237	*KRE1*	Cell wall glycoprotein involved in beta-glucan assembly
YNL323W	-2.476411	*LEM3*	Membrane protein of the plasma membrane and ER
YFR024C-A	-1.219669	*LSB3*	Protein containing a C-terminal SH3 domain
YAL024C	-4.649803	*LTE1*	Protein similar to GDP/GTP exchange factors
YEL053C	-1.131493	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YBR298C	-2.053853	*MAL31*	Maltose permease; high-affinity maltose transporter (alpha-glucoside transporter)
YBR297W	-1.288892	*MAL33*	MAL-activator protein
YNL307C	-1.577446	*MCK1*	Dual-specificity ser/thr and tyrosine protein kinase
YML062C	-1.884985	*MFT1*	Subunit of the THO complex; involved in telomere maintenance
YMR115W	-2.957041	*MGR3*	Subunit of the mitochondrial (mt) i-AAA protease supercomplex; i-AAA degrades misfolded mitochondrial proteins

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YNL100W	-0.903292	*MIC27*	Component of the MICOS complex
YMR036C	-1.002473	*MIH1*	Protein tyrosine phosphatase involved in cell cycle control
YEL007W	-1.580623	*MIT1*	Transcriptional regulator of pseudohyphal growth
YLR035C	-1.077794	*MLH2*	Protein involved in mismatch repair and meiotic recombination
YJR074W	-6.226682	*MOG1*	Conserved nuclear protein that interacts with GTP-Gsp1p
YDR162C	-4.373172	*NBP2*	Protein involved in the HOG (high osmolarity glycerol) pathway
YHR004C	-1.794501	*NEM1*	Probable catalytic subunit of Nem1p-Spo7p phosphatase holoenzyme
YNL123W	-1.637548	*NMA111*	Serine protease and general molecular chaperone; involved in response to heat stress and promotion of apoptosis
YDR432W	-6.182218	*NPL3*	RNA-binding protein; promotes elongation, regulates termination, and carries poly(A) mRNA from nucleus to cytoplasm
YCR026C	-1.769405	*NPP1*	Nucleotide pyrophosphatase/phosphodiesterase; mediates extracellular nucleotide phosphate hydrolysis along with Npp2p and Pho5p
YBR188C	-1.130229	*NTC20*	Member of the NineTeen Complex (NTC)
YBL079W	-3.119160	*NUP170*	Subunit of inner ring of nuclear pore complex (NPC)
YLR335W	-1.111987	*NUP2*	Nucleoporin involved in nucleocytoplasmic transport
YNL099C	-1.061782	*OCA1*	Putative protein tyrosine phosphatase; required for cell cycle arrest in response to oxidative damage of DNA
YHL020C	-2.891999	*OPI1*	Transcriptional regulator of a variety of genes
YHR179W	-2.580863	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN); may be involved in sterol metabolism, oxidative stress response, and programmed cell death; protein abundance increases in response to DNA replication stress
YBR295W	-1.542771	*PCA1*	Cadmium transporting P-type ATPase
YNR045W	-1.759150	*PET494*	Mitochondrial translational activator specific for the COX3 mRNA
YNL097C	-3.586818	*PHO23*	Component of the Rpd3L histone deacetylase complex
YBR296C	-1.034167	*PHO89*	Plasma membrane Na+/Pi cotransporter; active in early growth phase
YCR024C-A	-1.693138	*PMP1*	Regulatory subunit for the plasma membrane H(+)-ATPase Pma1p
YAL023C	-4.893265	*PMT2*	Protein O-mannosyltransferase of the ER membrane

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YBR022W	-1.178734	*POA1*	Phosphatase that is highly specific for ADP-ribose 1''-phosphate; a tRNA splicing metabolite; may have a role in regulation of tRNA splicing
YDL134C	-0.992481	*PPH21*	Catalytic subunit of protein phosphatase 2A (PP2A)
YNL098C	-1.615861	*RAS2*	GTP-binding protein; regulates nitrogen starvation response, sporulation, and filamentous growth
YJL217W	-1.151717	*REE1*	Cytoplasmic protein involved in the regulation of enolase (ENO1)
YGL250W	-2.139808	*RMR1*	Protein required for meiotic recombination and gene conversion
YIL148W	-1.437976	*RPL40A*	Ubiquitin-ribosomal 60S subunit protein L40A fusion protein
YDL020C	-3.985606	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YCR045C	-1.587381	*RRT12*	Probable subtilisin-family protease
YMR263W	-3.397736	*SAP30*	Component of Rpd3L histone deacetylase complex; involved in silencing at telomeres, rDNA, and silent mating-type loci; involved in telomere maintenance
YMR305C	-1.208042	*SCW10*	Cell wall protein with similarity to glucanases
YLR268W	-7.230171	*SEC22*	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p
YGL066W	-2.585032	*SGF73*	Subunit of DUBm module of SAGA and SLIK
YNL196C	-1.252873	*SLZ1*	Sporulation-specific protein with a leucine zipper motif
YDR006C	-1.081862	*SOK1*	Protein of unknown function
YDR392W	-2.764275	*SPT3*	Subunit of the SAGA and SAGA-like transcriptional regulatory complexes
YLR055C	-2.137948	*SPT8*	Subunit of the SAGA transcriptional regulatory complex
YHL007C	-2.448635	*STE20*	Cdc42p-activated signal transducing kinase
YJL004C	-2.040557	*SYS1*	Integral membrane protein of the Golgi
YHR167W	-2.153594	*THP2*	Subunit of the THO and TREX complexes
YPR074C	-4.096908	*TKL1*	Transketolase; catalyzes conversion of xylulose-5-phosphate and ribose-5-phosphate to sedoheptulose-7-phosphate and glyceraldehyde-3-phosphate in the pentose phosphate pathway
YDR457W	-3.634043	*TOM1*	E3 ubiquitin ligase of the hect-domain class
YDR120C	-1.847102	*TRM1*	tRNA methyltransferase
YDR108W	-1.838352	*TRS85*	Component of transport protein particle (TRAPP) complex III
YDL190C	-0.869943	*UFD2*	Ubiquitin chain assembly factor (E4)
YBR006W	-1.798114	*UGA2*	Succinate semialdehyde dehydrogenase

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YPL139C	-1.565075	*UME1*	Component of both the Rpd3S and Rpd3L histone deacetylase complexes
YIL017C	-2.181051	*VID28*	GID Complex subunit, serves as adaptor for regulatory subunit Vid24p;
YOR069W	-1.618123	*VPS5*	Nexin-1 homolog; required for localizing membrane proteins from a prevacuolar/late endosomal compartment back to late Golgi
YOR043W	-2.069352	*WHI2*	Protein required for full activation of the general stress response
YGR194C	-0.977812	*XKS1*	Xylulokinase; converts D-xylulose and ATP to xylulose 5-phosphate and ADP
YML007W	-3.703475	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YLR020C	-1.585105	*YEH2*	Steryl ester hydrolase
YER041W	-1.175374	*YEN1*	Holliday junction resolvase
YFR007W	-1.923440	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YCR059C	-1.570233	*YIH1*	Negative regulator of eIF2 kinase Gcn2p
YNL237W	-1.781428	*YTP1*	Probable type-III integral membrane protein of unknown function
YBR284W	-2.845409	*YBR284W*	Putative metallo-dependent hydrolase superfamily protein
YMR099C	-2.265864	*YMR099C*	Glucose-6-phosphate 1-epimerase (hexose-6-phosphate mutarotase)
YDL183C	-2.178185	*YDL183C*	Protein that may form an active mitochondrial KHE system
YBL100C	-1.884527	-	Dubious open reading frame
YJL193W	-1.857588	-	Putative protein of unknown function
YCR061W	-1.711321	-	Protein of unknown function
YCR087W	-1.661878	-	Dubious open reading frame
YDR149C	-1.660672	-	Dubious open reading frame
YCR050C	-1.655101	-	Non-essential protein of unknown function
YOL163W	-1.570261	-	Putative protein of unknown function
YDR537C	-1.521862	-	Dubious open reading frame
YCR087C-A	-1.501218	-	Putative protein of unknown function
YNR042W	-1.498250	-	Dubious open reading frame
YPR146C	-1.468608	-	Dubious open reading frame
YDR431W	-1.345096	-	Dubious open reading frame

2-MF 10G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YGR035C	-1.337407	-	Putative protein of unknown function
YNL013C	-1.279017	-	Dubious open reading frame
YBR292C	-1.236470	-	Dubious open reading frame
YCR085W	-1.137312	-	Putative protein of unknown function
YOR225W	-1.112290	-	Dubious open reading frame
YNL115C	-1.087630	-	Putative protein of unknown function
YGR259C	-1.023921	-	Dubious open reading frame
YDR048C	-1.003763	-	Dubious open reading frame
YGL042C	-0.985340	-	Dubious open reading frame
YCR049C	-0.942974	-	Dubious open reading frame
YOL085C	-0.890584	-	Putative protein of unknown function
YNR061C	-0.797835	-	Protein of unknown function

2-MF 15G Sensitive

2-MF 15G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YFL056Ca	-7.415635	*AAD6*	Putative aryl-alcohol dehydrogenase
YLR131C	-1.278635	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YIL087C	-1.220251	*AIM19*	Protein of unknown function
YBR286W	-2.649220	*APE3*	Vacuolar aminopeptidase Y
YNL077W	-6.859713	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation
YDL192W	-0.830221	*ARF1*	ADP-ribosylation factor; GTPase of the Ras superfamily involved in regulation of coated vesicle formation in intracellular trafficking within the Golgi
YIL130W	-2.699690	*ASG1*	Zinc cluster protein proposed to be a transcriptional regulator; regulator involved in the stress response
YCR068W	-5.218744	*ATG15*	Phospholipase; preferentially hydrolyses phosphatidylserine, with minor activity against cardiolipin and phosphatidylethanolamine
YPL078C	-4.002264	*ATP4*	Subunit b of the stator stalk of mitochondrial F1F0 ATP synthase
YBR290W	-1.386475	*BSD2*	Heavy metal ion homeostasis protein
YLR319C	-1.222312	*BUD6*	Actin- and formin-interacting protein
YMR275C	-1.825758	*BUL1*	Ubiquitin-binding component of the Rsp5p E3-ubiquitin ligase complex

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOR125C	-1.864373	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YKL208W	-1.635037	*CBT1*	Protein involved in 5' RNA end processing
YDR254W	-2.064220	*CHL4*	Outer kinetochore protein required for chromosome stability
YPL241C	-1.071311	*CIN2*	GTPase-activating protein (GAP) for Cin4p
YNR001C	-1.597698	*CIT1*	Citrate synthase; catalyzes the condensation of acetyl coenzyme A and oxaloacetate to form citrate
YML110C	-1.350594	*COQ5*	2-hexaprenyl-6-methoxy-1,4-benzoquinone methyltransferase
YPL172C	-0.909607	*COX10*	Heme A:farnesyltransferase; catalyzes first step in conversion of protoheme to heme A prosthetic group required for cytochrome c oxidase activity
YLR038C	-1.970031	*COX12*	Subunit VIb of cytochrome c oxidase
YLL018C-A	-0.993821	*COX19*	Protein required for cytochrome c oxidase assembly
YML078W	-3.356079	*CPR3*	Mitochondrial peptidyl-prolyl cis-trans isomerase (cyclophilin)
YLR087C	-2.220065	*CSF1*	Protein required for fermentation at low temperature
YCR086W	-0.817225	*CSM1*	Nucleolar protein that mediates homolog segregation during meiosis I
YOL159C	-3.048299	*CSS3*	Protein of unknown function
YBR291C	-1.481115	*CTP1*	Mitochondrial inner membrane citrate transporter
YML113W	-6.795769	*DAT1*	DNA binding protein that recognizes oligo(dA).oligo(dT) tracts
YGR092W	-3.931766	*DBF2*	Ser/Thr kinase involved in transcription and stress response; functions as part of a network of genes in exit from mitosis
YPL265W	-1.348863	*DIP5*	Dicarboxylic amino acid permease; mediates high-affinity and high-capacity transport of L-glutamate and L-aspartate
YMR276W	-1.423329	*DSK2*	Nuclear-enriched ubiquitin-like polyubiquitin-binding protein
YBR281C	-1.237421	*DUG2*	Component of glutamine amidotransferase (GATase II)
YMR299C	-5.688198	*DYN3*	Dynein light intermediate chain (LIC)
YNL136W	-1.186960	*EAF7*	Subunit of the NuA4 histone acetyltransferase complex
YLR436C	-2.570277	*ECM30*	Protein of unknown function
YJR129C	-2.019021	*EFM3*	S-adenosylmethionine-dependent methyltransferase
YCR034W	-2.028307	*ELO2*	Fatty acid elongase, involved in sphingolipid biosynthesis
YMR015C	-2.536330	*ERG5*	C-22 sterol desaturase

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YJL155C	-3.144793	*FBP26*	Fructose-2,6-bisphosphatase, required for glucose metabolism; protein abundance increases in response to DNA replication stress
YIL065C	-0.895339	*FIS1*	Protein involved in mitochondrial fission and peroxisome abundance
YLR454W	-1.432406	*FMP27*	Putative protein of unknown function
YJR040W	-4.754243	*GEF1*	Voltage-gated chloride channel
YGL020C	-5.765378	*GET1*	Subunit of the GET complex
YER083C	-1.135675	*GET2*	Subunit of the GET complex
YMR311C	-0.976193	*GLC8*	Regulatory subunit of protein phosphatase 1 (Glc7p)
YLL060C	-1.824773	*GTT2*	Glutathione S-transferase capable of homodimerization
YOR070C	-2.072129	*GYP1*	Cis-golgi GTPase-activating protein (GAP) for yeast Rabs
YGR187C	-1.821889	*HGH1*	Nonessential protein of unknown function
YDL066W	-1.256462	*IDP1*	Mitochondrial NADP-specific isocitrate dehydrogenase
YMR035W	-0.873326	*IMP2*	Catalytic subunit of mitochondrial inner membrane peptidase complex
YIR024C	-5.755591	*INA22*	F1F0 ATP synthase peripheral stalk assembly factor
YOL081W	-4.567548	*IRA2*	GTPase-activating protein
YKR019C	-1.263834	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YPL145C	-1.278238	*KES1*	One of seven members of the yeast oxysterol binding protein family
YPL155C	-1.992579	*KIP2*	Kinesin-related motor protein involved in mitotic spindle positioning; stabilizes microtubules by targeting Bik1p to the plus end
YNL322C	-3.367964	*KRE1*	Cell wall glycoprotein involved in beta-glucan assembly
YOR322C	-2.385248	*LDB19*	Alpha-arrestin involved in ubiquitin-dependent endocytosis
YNL323W	-1.421697	*LEM3*	Membrane protein of the plasma membrane and ER
YFR024C-A	-1.471976	*LSB3*	Protein containing a C-terminal SH3 domain
YEL053C	-1.844089	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YPR051W	-2.774235	*MAK3*	Catalytic subunit of the NatC type N-terminal acetyltransferase (NAT)
YBR298C	-2.177613	*MAL31*	Maltose permease
YBR297W	-2.461836	*MAL33*	MAL-activator protein;
YNL307C	-3.944692	*MCK1*	Dual-specificity ser/thr and tyrosine protein kinase
YML062C	-1.473154	*MFT1*	Subunit of the THO complex

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YMR036C	-2.313372	*MIH1*	Protein tyrosine phosphatase involved in cell cycle control
YEL007W	-1.487157	*MIT1*	Transcriptional regulator of pseudohyphal growth
YMR070W	-1.053629	*MOT3*	Transcriptional repressor, activator; role in cellular adjustment to osmotic stress including modulation of mating efficiency
YMR037C	-4.302750	*MSN2*	Stress-responsive transcriptional activator; activated in stochastic pulses of nuclear localization in response to various stress conditions; binds DNA at stress response elements of responsive genes
YKR048C	-3.171312	*NAP1*	Histone chaperone
YHR004C	-3.200743	*NEM1*	Probable catalytic subunit of Nem1p-Spo7p phosphatase holoenzyme
YCR026C	-1.919325	*NPP1*	Nucleotide pyrophosphatase/phosphodiesterase
YNL183C	-7.553306	*NPR1*	Protein kinase; stabilizes several plasma membrane amino acid transporters by antagonizing their ubiquitin-mediated degradation; phosphorylates Aly2p
YBL079W	-4.114137	*NUP170*	Subunit of inner ring of nuclear pore complex (NPC)
YLR335W	-1.087515	*NUP2*	Nucleoporin involved in nucleocytoplasmic transport
YNL099C	-6.775693	*OCA1*	Putative protein tyrosine phosphatase; required for cell cycle arrest in response to oxidative damage of DNA
YHL020C	-3.807469	*OPI1*	Transcriptional regulator of a variety of genes
YHR179W	-1.977984	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN)
YBR295W	-2.343834	*PCA1*	Cadmium transporting P-type ATPase
YOR360C	-2.566723	*PDE2*	High-affinity cyclic AMP phosphodiesterase
YJL053W	-1.548599	*PEP8*	Vacuolar protein component of the retromer
YNR045W	-1.501973	*PET494*	Mitochondrial translational activator specific for the COX3 mRNA
YGR077C	-6.251252	*PEX8*	Intraperoxisomal organizer of the peroxisomal import machinery
YGL025C	-3.299959	*PGD1*	Subunit of the RNA polymerase II mediator complex
YNL097C	-1.581213	*PHO23*	Component of the Rpd3L histone deacetylase complex
YBR296C	-2.489030	*PHO89*	Plasma membrane Na+/Pi cotransporter
YCR024C-A	-1.430018	*PMP1*	Regulatory subunit for the plasma membrane H(+)-ATPase Pma1p
YGL063W	-3.377646	*PUS2*	Mitochondrial tRNA:pseudouridine synthase
YLR204W	-1.487115	*QRI5*	Mitochondrial inner membrane protein
YNL098C	-1.626706	*RAS2*	GTP-binding protein; regulates nitrogen starvation response, sporulation, and filamentous growth

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YDR202C	-3.245415	*RAV2*	Subunit of RAVE complex (Rav1p, Rav2p, Skp1p)
YCR036W	-1.141943	*RBK1*	Putative ribokinase
YMR274C	-1.868680	*RCE1*	Type II CAAX prenyl protease
YLR248W	-1.473914	*RCK2*	Protein kinase involved in response to oxidative and osmotic stress
YNL022C	-1.088232	*RCM1*	rRNA m5C methyltransferase
YJL217W	-2.387224	*REE1*	Cytoplasmic protein involved in the regulation of enolase (ENO1)
YGL250W	-1.945067	*RMR1*	Protein required for meiotic recombination and gene conversion
YIL148W	-7.692187	*RPL40A*	Ubiquitin-ribosomal 60S subunit protein L40A fusion protein
YDL020C	-1.895597	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YCR045C	-4.039654	*RRT12*	Probable subtilisin-family protease
YMR263W	-5.638795	*SAP30*	Component of Rpd3L histone deacetylase complex
YLR268W	-1.042550	*SEC22*	R-SNARE protein; assembles into SNARE complex with Bet1p, Bos1p and Sed5p
YGR208W	-1.310761	*SER2*	Phosphoserine phosphatase of the phosphoglycerate pathway
YDL225W	-4.259420	*SHS1*	Component of the septin ring that is required for cytokinesis
YNL236W	-1.055804	*SIN4*	Subunit of the RNA polymerase II mediator complex; associates with core polymerase subunits to form the RNA polymerase II holoenzyme
YGR271W	-1.548723	*SLH1*	Putative RNA helicase related to Ski2p
YDR006C	-1.196946	*SOK1*	Protein of unknown function
YNL012W	-7.164570	*SPO1*	Meiosis-specific prospore protein
YDR392W	-3.049240	*SPT3*	Subunit of the SAGA and SAGA-like transcriptional regulatory complexes
YLR055C	-1.271974	*SPT8*	Subunit of the SAGA transcriptional regulatory complex
YAL005C	-1.276398	*SSA1*	ATPase involved in protein folding and NLS-directed nuclear transport
YLL024C	-2.974996	*SSA2*	HSP70 family ATP-binding protein; involved in protein folding, vacuolar import of proteins
YHL007C	-1.072217	*STE20*	Cdc42p-activated signal transducing kinase
YMR039C	-3.085640	*SUB1*	Transcriptional regulator; facilitates elongation through factors that modify RNAP II; role in peroxide resistance involving Rad2p

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YJL004C	-1.144495	*SYS1*	Integral membrane protein of the Golgi; required for targeting of the Arf-like GTPase Arl3p to the Golgi; multicopy suppressor of ypt6 null mutation
YEL048C	-2.595533	*TCA17*	Component of transport protein particle (TRAPP) complex II
YCR053W	-2.161817	*THR4*	Threonine synthase; conserved protein that catalyzes formation of threonine from O-phosphohomoserine
YOL018C	-4.504234	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi
YDR457W	-1.884953	*TOM1*	E3 ubiquitin ligase of the hect-domain class; has a role in mRNA export from the nucleus and may regulate transcriptional coactivators
YNL070W	-2.884919	*TOM7*	Component of the TOM (translocase of outer membrane) complex
YDR120C	-1.021030	*TRM1*	tRNA methyltransferase
YBR058C	-1.060611	*UBP14*	Ubiquitin-specific protease; specifically disassembles unanchored ubiquitin chains; involved in fructose-1,6-bisphosphatase (Fbp1p) degradation; similar to human isopeptidase T
YDL190C	-1.140738	*UFD2*	Ubiquitin chain assembly factor (E4)
YKL010C	-2.027636	*UFD4*	Ubiquitin-protein ligase (E3)
YBR006W	-1.706890	*UGA2*	Succinate semialdehyde dehydrogenase
YML021C	-1.184874	*UNG1*	Uracil-DNA glycosylase; required for repair of uracil in DNA formed by spontaneous cytosine deamination; efficiently excises uracil from single-stranded DNA in vivo; not required for strand-specific mismatch repair; cell-cycle regulated, expressed in late G1; localizes to mitochondria and nucleus
YJL130C	-1.476761	*URA2*	Bifunctional carbamoylphosphate synthetase/aspartate transcarbamylase
YJR049C	-1.970408	*UTR1*	ATP-NADH kinase; phosphorylates both NAD and NADH
YOR068C	-2.652752	*VAM10*	Protein involved in vacuole morphogenesis
YIL017C	-7.748956	*VID28*	GID Complex subunit, serves as adaptor for regulatory subunit Vid24p
YPL253C	-2.240708	*VIK1*	Protein that forms a kinesin-14 heterodimeric motor with Kar3p; localizes Kar3p at mitotic spindle poles
YJL154C	-1.749445	*VPS35*	Endosomal subunit of membrane-associated retromer complex
YLR360W	-3.246014	*VPS38*	Part of a Vps34p phosphatidylinositol 3-kinase complex

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOR043W	-4.119905	*WHI2*	Protein required for full activation of the general stress response
YHR134W	-5.972651	*WSS1*	Metalloprotease involved in DNA repair, removes DNA-protein crosslinks at stalled replication forks during replication of damaged DNA; localizes to a single spot on the nuclear periphery of mother cells but not daughters; interacts genetically with SMT3; activated by DNA binding
YML007W	-1.323049	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YHL009C	-1.588694	*YAP3*	Basic leucine zipper (bZIP) transcription factor
YBR216C	-1.635419	*YBP1*	Protein involved in cellular response to oxidative stress; required for oxidation of specific cysteine residues of transcription factor Yap1p
YLR020C	-2.303755	*YEH2*	Steryl ester hydrolase; catalyzes steryl ester hydrolysis at the plasma membrane
YER041W	-1.928270	*YEN1*	Holliday junction resolvase; promotes template switching during break-induced replication (BIR), causing non-reciprocal translocations (NRTs)
YFR007W	-2.042895	*YFH7*	Putative kinase with similarity to the PRK/URK/PANK kinase subfamily
YCR059C	-1.191655	*YIH1*	Negative regulator of eIF2 kinase Gcn2p
YJR142W	-2.037867	*YJR142W*	8-oxo-dGTP diphosphatase of the Nudix hydrolase family
YHL014C	-3.175322	*YLF2*	Protein of unknown function
YMR099C	-1.770541	*YMR099C*	Glucose-6-phosphate 1-epimerase (hexose-6-phosphate mutarotase)
YDR349C	-4.760298	*YPS7*	Putative GPI-anchored aspartic protease
YNL237W	-3.431980	*YTP1*	Protein of unknown function
YJL056C	-3.177184	*ZAP1*	Zinc-regulated transcription factor
YBL100C	-3.165495	-	Dubious open reading frame
YKR047W	-3.020315	-	Dubious open reading frame
YCR087C-A	-2.887457	-	Putative protein of unknown function
YCR061W	-2.764908	-	Protein of unknown function
YCL001W-A	-2.629919	-	Putative protein of unknown function
YJL193W	-2.591022	-	Putative protein of unknown function
YCR087W	-2.462803	-	Dubious open reading frame
YCR050C	-2.304166	-	Protein of unknown function
YIL077C	-2.247159	-	Putative protein of unknown function

2-MF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YOL050C	-2.061481	-	Dubious open reading frame
YGR164W	-1.978256	-	Dubious open reading frame
YDR537C	-1.966123	-	Dubious open reading frame
YDR431W	-1.944835	-	Dubious open reading frame
YDL183C	-1.872611	-	Protein of unknown function
YCR085W	-1.838758	-	Putative protein of unknown function
YGR035C	-1.800752	-	Putative protein of unknown function
YDL023C	-1.735333	-	Dubious open reading frame
YCR025C	-1.663695	-	Dubious open reading frame
YPR146C	-1.600185	-	Dubious open reading frame
YPR197C	-1.596537	-	Dubious open reading frame
YDR149C	-1.543165	-	Dubious open reading frame
YKL023W	-1.540832	-	Putative protein of unknown function
YNL013C	-1.386039	-	Dubious open reading frame
YPL182C	-1.372995	-	Dubious open reading frame
YCR049C	-1.325314	-	Dubious open reading frame
YAL004W	-1.297270	-	Dubious open reading frame
YNL115C	-1.247412	-	Putative protein of unknown function
YBR284W	-1.245773	-	Putative metallo-dependent hydrolase superfamily protein
YPR039W	-1.238496	-	Dubious open reading frame
YNL205C	-1.155408	-	Dubious open reading frame
YBR292C	-1.134037	-	Dubious open reading frame
YGR259C	-1.097749	-	Dubious open reading frame
YJL215C	-1.024835	-	Dubious open reading frame
YJR128W	-1.005754	-	Dubious open reading frame
YAL058C-A	-0.983666	-	Dubious open reading frame
YJR087W	-0.890169	-	Dubious open reading frame;
YOL085C	-0.841583	-	Putative protein of unknown function
YNR061C	-0.727857	-	Protein of unknown function
YGR022C	-0.663654	-	Dubious open reading frame

2-MF Sensitive across all time points

2-MF Sensitive across all time points		
ORF Name	**Gene Name**	**Deleted Gene Function**
YNL077W	*APJ1*	Chaperone with a role in SUMO-mediated protein degradation
YBR290W	*BSD2*	Heavy metal ion homeostasis protein
YOR125C	*CAT5*	Protein required for ubiquinone (Coenzyme Q) biosynthesis
YGR092W	*DBF2*	Ser/Thr kinase involved in transcription and stress response; functions as part of a network of genes in exit from mitosis
YPL265W	*DIP5*	Dicarboxylic amino acid permease; mediates high-affinity and high-capacity transport of L-glutamate and L-aspartate
YMR299C	*DYN3*	Dynein light intermediate chain (LIC)
YJR129C	*EFM3*	S-adenosylmethionine-dependent methyltransferase
YKR019C	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YFR024C-A	*LSB3*	Protein containing a C-terminal SH3 domain
YBR298C	*MAL31*	Maltose permease; high-affinity maltose transporter (alpha-glucoside transporter)
YNL307C	*MCK1*	Dual-specificity ser/thr and tyrosine protein kinase
YML062C	*MFT1*	Subunit of the THO complex
YCR026C	*NPP1*	Nucleotide pyrophosphatase/phosphodiesterase
YHR179W	*OYE2*	Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN)
YNL097C	*PHO23*	Component of the Rpd3L histone deacetylase complex
YNL098C	*RAS2*	GTP-binding protein
YDL020C	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YMR263W	*SAP30*	Component of Rpd3L histone deacetylase complex
YLR268W	*SEC22*	R-SNARE protein
YLR055C	*SPT8*	Subunit of the SAGA transcriptional regulatory complex
YHL007C	*STE20*	Cdc42p-activated signal transducing kinase
YDR457W	*TOM1*	E3 ubiquitin ligase of the hect-domain class
YML007W	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YCR059C	*YIH1*	Negative regulator of eIF2 kinase Gcn2p
YMR099C	*YMR099C*	Glucose-6-phosphate 1-epimerase (hexose-6-phosphate mutarotase)
YNR061C	-	Protein of unknown function
YCR061W	-	Protein of unknown function

Appendix 4: Mutants with altered growth in 2-EF

2-EF 5G Resistant

2-EF 5G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YGR041W	2.628554	*BUD9*	Protein involved in bud-site selection
YNL278W	1.488499	*CAF120*	Part of the CCR4-NOT transcriptional regulatory complex
YJL099W	3.400330	*CHS6*	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family
YGL002W	1.730549	*ERP6*	Member of the p24 family involved in ER to Golgi transport
YJR030C	0.979189	*RBH2*	Putative protein of unknown function
YKL020C	1.931782	*SPT23*	ER membrane protein involved in regulation of OLE1 transcription
YBR044C	2.444186	*TCM62*	Protein involved in assembly of the succinate dehydrogenase complex
YOR132W	1.803843	*VPS17*	Subunit of the membrane-associated retromer complex
YPR059C	1.261990	-	Dubious open reading frame
YMR310C	1.330129	-	Putative protein of unknown function
YLR269C	1.332697	-	Dubious open reading frame
YOL163W	1.370128	-	Putative protein of unknown function
YOR364W	2.496378	-	Dubious open reading frame

2-EF 15G Resistant

2-EF 15G Resistant Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YHR094C	1.021111	*HXT1*	Low-affinity glucose transporter of the major facilitator superfamily
YOL122C	0.859951	*SMF1*	Divalent metal ion transporter; broad specificity for divalent and trivalent metals
YDR109C	1.342878	-	Putative kinase
YLR012C	0.880762	-	Putative protein of unknown function
YLR434C	1.246196	-	Dubious open reading frame
YEL023C	1.140432	-	Putative protein of unknown function
YDL086W	1.221329	-	Putative protein of unknown function

2-EF 5G Sensitive

2-EF 5G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YLR131C	-2.153724	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YNL051W	-3.386425	*COG5*	Component of the conserved oligomeric Golgi complex
YNL041C	-2.093056	*COG6*	Component of the conserved oligomeric Golgi complex
YIL002C	-1.146930	*INP51*	Phosphatidylinositol 4,5-bisphosphate 5-phosphatase
YKR019C	-2.351672	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YEL037C	-0.874186	*RAD23*	Protein with ubiquitin-like N terminus; subunit of Nuclear Excision Repair Factor 2 (NEF2) with Rad4p that binds damaged DNA; enhances protein deglycosylation activity of Png1p; also involved, with Rad4p, in ubiquitylated protein turnover; Rad4p-Rad23p heterodimer binds to promoters of DNA damage response genes to repress their transcription in the absence of DNA damage
YDL020C	-1.921801	*RPN4*	Transcription factor that stimulates expression of proteasome genes upon DNA replication stress
YHR206W	-1.578703	*SKN7*	Nuclear response regulator and transcription factor
YDR463W	-3.244005	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YJL004C	-2.466030	*SYS1*	Integral membrane protein of the Golgi
YBR069C	-1.752176	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YDL080C	-1.397818	*THI3*	Regulatory protein that binds Pdc2p and Thi2p transcription factors
YML007W	-1.187031	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YNL237W	-0.994653	*YTP1*	Protein of unknown function

2-EF 10G Sensitive

2-EF 10G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YLR131C	-2.949282	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YDR101C	-1.767640	*ARX1*	Nuclear export factor for the ribosomal pre-60S subunit
YNL041C	-7.168183	*COG6*	Component of the conserved oligomeric Golgi complex
YML071C	-5.526670	*COG8*	Component of the conserved oligomeric Golgi complex
YCR071C	-2.437157	*IMG2*	Mitochondrial ribosomal protein of the large subunit
YKR019C	-4.587350	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YAL024C	-4.947588	*LTE1*	Protein similar to GDP/GTP exchange factors
YEL053C	-1.468892	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YHR206W	-1.802560	*SKN7*	Nuclear response regulator and transcription factor
YDR463W	-7.957441	*STP1*	Transcription factor; contains a N-terminal regulatory motif (RI) that acts as a cytoplasmic retention determinant and as an Asi dependent degron in the nucleus
YJL004C	-3.690139	*SYS1*	Integral membrane protein of the Golgi
YBR069C	-3.626629	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YML007W	-1.662550	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YMR262W	-2.355676	-	Protein of unknown function
YMR141C	-1.376591	-	Putative protein of unknown function
YGR035C	-1.101691	-	Putative protein of unknown function

2-EF 15G Sensitive

2-EF 15G Sensitive Strains			
Deleted ORF Name	**Log_2 Fold Change**	**Deleted Gene Name**	**Deleted Gene Function**
YLR131C	-7.399615	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YER017C	-6.230796	*AFG3*	Mitochondrial inner membrane m-AAA protease component

2-EF 15G Sensitive Strains (continued)			
Deleted ORF Name	**Deleted ORF Name**	**Deleted ORF Name**	**Deleted ORF Name**
YER167W	-3.764460	*BCK2*	Serine/threonine-rich protein involved in PKC1 signaling pathway
YAR014C	-2.708735	*BUD14*	Protein involved in bud-site selection
YGL027C	-1.138990	*CWH41*	Processing alpha glucosidase I; ER type II integral membrane N-glycoprotein
YAL026C	-2.130574	*DRS2*	Trans-golgi network aminophospholipid translocase (flippase)
YMR276W	-1.795741	*DSK2*	Nuclear-enriched ubiquitin-like polyubiquitin-binding protein
YCR034W	-1.240562	*ELO2*	Fatty acid elongase, involved in sphingolipid biosynthesis
YGL054C	-2.875402	*ERV14*	COPII-coated vesicle protein
YIL065C	-0.931000	*FIS1*	Protein involved in mitochondrial fission and peroxisome abundance
YGL020C	-6.128761	*GET1*	Subunit of the GET complex
YOR070C	-1.758608	*GYP1*	Cis-golgi GTPase-activating protein (GAP) for yeast Rabs
YOL081W	-2.700860	*IRA2*	GTPase-activating protein
YKR019C	-8.159027	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YEL053C	-2.250082	*MAK10*	Non-catalytic subunit of the NatC N-terminal acetyltransferase
YOL116W	-1.239047	*MSN1*	Transcriptional activator; involved in regulation of invertase and glucoamylase expression, invasive growth and pseudohyphal differentiation, iron uptake, chromium accumulation, and response to osmotic stress
YNR049C	-1.784849	*MSO1*	Lipid-interacting protein in SNARE complex assembly machinery
YMR080C	-2.193308	*NAM7*	ATP-dependent RNA helicase of the SFI superfamily
YHR004C	-1.717609	*NEM1*	Probable catalytic subunit of Nem1p-Spo7p phosphatase holoenzyme
YBL079W	-3.181725	*NUP170*	Subunit of inner ring of nuclear pore complex (NPC)
YCR079W	-1.047813	*PTC6*	Mitochondrial type 2C protein phosphatase (PP2C)
YMR274C	-2.707464	*RCE1*	Type II CAAX prenyl protease
YDL020C	-3.434714	*RPN4*	Transcription factor that stimulates expression of proteasome genes
YCR045C	-1.065848	*RRT12*	Probable subtilisin-family protease
YJL047C	-2.464907	*RTT101*	Cullin subunit of a Roc1p-dependent E3 ubiquitin ligase complex
YHR206W	-2.748666	*SKN7*	Nuclear response regulator and transcription factor
YOR327C	-1.598279	*SNC2*	Vesicle membrane receptor protein (v-SNARE)

YJL036W	-0.982127	*SNX4*	Sorting nexin; involved in retrieval of late-Golgi SNAREs from post-Golgi endosomes to the trans-Golgi network and in cytoplasm to vacuole transport
YJL004C	-7.359825	*SYS1*	Integral membrane protein of the Golgi
YBR069C	-4.552861	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine
YCR053W	-2.018128	*THR4*	Threonine synthase; conserved protein that catalyzes formation of threonine from O-phosphohomoserine
YOL018C	-4.553734	*TLG2*	Syntaxin-like t-SNARE; forms a complex with Tlg1p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi
YNL300W	-1.113109	*TOS6*	Glycosylphosphatidylinositol-dependent cell wall protein
YDR120C	-1.530404	*TRM1*	tRNA methyltransferase
YGR166W	-1.730845	*TRS65*	Component of transport protein particle (TRAPP) complex II
YDR108W	-2.035670	*TRS85*	Component of transport protein particle (TRAPP) complex III
YJR049C	-1.923595	*UTR1*	ATP-NADH kinase; phosphorylates both NAD and NADH
YIL017C	-1.511648	*VID28*	GID Complex subunit, serves as adaptor for regulatory subunit Vid24p
YLL040C	-1.423604	*VPS13*	Protein involved in prospore membrane morphogenesis
YOR043W	-1.012060	*WHI2*	Protein required for full activation of the general stress response
YHR134W	-5.585626	*WSS1*	Metalloprotease involved in DNA repair, removes DNA-protein crosslinks at stalled replication forks during replication of damaged DNA; localizes to a single spot on the nuclear periphery of mother cells but not daughters; interacts genetically with SMT3; activated by DNA binding
YML007W	-2.302138	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress
YDR349C	-1.801126	*YPS7*	Putative GPI-anchored aspartic protease
YJL056C	-1.455325	*ZAP1*	Zinc-regulated transcription factor
YOR364W	-2.583798	-	Dubious open reading frame
YPR050C	-1.985367	-	Dubious open reading frame
YJL022W	-1.346337	-	Dubious open reading frame
YCR050C	-1.264936	-	Protein of unknown function
YNL205C	-1.047102	-	Dubious open reading frame

2-EF Sensitive across all time points

2-EF Sensitive Strains Across All Time Points		
Deleted ORF Name	**Deleted Gene Name**	**Deleted Gene Function**
YLR131C	*ACE2*	Transcription factor required for septum destruction after cytokinesis
YKR019C	*IRS4*	EH domain-containing protein; involved in regulating phosphatidylinositol 4,5-bisphosphate levels and autophagy
YHR206W	*SKN7*	Nuclear response regulator and transcription factor
YJL004C	*SYS1*	Integral membrane protein of the Golgi
YBR069C	*TAT1*	Amino acid transporter for valine, leucine, isoleucine, and tyrosine; low-affinity tryptophan and histidine transporter
YML007W	*YAP1*	Basic leucine zipper (bZIP) transcription factor; required for oxidative stress tolerance; relative distribution to the nucleus increases upon DNA replication stress

Appendix 5: Yeast Functional Toxicogenomic Assay Sequencing Design

Lane1				
Plate	**Sample Code/ Tube Label**	**Sample Name and Replicate #**	**Sample Type**	**Index**
UP	1A2Up	DMSO 5G - 1	Control 1	AACCGTGT
UP	1A3Up	DMSO 5G - 2	Control 1	AAGATTGC
UP	1A4Up	DMSO 5G - 3	Control 1	AAGCGGTC
UP	1A5Up	DMSO 10G - 1	Control 2	AATCACAC
UP	1A7Up	DMSO 10G - 2	Control 2	ACAGTGCA
UP	1A8Up	DMSO 10G - 3	Control 2	ACCGTTAT
UP	1B3Up	DMSO 15 G - 1	Control 3	AGTAGTGG
UP	1B4Up	DMSO 15 G - 2	Control 3	AGTTGCTA
UP	1B5Up	DMSO 15 G - 3	Control 3	ATATAGGA
UP	1B6Up	2-MF 5G - 1	Treatment 1	ATCCTATT
UP	1B7Up	2-MF 5G - 2	Treatment 1	ATCGCCAG
UP	1B9Up	2-MF 5G - 3	Treatment 1	ATGCATCC
UP	1C2Up	2-MF 10G - 1	Treatment 2	CACGTGTT
UP	1C3Up	2-MF 10G - 2	Treatment 2	CAGGAGGC
UP	1C4Up	2-MF 10G - 3	Treatment 2	CATTCCAA
UP	1C5Up	2-MF 15G - 1	Treatment 3	CCAGCACG
UP	1C6Up	2-MF 15G - 2	Treatment 3	CCATACAC
UP	1C7Up	2-MF 15G - 3	Treatment 3	CCGGATAG
UP	1C8Up	2-EF 5G - 1	Treatment 1	CCGTCTGA
UP	1C9Up	2-EF 5G - 2	Treatment 1	CCTACAAC
UP	1D2Up	2-EF 5G - 3	Treatment 1	CGCTTCTG
UP	1D3Up	2-EF 10G - 1	Treatment 2	CGGACGTG
UP	1D4Up	2-EF 10G - 2	Treatment 2	CGGTTGAT
UP	1D5Up	2-EF 10G - 3	Treatment 2	CGTCGGCT
UP	1D6Up	2-EF 15G - 1	Treatment 3	CTAGATTC
UP	1D7Up	2-EF 15G - 2	Treatment 3	CTAGTCAT
UP	1D9Up	2-EF 15G - 3	Treatment 3	CTTAAGAT
UP	1E1Up	2,3-DMF 5G - 1	Treatment 1	GACGTCAA
UP	1E2Up	2,3-DMF 5G - 2	Treatment 1	GAGAACTC
UP	1E3Up	2,3-DMF 5G - 3	Treatment 1	GAGTTAAC
UP	1E4Up	2,3-DMF 10G - 1	Treatment 2	GATCCAGC
UP	1E5Up	2,3-DMF 10G - 2	Treatment 2	GATGGAAT
UP	1E6Up	2,3-DMF 10G - 3	Treatment 2	GCAAGTAG
UP	1E8Up	2,3-DMF 15G - 1	Treatment 3	GCGGCGAA
UP	1E9Up	2,3-DMF 15G - 2	Treatment 3	GCGTTTCG
UP	1F1Up	2,3-DMF 15G - 3	Treatment 3	GGATATGG
UP	1F2Up	2,5-DMF 5G - 1	Treatment 1	GGCAGACG
UP	1F3Up	2,5-DMF 5G - 2	Treatment 1	GGCGAGGA
UP	1F4Up	2,5-DMF 5G - 3	Treatment 1	GGTCCTTG
UP	1F5Up	2,5-DMF 10G - 1	Treatment 2	GGTCTGAC
UP	1F6Up	2,5-DMF 10G - 2	Treatment 2	GTACTTGC
UP	1G1Up	2,5-DMF 10G - 3	Treatment 2	GTTTCACT
UP	1G2Up	2,5-DMF 15G - 1	Treatment 3	TACGAATC
UP	1G3Up	2,5-DMF 15G - 2	Treatment 3	TACTGCGC
UP	1H4Up	2,5-DMF 15G - 3	Treatment 3	TGATCCGA
	Total Number of Samples in this Library: 45			

Lane2

Plate	Sample Code/ Tube Label	Sample Name and Replicate #	Sample Type	Index
DWN	1A2DWN	DMSO 5G - 1	Control 1	AACCGTGT
DWN	1A3DWN	DMSO 5G - 2	Control 1	AAGATTGC
DWN	1A4DWN	DMSO 5G - 3	Control 1	AAGCGGTC
DWN	1A5DWN	DMSO 10G - 1	Control 2	AATCACAC
DWN	1A7DWN	DMSO 10G - 2	Control 2	ACAGTGCA
DWN	1A8DWN	DMSO 10G - 3	Control 2	ACCGTTAT
DWN	1B3DWN	DMSO 15 G - 1	Control 3	AGTAGTGG
DWN	1B4DWN	DMSO 15 G - 2	Control 3	AGTTGCTA
DWN	1B5DWN	DMSO 15 G - 3	Control 3	ATATAGGA
DWN	1B6DWN	2-MF 5G - 1	Treatment 1	ATCCTATT
DWN	1B7DWN	2-MF 5G - 2	Treatment 1	ATCGCCAG
DWN	1B9DWN	2-MF 5G - 3	Treatment 1	ATGCATCC
DWN	1C2DWN	2-MF 10G - 1	Treatment 2	CACGTGTT
DWN	1C3DWN	2-MF 10G - 2	Treatment 2	CAGGAGGC
DWN	1C4DWN	2-MF 10G - 3	Treatment 2	CATTCCAA
DWN	1C5DWN	2-MF 15G - 1	Treatment 3	CCAGCACG
DWN	1C6DWN	2-MF 15G - 2	Treatment 3	CCATACAC
DWN	1C7DWN	2-MF 15G - 3	Treatment 3	CCGGATAG
DWN	1C8DWN	2-EF 5G - 1	Treatment 1	CCGTCTGA
DWN	1C9DWN	2-EF 5G - 2	Treatment 1	CCTACAAC
DWN	1D2DWN	2-EF 5G - 3	Treatment 1	CGCTTCTG
DWN	1D3DWN	2-EF 10G - 1	Treatment 2	CGGACGTG
DWN	1D4DWN	2-EF 10G - 2	Treatment 2	CGGTTGAT
DWN	1D5DWN	2-EF 10G - 3	Treatment 2	CGTCGGCT
DWN	1D6DWN	2-EF 15G - 1	Treatment 3	CTAGATTC
DWN	1D7DWN	2-EF 15G - 2	Treatment 3	CTAGTCAT
DWN	1D9DWN	2-EF 15G - 3	Treatment 3	CTTAAGAT
DWN	1E1DWN	2,3-DMF 5G - 1	Treatment 1	GACGTCAA
DWN	1E2DWN	2,3-DMF 5G - 2	Treatment 1	GAGAACTC
DWN	1E3DWN	2,3-DMF 5G - 3	Treatment 1	GAGTTAAC
DWN	1E4DWN	2,3-DMF 10G - 1	Treatment 2	GATCCAGC
DWN	1E5DWN	2,3-DMF 10G - 2	Treatment 2	GATGGAAT
DWN	1E6DWN	2,3-DMF 10G - 3	Treatment 2	GCAAGTAG
DWN	1E8DWN	2,3-DMF 15G - 1	Treatment 3	GCGGCGAA
DWN	1E9DWN	2,3-DMF 15G - 2	Treatment 3	GCGTTTCG
DWN	1F1DWN	2,3-DMF 15G - 3	Treatment 3	GGATATGG
DWN	1F2DWN	2,5-DMF 5G - 1	Treatment 1	GGCAGACG
DWN	1F3DWN	2,5-DMF 5G - 2	Treatment 1	GGCGAGGA
DWN	1F4DWN	2,5-DMF 5G - 3	Treatment 1	GGTCCTTG
DWN	1F5DWN	2,5-DMF 10G - 1	Treatment 2	GGTCTGAC
DWN	1F6DWN	2,5-DMF 10G - 2	Treatment 2	GTACTTGC
DWN	1G1DWN	2,5-DMF 10G - 3	Treatment 2	GTTTCACT
DWN	1G2DWN	2,5-DMF 15G - 1	Treatment 3	TACGAATC
DWN	1G3DWN	2,5-DMF 15G - 2	Treatment 3	TACTGCGC
DWN	1H4DWN	2,5-DMF 15G - 3	Treatment 3	TGATCCGA

Total Number of Samples in this Library: 45