# Reporter-based synthetic genetic analysis of budding yeast reveals novel MMS-induced effectors of the *RNR3* promoter

Nada Elnour

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
in partial fulfillment of the requirements for the
MSc Degree in
Cellular and Molecular Medicine

Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa

# Abstract

The DNA damage response is a cell-wide response that coordinates repair and cell-cycle progression. Crucial to fidelity of genetic propagation, survival, and apoptosis, dysfunctions in the response are at the root of genome instability syndromes and cancer predisposition in mammalian cells. Within the response lie hubs of coordination, called checkpoints, whose members and organization are ubiquitous amongst eukaryotes. The high conservation of these checkpoints enable the study of their dynamics by proxy via simpler model organisms. We use the budding yeast, *Saccharomyces cerevisiae*, to study the replication and DNA damage checkpoints — both implicated in DNA damage repair. Using a yEGFP reporter driven by the *RNR3* promoter and reporter-based synthetic genetic array analysis, we created a detector of potential checkpoint activation in response to two doses of MMS, 0.015% and 0.060% (v/v). The high-throughput screens and differential epistasis miniarray analyses (EMAPs) yield unanticipated involvement of oxidative stress response, ribosomal biogenesis, and chromatin remodelling genes.

# Acknowledgments

I thank members of the Kærn lab for their companionship.

Thank you Lioudmila Tepliakova for your help and patience with the seven-to-eleven-hour protocols. Without you, I imagine the strains would tauntingly dance to the rhythm of multichannel pipettes. More importantly, thank you for your guidance and mentorship throughout these two years.

Also, thank you Dr. Kærn for giving me the opportunity to explore research and not kicking me out of the lab.

I thank:
Dr. Theodore Perkins for your help with the statistical challenges of high-throughput data analysis;
Dr. Kristin Baetz for providing the Singer RoToR and for your insights during the committee meetings; and
Dr.Vera Tang for your tutelage on the black box that is the flow cytometer.

Finally, I thank my parents and sisters for their support and patience with my journey riddled with highs, lows, and recovery. To you, I am indebted the most.

Even this clear sky is
Complex in its simplicity;
Where it begins and ends,
Our minds only profess to see.

# Contents

# List of Figures

# List of Tables

## List of Abbreviations

| | |
|---|---|
| 4NQO | 4-nitroquinoline 1-oxide |
| ADC | Analogue-to-Digital Converter |
| ARG | Arginine |
| AUC | Area Under Curve |
| BER | Base Excision Repair |
| BFP | Blue Fluorescent Protein |
| BP | Biological Process |
| CC | Cellular Component |
| CDF | Cumulative Distribution Function |
| DAmP | Decreased Abundance by mRNA disruption |
| DDC | DNA Damage Checkpoint Pathway |
| DDR | DNA Damage Response |
| DMA | Deletion Mutant Array |
| DNA | Deoxyribonucleic Acid |
| DNA-PK | DNA-dependent Protein Kinase |
| dRP | deoxyribose Phosphate |
| DSB | Double-stranded Break |
| EMAP | Epistasis Miniarray Profile |
| FSC | Forward Scatter Channel |
| GFP | Green Fluorescent Protein |
| GO | Gene Ontology |
| HAT | Histone Acetyltransferase |
| HDAC | Histone Deacetylation Complex |
| HR | Homologous Recombination |
| HTFM | High-throughput Fluorescence Microscopy |
| kanMX | Kanamycin cassette |
| MAPK | Mitogen-activated Protein Kinase |
| MF | Molecular Function |
| MMR | Mismatch Repair |
| MMS | Methyl Methanesulfonate |
| natMX | Nourseothricin (ClonNAT) cassette |
| NER | Nucleotide Excision Repair |
| NHEJ | Non-homologous End-Joining |
| PCNA | Proliferating Cell Nuclear Antigen |
| PCR | Polymerase Chain Reaction |
| PDF | Probabiliy Density Function |
| PMT | Photomultiplier Tube |

| | |
|---|---|
| Pol | Polymerase |
| RMSE | Root-Mean-Square-Srror |
| RNA | Ribonucleic Acid |
| RNR | Ribonucleotide Reductase |
| ROS | Reactive Oxygen Species |
| RPA | Replication Protein A |
| R-SGA | Reporter-based SGA |
| SAGA | Spt-Ada-Gcn5-Acetylatransferase complex |
| SGA | Synthetic Genetic Array Analysis |
| SM | Synthetic Media |
| SPO | Sporulation |
| SSB | Single-stranded Break |
| SSC | Side Scatter Channel |
| TBP | TATA-binding Protein |
| TE | Tris-EDTA buffer |
| URA | Uracil |
| WT | neutral mutants |
| YPD | Yeast extract Peptone Dextrose |

# 1 | Introduction

Cell survival is dependent on the cell's ability to maintain its structural and functional integrity in the face of a changing intracellular and extracellular environment. To do so, the cell degrades and replaces damaged structures, and regulates the production of proteins that buffer the intracellular environment to maintain chemical homeostasis. At the heart of these strategies sits *ab initio* replacement, wherein genes are expressed to replace damaged macromolecules and modulate the pre-existing proteome and metabolome for repair [38]. Because *ab initio* replacement involves gene expression, the maintenance of DNA integrity within a reasonable range is imperative for maintaining homeostasis.

The consequences of poor genetic integrity are best illustrated in diseases of dysfunctional DNA maintenance. Dysfunctions in DNA damage repair pathways often result in higher mutation rates — the culmination of which is a mutator phenotype that predisposes cells to unregulated proliferation — malignancy tran sitions, in mammals. These mutations often form genetic signatures that are enriched in genome instability syndromes (e.g. ataxia telangiectasia and xeroderma pigmentosa), neurodegenerative diseases (e.g. microcephaly and cerebellar ataxia), and cancer (e.g. broadly lymphomas, melanoma, carcinomas) [6, 11, 52, 77].

# 2 | DNA Damage and Repair

This chapter will present the importance of upstream checkpoint members in mobilizing a multitude of pathways in response to DNA damage. Checkpoint members act as hubs of coordination at which DNA damage integrity sensors and effectors can communicate. Their importance is highlighted in that the disintegration of the checkpoint leads to the cell-cycle, growth, and proliferation abnormalities common in genome instability syndromes. Because the checkpoints in higher eukaryotes are complex, we often need a simpler model. In this chapter, we present Saccharomyces cerevisiae as one containing the highly conserved orthologues of the replication and DNA damage checkpoints members commonly found in mammals. Furthermore, the presence of a relatively simple and well-characterized transcriptional system ($RNR3$) downstream of the checkpoints allows us to use its promoter ($RNR3pr$) dynamics as a detector system of checkpoint activation in the absence of a gene.

## 2.1 DNA Damage

DNA damage can be endogenous (from intracellular reactions) or exogenous (from extracellular sources). Common endogenous sources of lesions include deamination, spontaneous mutations (occurs in $10^4$ bases/cell/day in *Homo sapiens* [47]) and replication errors (a result of inherent error in misincorporation and proofreading by replication machinery), and metabolism by-products, such as reactive oxygen species (ROS).

Deamination refers to the process of amine hydrolysis from nucleobases, and can lead to mismatched pairing if not repaired. For example, 5-methylcytosine deamination produces

thymine, inceasing the risk of a $C \bullet G \rightarrow T \bullet A$ transition. Similarly, cytosine deamination creates uracil, leading to $C \bullet G \rightarrow T \bullet A$ transition either via pairing with adenine or via creating an ambiguous template in the event of uracil excision. Deamination at CpG islands in humans and the subsequent transitions represent the two most common mutagenic signatures observed in $> 25$ different cancers [77]. Spontaneous mutations and replication errors can also result in ambiguous pairing and point mutations because of nucleobase loss and mismatch insertions. Finally, ROS can also lead to $C \bullet G \rightarrow T \bullet A$ transitions via the oxidation of guanine into 8-oxo-2'-deoxyguanosine.

Exogenous agents include ROS and free radicals, in addition to non-ionizing radiation (UV), ionizing radiation (X- and $\gamma$-radiation), radiomimetic drugs, alkylating agents, intercalating agents, and psoralens to name a few.
Ionizing radiation and radiomimetic drugs produce free radicals that introduce double-stranded breaks (DSBs) in DNA, while non-ionizing radiation mediates the covalent bonding of two adjacent pyrimidines forming dipyrimidine. Dipyrimidine formation is susceptible to $CC \bullet GG \rightarrow TT \bullet AA$ transitions, a mutation signature that makes up to 25% of cutaneous cancers [6]. Finally, interacalating and alkylating agents can result in overlapping signatures. Both can form adducts with nucleotides depending on the agent, e.g. temozolomide is thought to methylate $O^6$ of guanine [47], benzo[$\alpha$]pyrene forms a bulky adduct with guanine [77], and aristolochic acid forms an adduct with adenine. Interacalating agents often lead to insertions, deletions, and frameshift mutations if not excised. Alkylating agents, on the other hand, can lead to point base substitutions, single stranded breaks (SSBs), or DSBs (rev.in [47]).

## 2.2 Repair in Mammals

Mutational outcomes appear to fall into categories regardless of the agents' mechanism of action: depurination (and less frequently, depyrimidination), point mutations or mismatches, single-stranded breaks (SSBs), double-stranded breaks (DSBs), pyrimidine dimers, interstrand linkages, or bulky adducts. These broader categories are what activate some

of the five repair pathways ubiquitous to eukaryotes: mismatch repair (MMR), nucleotide excision repair (NER), base excision repair (BER), non-homologous end-joining (NHEJ), or homologous recombination (HR).

Small mutation loads are resolved quickly. Replication errors are repaired by MMR, whose members reduce the rate of error by 100-fold (rev. in [62]). Pair mismatches are thought to produce a kink in the helix, and are recognized by the sensors, MSH2 and MSH6 [38, 62]. These sensors enhance the recruitment of the endonuclease PMS2, exonuclease 1 (EXO1), MLH1, and the proliferating cell nuclear antigen (PCNA) clamp [38, 62, 70]. The clamp anchors PMS2, EXO1 and MLH1 so that the latter can cleave at points flanking the mismatched nucleotide. PCNA, finally, interacts with Polymerase (Pol) $\delta$ [11, 70, 92] to replace the missing bases. This theme of damage excision and replacement extends to BER and NER, as their names imply, with the main difference between the repair mechanisms is the length of the removed segment.

In NER, damage can be detected during transcription by Pol II-CSA-CSB [49, 70], or during global genome repair by the DNA repair complementing (XP-C)-RAD23B complex [70]. Despite the different detection pathways, the signal is believed to converge downstream as transcription factor II subunit H ($TF_{II}$ H)'s XPB and XPD helicases unravel a 27–29-bp long segment around the lesion [62]. Then, endonucleases XPG and ERCC1/XPF nick the 3' and 5' end of the unraveled segment, while replication protein A (RPA) coats and protects the exposed single stranded DNA (ssDNA) ( [47], rev. in [16]). Again, the PCNA clamp was implicated in this process and is thought to latch onto the ssDNA to recruit any high-fidelity DNA polymerase for synthesis. Generally, NER is activated in response to bulky adducts and pyrimidine dimers.

Base excision repair, on the other hand, excises one base at a time, and is activated in response to oxidative damage, depurination, SSBs, and alkylations. Damaged bases are depurinated (or dyprimidinated) by a nucleotide-specific glycosylase [49]. The abasic

site is first incised by APEX1, a DNA lyase, followed by 5'-deoxyribose phosphate (dRP) lyase-mediated removal of the sugar backbone [49,70]. Finally, Pol $\beta$ replaces the nucleotide and the ligase 3-XRCC1 complex ligates the new moiety to the rest of the strand (rev. in [16, 70]).

The last two repair pathways do not rely on excision. Instead, they repair via forced (NHEJ) or pre-existing (HR) homology in response to DSBs and interstrand crosslinks. They are thought to resolve damage done by ionizing radiation and certain carcinogens. Interestingly, their predominance is complementary and cell-cycle dependent [23]. The current model for homology repair places NHEJ's and HR's predominance as a gradient beginning in G1 (NHEJ solely) to G2 (HR solely), with equal probability of repair by either at mid-S phase.

NHEJ does not need a sister chromatid or second copy of a gene to operate. Double stranded breaks usually leave blunt break ends with little overhangs. The KU70/KU80 complex sense these ends, and recruit DNA-dependent protein kinase (DNA-PK) and Artemis for resection (rev. in [95]). Resection involves the removal of a few bases from one of the antiparallel strands to create ssDNA microhomologies. With microhomology formation, the LIG4-XRCC1 complex forces the ligation of the two segments while DNA Pol $\lambda/\mu$ fills in the missing bases [96, rev. in [95]]. Because of forced ligation of resected ends and lack of a second copy to serve as a template, NHEJ-repaired DNA is prone to translocations and deletions.

Unlike NHEJ, HR begins operating in mid-S phase until just after the G2/M checkpoint in human cells — i.e. in all cell-cycle stages containing two copies of the genome. Homologous recombination repair can be accomplished in one of two ways: one-end, or two-end strand invasion (rev. in [95]). During one-end invasion, a strand of the intact sister chromatid invades and serves as a template for damaged sister chromatid repair. Following synthesis, the invaded strand is released and microhomology–mediated end-joining is believed to occur à la NHEJ. Two-end strand invasion, on the other hand, involves invasion by one strand of each sister chromatid, damaged and intact. Again, after homologous

pairing, DNA is synthesized resulting in a heteroduplex structure known as a Holliday junction. At this stage, the sister chromatids can either crossover — a process termed resolution and mediated by resolvases — or dissociate without crossover — a process called dissolution and involves the release of the original strands with their synthesized homologies. Two-end invasion HR is therefore error-free.

## 2.3 The DNA Damage Response

Under a large mutation load, the cell relies on successful coordination of its cell cycle with damage repair. Continuation of the cell cycle in the presence of irreparable damage, especially adducts, crosslinks, and strand breaks, usually leads to a cell-fate decision to die. If apoptosis fails, however, cells cycling with the excessive aberrations can transition to malignancy — a fate detrimental to organisms that rely on the proper operation of cellular communities. Cells with damaged genomes do not automatically become rogue, however. If the large mutation load is still within a tolerable range, the DNA Damage Response (DDR) is activated by the kinases ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad3-related protein (ATR) [11]. The DDR is comprised of sensors, mediators, and effectors that collectively make up checkpoints relaying DNA health status to the rest of cellular processes[11].

With the onset of extensive adduct formation or nucleotide lesions, BER and NER processing can expose RPA-coated ssDNA foci [70]. Both ATR and its regulatory subunit ATRIP interact with RPA chains [70]. ATR-ATRIP phosphorylates the 9-1-1 clamp (RAD9-HUS1-RAD1) and RAD17-RFC, which in turn recruit the mediators claspin, RAD9, BRCA1, 53BP1, and MDC1 [11,70,rev. in [95]]. With the help of chromatin remodeling complexes, the mediators lead to the activation of the effector CHK1.

DSBs, on the other hand, are sensed by the MRN (MRE11-RAD50-NBS1) complex that recruits an ATM dimer [11]. Phosphorylation of ATM transduces the damage signal to the mediators 53BP1 and MDC1 [11]. Concurrently, ATM and ATR coordinate cell-

cycle progression via a multitude of downstream targets. For example, RAD9 can arrest cells in early G1; NBS1, SMC1, and BRCA1 at early, mid-, and late S phase, respectively [91]. ATR can also arrest the cell at early G1 and the G2/M checkpoint in humans via RAD17. Therefore, ATM and ATR are essential hubs coordinating DNA integrity status with cell-cycle progression, repair, and cell-fate decisions. They, along with the mediators, comprise two checkpoints: the replication checkpoint, downstream of which is CHK2, and the DNA damage checkpoint (DDC) with CHK1 [11, 70, 91, 96].

What is truly fascinating is the conservation of organization and function of the DDR across prokaryotes and eukaryotes (Table 2.1). While functional conservation may not apply to all members of the DDR, that which remains enables the use of simpler organisms as proxy models. The benefit of translatability of insights to mammalian cells is crucial considering the importance of DDR members to genome instability syndromes. To this end, we chose the replication and DNA damage checkpoints in the budding yeast, *Saccharomyces cerevisiae*, as a model of mammalian checkpoint dynamics in response to the carcinogen, methyl methanesulfonate (MMS).

**Table 2.1:** Homology of DDR members ([42, 57, 88]; rev. in [53]).

| Category | Mammals | *S. cervisiae* |
|---|---|---|
| Sensors | RPA | Rfa |
| | MRN(MRE11-RAD50-NBS1) | MRX(Mre11-Rad50-Xrs2) |
| | Rad9-Hus1-Rad1 | Ddc1-Mec3-Rad17 |
| | Rad17 | Rad24 |
| | Rfc2-5 | Rfc2-5 |
| Transducers | ATM | Tel1 |
| | ATR | Mec1 |
| | ATRIP | Ddc2 |
| Mediators | BRCA1, MDC1, 53BP1 | Rad9 |
| | TopBP1 | Dpb11 |
| | Claspin | Mrc1 |
| Effectors | CHK1 | Chk1 |
| | CHK2 | Rad53 |

## 2.4 DDR in Budding Yeast

The DDR in *S. cerevisiae* is organized much the same as it is in mammalian cells, with sensors, transducers, mediators and effectors. At the core of the checkpoint pathways are Mec1p (ATR) or Tel1p (ATM), and Rad53p (CHK2) as sensors and transducers of the damage signal, respectively [53, 56]. A major difference between *S. cerevisiae* and mammalian cells is that the former does not need to repair its damaged genome fully before proceeding to replication. In fact, much of the damage leads to brief delays in G1. Minor damage is repaired quickly whereas extensive damage is left as is until replication commences. Upon encountering the lesions in S phase, the replication fork may stall, exposing stretches of ssDNA (Fig. 2.1). Like $RPA^{Hs}$ , $RPA^{Sc}$ (also known as Rfa) has a high affinity for ssDNA ($10^9$ $M^{-1}$) and low affinity for dsDNA [97, 98]. The

exposed stretches of ssDNA are thus coated by Rfa proteins, triggering the recruitment of Mec1p–Ddc2p, the PCNA-like clamp (Ddc1p-Rad17p-Mec3p), and the mediator Rad9p. Collectively, they promote the activation of the effector Rad53p which amplifies the signal via kinases (e.g. Dun1p) and transcription factors (e.g. Sod1p) [7, 42, 91]. Activation of this replication checkpoint promotes transcription of repair genes, regulation of dNTP biosynthesis, activation of pre-existing repair machinery, and origin control [7]. Moreover, the stalled fork can collapse creating a DSB. The DDC branch of DDR is consequently activated by Tel1p, leading to resection, HR repair, and Chk1-mediated intra-S cell cycle delay and fork stabilization. The greater the damage, the slower the progression through S phase until the cell appears to arrest.

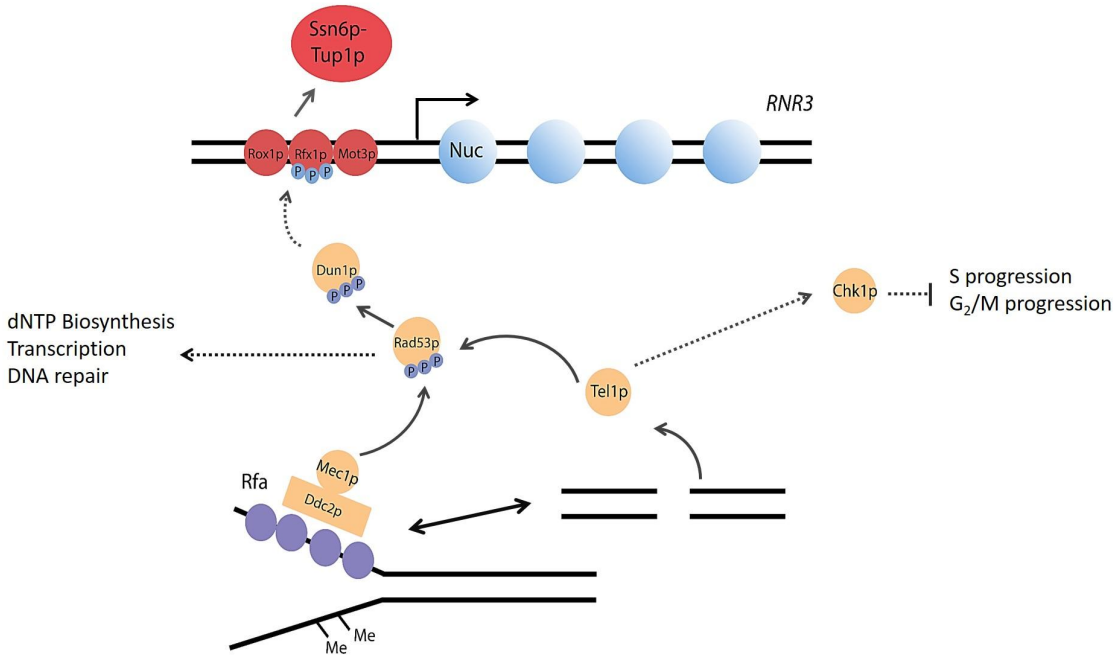



**Figure 2.1:** The working model of the replication and DNA damage checkpoints in *S. cerevisiae*. Upon detection of single-stranded damage (replication checkpoint) or double-stranded breaks (DDC), a cascade of kinases work to activate effectors such as Rad53p and Dun1p to bring about cell-cycle arrest, damage repair, and dNTP biosynthesis via derepression of ribonculeotide reductase genes.

### 2.4.1 *RNR3* is a simple reporter system

The *RNR3* transcriptional system is relatively simple, well-studied, and very sensitive to DNA damage. mRNA transcript measurements of *RNR3* show > 50-fold increase when treated with 0.02% MMS relative to untreated levels — by far the highest signal-to-noise ratio among the *RNR* genes ([30], Phenix, unpublished data). Taken together, $RNR3pr$ provides a suitable detector of defects in replication checkpoint activation.

$RNR3$ encodes Rnr3p, a nonessential component of the enzyme ribonucleotide reductase (RNR) that catalyzes the rate-limiting step of *de novo* dNTP biosynthesis in eukaryotes. Best described as an "obligate aerobe", RNR is a class Ia reductase in *S. cerevisiae* [27, 65]. It is a heterotetramer made up of a large (R1) and small subunit (R2) [65]. In mammals, the heterotetramer is made up of two different homodimers. In contrast, *S. cerevisiae* has a heterodimer for a small subunit consisting of the isomers Rnr2p and Rnr4p [39]; its large subunit can be made of a homodimer with Rnr1p (most abundant and catalytically potent), or a heterodimer with Rnr1p-Rnr3p (also isomers) when $RNR3$ is expressed [28]. Basal levels of Rnr1p, 2p, and 4p are usually high within the cell with the proteins kept inhibited by Sml1p (Rnr1p) or Dif1p (for Rnr2p and Rnr4p) [94]. Upon checkpoint activation, Mec1p/Rad53p are believed to phosphorylate Sml1p to disinhibit Rnr1p, and Dif1p to allow Rnr2p and Rnr4p to relocalize to the cytoplasm for tetramerization [29, 58, 89]. As a gene, $RNR3$ is conserved and highly repressed relative to $RNR2$ and $RNR4$; all three appear to have the same transcriptional and post-translational repression mechanisms. The role that the Rnr3-RNR variant plays in the cell is still largely uncharacterized, however, unlike the Rnr1 homodimer variant: overexpression of $RNR3$ rescues $rnr1\Delta$, yet $rnr3\Delta$ has not shown effects on fitness; granted, these evaluations were done on colony viability [28].

The gene's existence is paradoxical from an evolutionary standpoint:it is highly repressed and well-conserved among eukaryotes but its function is remains elusive and non-essential. For the purpose of our study, we ignore the function of Rnr3p and leave

it intact within the cell. Instead, we focus on its promoter dynamics: *RNR3* is a useful tool as a reporter of DDR because of its promoter [35, 46]. The *RNR3pr* contains multiple binding sites for its different repressors, making *RNR3pr* a multi-input promoter. Although the combinatorial effects of repressor site locations on *RNR3* expression have not been characterized yet, three of these repressors — Rfx1p, Rox1p, and Mot3p — have been shown to act synergistically using a *lacZ* colorimetric assay in *S. cerevisiae* [46]. Upon checkpoint activation, Rad53p and Dun1p mediate the derepression of *RNR3* by hyperphosphorylating Rfx1p (Fig. 2.1) [40]. While the actual kinase responsible for Rfx1p inhibition is still unknown, it was found that *rad*53Δ and *mec*1Δ strains eliminated the presence of Rfx1p phosphorylation; *dun*1Δ reduced phosphorylation [32,40]. Rfx1p was, nevertheless, found to be downstream of Dun1p using epistatic analysis [23, 40]. Interestingly, it was recently found that Dun1p activation also directly led to Sod1p phosphorylation, and that Sod1p binds *RNR3pr* [81]. It is possible then that Dun1p may be acting through the transcription factor to mediate *RNR3* expression, making the gene an oxidative stress response gene.

Despite the gap in knowledge, *RNR3* expression has been shown repeatedly to correlate with Mec1p/Rad53p checkpoint activation [7, 23, 28, 29, 32, 57, 58]. The effect of Sod1p suggests the involvement of oxidative stress response, and its integration within the checkpoint pathway. As of yet, no direct link between Rox1p, Mot3p, and the checkpoint pathway is found, although the two repressors are part of singlet oxygen and oxidative stress response pathways. Whether they regulate *RNR3* expression in a checkpoint- dependent or independent manner remains to be discovered.

# 3 | High-throughput Technology

## 3.1 Overview

Even in simpler organisms the DDR is complex and involves processes that may not appear to be linked to DNA repair at first glance, e.g. cobalamin biosynthesis and ribosomal biogenesis in *Halobacterium* $NRC-1$ [9, 85]. The damage checkpoint in *S. cerevisiae* also involves a myriad of structures and processes that are indirectly reparative. For example, Mec1p phosphorylates the C-terminal of $\gamma$H2A (H2A.X in humans) Ser 122 and Ser129 [25,59] which in turn act as recruitment signals for chromatin remodeling complexes, such as Swr1, Ino80, SAGA and NuA4, to remove nucleosomes at the site of damage [14, 25, 61, 82]. Mec1p and Tel1p are also thought to facilitate the relocalization of damaged segments to nuclear pores where they are anchored, so as to reduce the damage-induced increase in mobility. To complicate matters further, interactions between the checkpoint and these effectors may not be unidimensional. Kapoor *et al.* have recently found that the Ino80 complex can enhance Rad53p phosphorylation through its Ies4p [44]. Deletion of Ies4p reduced Rad53p hyperphosphorylation relative to wild-type, as well as the relative physical association of Rad53p to the Ino80 complex in the presence of MMS. They also discovered that this interaction may not be generalizable to other members of the Ino80 complex family: Swr1 failed to activate Rad53p hyperphosphorylation in vitro in the presence of MMS, suggesting that only remodellers with Ies4p are capable of enhancing Rad53p hyperphosphorylation. The lack of generalizability illustrates that the DDR is not linear in its hierarchy; one needs to survey numerous processes at a time to define a framework of what happens during genetic damage responses. "Bottom-up"

approaches, while hypothesis-driven and focused, can encounter problems with studying and organizing this complexity — i.e. top-down, systems-level techniques are needed.

The last fifty years have witnessed the development of systems biology and 'omics' technology into the integrative and high-dimensional fields they are today. The study of life is inherently complex, beginning with the spatio-temporal dynamics within the cell and pervading cell–environment and multicellular interactions. Omics technologies can screen thousands of entities, be they cells (e.g. genomics, phenomics, epigenomics,and interactomics), chemicals (e.g. metabolomics and chemical genetic screens), or macromolecules (e.g. lipidomics, proteomics, and transcriptomics). At the crux of these fields lie challenges with:

1. the systematization of high-throughput techniques to reduce technical variability; and

2. finding meaningful quantifiable models from which we can glean insights.

## 3.2 SGA

Synthetic genetic array (SGA) is a technique that enables high-throughput creation of mutant libraries of model organisms, such as budding or fission yeast. It exploits the organism's mating dynamics, as well as drug resistance cassettes to generate an array of strains with multiple mutations of interest.

### 3.2.1 Budding Yeast Mating Dynamics

*S. cerevisiae* cells can grow and divide as diploid or haploid. As haploids, the cells take on one of two mating types, A or $\alpha$, that regulate the types of pheromones and cell-surface receptors expressed. Determination of mating type occurs at the $MAT$ locus, wherein the presence of the gene $A1$ signals the profile of gene expression and repressions associated with the $MAT\ A$ type. Similarly, in $\alpha$ types, the $MAT$ locus harbours the $\alpha 1$ and $\alpha 2$ genes that regulate the expression of a different gene set. $MAT\ A$ cells secrete A pheromones,

and express the $\alpha$ factor recognizing cell-surface receptor Ste2p [15]; $MAT\ \alpha$ cells secrete $\alpha$ pheromones and the A-factor recognizing surface receptor Ste3p [33]. The cell-surface receptors provide reciprocal recognition; cells of one mating type thereby mate with those of the opposite type by secreting their type-specific pheromones.

Binding of pheromones activates a heterotrimeric G-protein-mediated mitogen- activated protein kinase (MAPK) signal cascade (rev. in [36]). Downstream of the signal transduction pathway is the formation of a projection, known as a shmoo, in each cell type towards its mate —- i.e. they prepare to merge themselves. Another downstream effect of the MAPK pathway is G1 cell-cycle arrest via Fus3p-activated Far1p [66]. Phosphorylated Far1p sequesters Cdc28p-Cln1/2p, prohibiting cell-cycle progression into S phase [66, 67].

After cell- wall and membrane fusion, the two nuclei merge in a process called karyogamy, resulting in a zygote. Generally, the A/$\alpha$ diploid cell continues its cell cycle without sporulating (returning from the spore state). In response to starvation, however, diploids undego meiosis and sporulate; the resultant spore can lay dormant until environmental conditions are permissive to growth again [20]. Sporulation produces four haploid (2 $\alpha$ and 2 $A$) cells in a spore sac or ascus.

### 3.2.2 Synthetic Genetic Array Analysis

When available, a suitable biological system allows high-throughput technologies to be useful in three ways. First, high-throughput genomic, transcriptomic, and proteomic screens can be done at condition gradients (e.g. drug doses, temperature, or time courses) to track changes in phenomes. Second, the integration of information from several high-throughput techniques utilize manually-curated datatbases to uncover higher-order organization and emergent properties. Third, the development of yeast two-hybrid and synthetic genetic array (SGA) analyses facilitated large-scale exploratory studies of protein-protein and genetic interactions, respectively. In 2010, for example, Bandyopadhyay *et*

*al.* reported how the genetic networks of *S. cerevisiae* "rewire" themselves in response to DNA damage [10]. They generated 80,0000 pairwise double-mutant strains using synthetic genetic arrays (SGA) whose growth was scored to assess genetic interactions. Their findings enabled them to construct several MMS-dependent genetic interaction networks of genes that are synergistic and antagonistic to one another in the DDR. Synthetic genetic array (SGA) analysis involves the mating of a query strain with a deletion array, a set of strains wherein a different gene is replaced by a selection cassette [5]. Strains that have successfully mated are selected for and sporulated, before haploid selection for either MAT A or $\alpha$ types to be assayed (Fig. 3.1). Different query genotype identities expand the usefulness of SGA to yield multi-dimensional results:

1. One strain with a mutation or mutations: the assay will consequently yield information about genetic interactions and their effect on viability. This approach has been the traditional use of SGA, with colony sizes measured by flatbed scanners or plate readers, such as PhenoBooth™. Alternatively,growth rates can be quantified using a plate spectrophotometer, such as Bioscreen C.

2. One strain with a fluorescent reporter: R-SGA's main advantage is its ability to quantify combinatorial phenotype enhancements and suppressions outside of viability.

3. One strain with a fluorescent reporter and mutation(s): if done, both viability and non-viability based phenotypes can be quantified for combinatorial mutations;

4. 1–3 but with a query array: these impart higher dimensionality, especially when creating pairwise combinatorial mutants, or, more specifically, a query set enriched for a process or complex of interest.

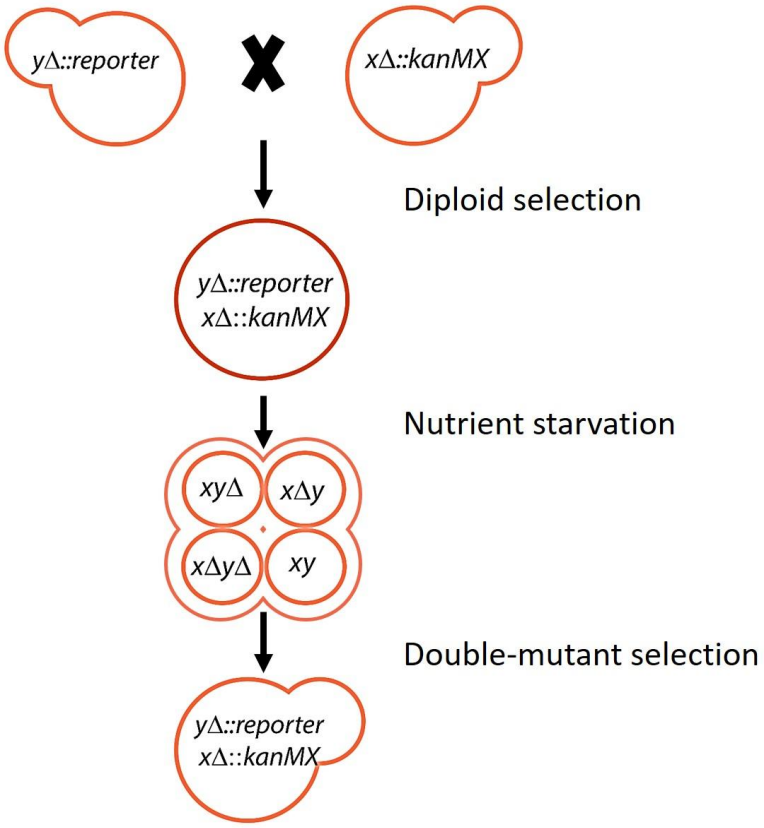



**Figure 3.1:** Schematic of reporter-based SGA (R-SGA). A query haploid (usually MAT $\alpha$) is mated against a commercial library carrying different mutations. The resultant mixture is left to grow and proliferate in media supporting positive selection of successfully generated diploids. Diploids are then forced into meiosis and sporulation by nutrient starvation, an event that can be checked using microscopy. Finally, the haploid carrying the desired genotype (i.e. reporter and deletion) is again selected for, grown, and preserved for screening.

Reporter-based-SGA has two other benefits. First, the introduction of quantification via fluorescence has the added benefit of providing *in vivo* data for statistical analysis. Many high-throughput techniques measure targets *ex vivo*, with the accuracy, precision, and yield compromised by the amount of processing. The loss in yield can additionally compromise data quality and limit inferences. For example, gene expression microarrays rely on the isolation of RNA from cells, followed by reverse transcription and labelling, and hybridization of the labelled cDNA to a microarray of genes to quantify luminescence [73]. Relative losses in yield during RNA isolation can increase type I or II errors depending on what the experimenter defines as a hit. Microarrays are also sensitive to sample

variability, and probe sequence: probes designed to hybridize to the same sequence but with slightly different sequences — i.e. potentially different binding affinities — were reported to yield intensities that are a decade apart [26]. Proteomic techniques also have their share of challenges. They need to be well-controlled at several steps sample purification, digestion, fractionation, concentration, and affinity capture.

Second, R-SGA enables the acquisition of both population-level and single-cell level information by simply varying the acquisition instrument. What this advantage translates to is the ability to measure phenotypes beyond viability and growth.

At the population level, high-throughput fluorescence microscopy (HTFM) has been widely developed to screen plates of strains grown on solid or liquid media. Solid-phase HTFMs essentially take a snapshot of the entire plate after laser- induced excitation of the reporter protein. While generally faster than the liquid counterpart, solid-medium arrays suffer from position-dependent growth effects: colonies at the periphery of the array grow larger as they share nutrients with 2–3 other colonies on one side, and have access to a substantial portion of unoccupied medium on the other side; colonies in the middle are smaller. Statistical normalization becomes very important here.

Liquid-phase HTFM should not suffer from the same positional effects provided that the entire protocol is executed in liquid-phase. The main disadvantage is that it does not distinguish the presence of subpopulations within a colony. A mutation (or collection thereof) that promotes the divergence of an isogenic line into two distinct reporter-expressing subpopulations can be easily missed when quantifying fluorescence using a classical statistic, such as mean, median, or total fluorescence. A corollary of simple statistics' low resolving power is that these subpopulations can result in mixed phenotypes in follow-up studies, or even type II errors if the wild type distribution is actually different.

## 3.3 Flow Cytometry

While different instruments are organized differently, all flow cytometers combine fluidics, optics, and electronics to assess the fluorescence properties of labeled cells along with size

and internal complexity (Fig. 3.2).

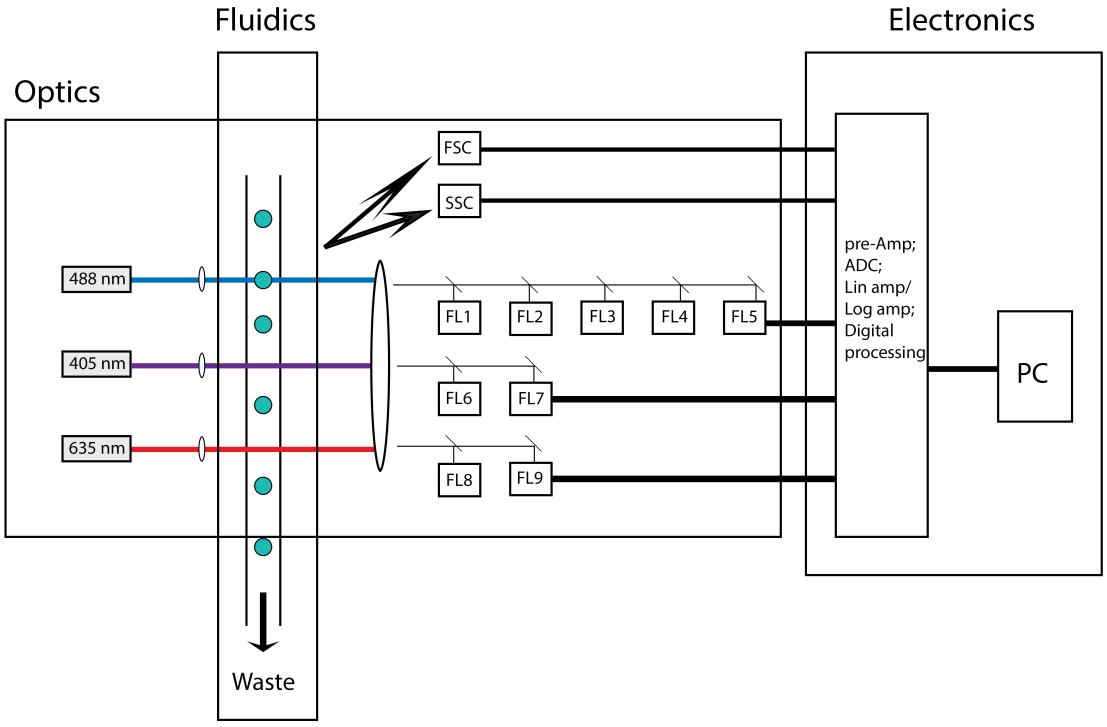



**Figure 3.2:** Schematic of data acquisition order in Cyan ADP 9.0. Wavelengths denotes lasers; ellipses denote focusing lenses; FLx refers to photomultiplier tubes; 45 deg slashes denote dichroic mirrors.

When sample of the particles or cells is injected into the cytometer's flow cell, the stream containing the cells is centered surrounded by isotone sheath fluid flowing at a higher velocity [2, 3]. The differential speed — and subsequent pressure — between the two streams allows them to flow coaxially without ever mixing (laminar flow) [3]. Reynolds number ($R_e$) can be calculated to estimate whether laminar flow will happen given the characteristics of the cytometer

$$R_e = \frac{\rho V D}{\mu} \tag{3.1}$$

where $\rho$ is central fluid density, V is the mean fluid velocity, D is the tube diameter, $\mu$ is fluid viscosity, and $R_e < 2300$ indicates laminar flow [3]. The drag effect created by the faster sheath fluid also modulates the central fluid's velocity causing particles in the center of the central fluid to have a higher velocity than those closest to the periphery. Consequently, the particles flow down the tube and become exposed to the laser beam in

a single file. This phenomenon is called hydrodynamic focusing [2, 3, 4].

As the cells flow, each one is illuminated by a laser beam of a specific wavelength. The light is diffracted and scattered, and, if applicable, absorbed by fluorophores. Upon illumination with the appropriate excitation wavelength, a fluorophore's valence electrons absorb the photons, and emit photons of lower energy as the former return to their ground state [4]. Emitted photons scatter in all directions depending on the complexity and size of the cell. Photons that are scattered forward ($\pm 20$ deg from the beam axis in all directions) are detected by a silicon photodiode plate (in the forwards scatter channel, FSC) on the opposite end of the laser [3,4]. These photons allow the estimation of cell size, wherein the greater the intensity of the forward scatter, the larger the particle. Side scatter, on the other hand, correlates with the granularity or internal complexity of the cells [2]. The side-scattered rays tend to be less intense. Consequently, the more sensitive photomultiplier tubes (PMTs) are used as detectors of side-scattered photons instead. Side-scattered light is gathered at 90 ° to the beam axis and redirected by a dichroic mirror to the PMT-containing side scatter channel (SSC) [3,4]. In the presence of light rays whose wavelengths are not reflected by the dichroic lens, the beams continue until they meet the appropriate filter that reflects them into a fluorescence channel. The most commonly used filters are long pass, short pass, and band pass filters [3]. Long pass filters allow light with wavelengths longer than a pre-set value through; short pass, conversely, allow wavelengths shorter than the thresh- old. Finally, band pass filters allow a narrow range of wavelengths to pass through.

Once photons reach the photomultiplier tubes, an electrical current is generated that is then converted into a voltage through a series of linear or logarithmic amplifiers and an analogue-to-digital converter (ADC) [2, 3]. Whether the amplification should be linear or logarithmic depends on the investigator's interest. Usually, linear amplification is used when the results are expected to lie within a narrow range of signals. Fluorescence experiments, however, tend to result in a broad range of responses, and, thus, linear

amplification cannot be used for those [3, 4]. Instead, logarithmic amplification is used because it has the added advantage of scaling weak and strong signals such that the former are magnified and the latter are compressed. The investigator can also opt to collect peak intensity (maximum registered fluorescence intensity within the cell), or fluorescence area (integral of single-cell fluorescence curve from the moment a cell passes enters the laser field to the moment it leaves). Either way, acquired data is single-cell data, and we can observe the shape of the reporter distribution within samples.

## 3.4 Epistasis Analysis

Since its conception in 1909, the term epistasis has taken various definitions depending on the field and trait dynamics. Originally, epistasis referred to the deviation from independence of genes at two different loci in their ability to effect a trait [1]. Bateson defined epistasis asymmetrically in a diploid system wherein a predisposing allele (one dominant allele) at locus 1 effects a trait despite heterozyosity or predisposing genotype (two dominant alleles) at locus 2 (Table 3.1) [1]. The absence of locus 1 predisposing allele, however, effects a different trait when combined with the predisposing alleles at locus 2, and no trait expression when predisposing alleles are absent in both loci.

**Table 3.1:** Sample penetrance table demonstrating asymmetric epistasis.

| | $a/a$ | $a/A$ | $A/A$ |
|---|---|---|---|
| $b/b$ | 0 | 1 | 1 |
| $b/B$ | 2 | 1 | 1 |
| $B/B$ | 2 | 1 | 1 |

In this case, epistasis can be though of as “interlocus dominance”, mimicking the effect of dominance of a single allele at a specific locus. The “dominant” locus is said to be epistatic to the “recessive” locus but not *v.v.*.

Neuman and Rice later defined epistasis in terms of masking in a symmetric model [63]. Here, the absence of a predisposing allele at either locus suppresses trait expression, and presence of at least one predisposing allele at both loci effects the trait [63]. Here, symmetry allows bidirectional inference of epistasis. Because the effects of genotype variation at one locus cannot be discerned in cases where no predisposing allele is present at the second locus and *v.v.*, each gene masks the other. The complexity of gene-gene interactions, however, can escape characterization even in symmetric penetrance (population trait distribution) tables.

A symmetric penetrance model of epistasis where only predisposing genotypes at either locus can effect trait displays masking of genetic variability in one locus in the presence of the predisposing genotype at the second locus. The expression of the trait, however, does not depend on interaction between the predisposing alleles since the presence of only one predisposing allele does not effect the trait. Therefore, in this model, also called a heterogeneity model, the two loci are independent yet epistatic by Neuman and Rice's definition.

What is the most intuitive definition of epistasis then?

If we want to infer genetic interactions, we are asking for the degree to which changes at different loci combinatorially affect the expression of a trait. That is, how does homozygosity or heterozygosity at one locus interact with those of a different locus? Noting that at this level of analysis we cannot infer biological epistasis (defined as physical interactions), several statistical epistasis models are being developed to bridge the gap between statistical and biological epistases.

The earliest statistical model defines epistasis as departure from linearity (Eq. 3.2), with the underlying assumption that genes are independent in effecting the trait [31]. Say there are hypothetical genes x and y. If they are independent, deletion of both genes should enhance fitness by a magnitude equal to the sum of the genes' independent effects

— i.e. 20-fold (Fig. 3.3).

$$F_{ij} = F_i + F_j \tag{3.2}$$

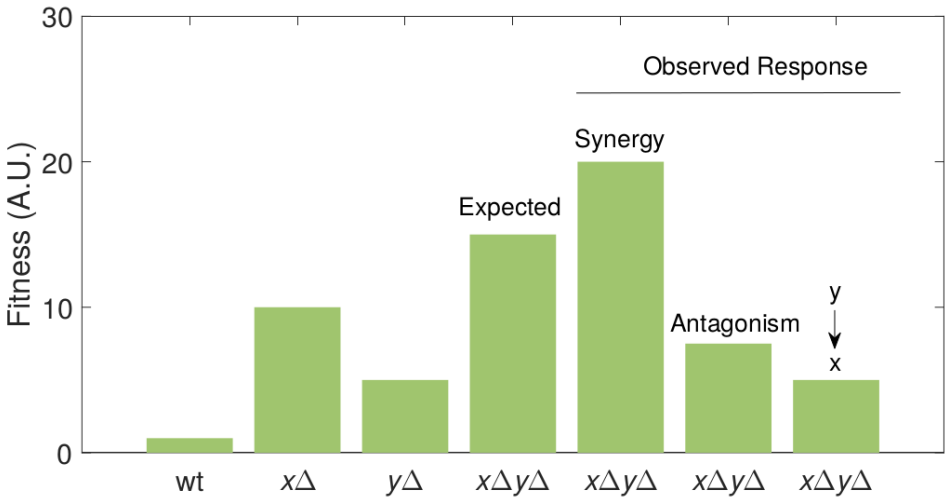



**Figure 3.3:** Statistical epistatic interactions between genes are deviations from linearity ("Expected"). Here, $x\Delta$ and $y\Delta$ are advantageous mutations since they enhance fitness relative to $wt$.

Deviation from the expectation implies interactions. Whether these interactions can be labeled as positive, negative, synergistic, or antagonistic depends on whether the mutations are advantageous (enhance fitness) or deleterious (diminish fitness). Figure 3.3 depicts the behaviour of advantageous mutations. Where $x\Delta y\Delta$ enhances fitness more than expected (25 A.U. vs. 20 A.U.), the interaction is termed positive and synergistic. Similarly, where $x\Delta y\Delta$ decreased fitness their interaction is negative and antagonistic. If $x\Delta$ and $y\Delta$ were deleterious mutations, enhancement of diminished fitness would be considered negative and synergistic, whereas reduction of diminished fitness is considered positive and antagonistic. Finally, if $x\Delta$ and $y\Delta$ were both advantageous or deleterious but $x\Delta y\Delta$ appears inviable or viable, respectively, theirs is a reciprocal sign interaction.

The heterogeneity model also assumes that the loci are independent, and is given by:

$$F_{ij} = F_i + F_j - F_i F_j \tag{3.3}$$

Because epistasis is assessed with respect to a linear regression model, the scale of the coordinate system within which the model is projected matters.

A multiplicative model of epistasis, in contrast, assumes epistasis wherein the expected effects of the two loci is simply the product of the phenotypes conferred by the individual loci (Eq. 3.4). Both the multiplicative and heterogeneity mod- els can also display epistasis as departure from additivity when logarithmically- transformed.

$$F_{ij} = F_i \times F_j \tag{3.4}$$

While a few other statistical models are developed, such as minimax, logistic, and multifactor dimensionality reduction, there is no single model that encompasses the information encoded in combinatorial genetic variability. First, scale- dependence often complicates the interpretation of an interaction score. Second, epistasis as a concept still carries various definitions, and the choice of the statisti- cal model depends on what the investigator aims to characterize. Finally, epistasis models assume a closed system where genes under study are the only ones that can be characterized. Gene-gene interactions can be complex, however, requiring high-throughput pairwise screens to mitigate omitted-variable bias in epistasis analysis. Often, multiple models are used in tandem to ascertain the presence of interactions.

# 4 | Objective and Hypotheses

## 4.1 Purpose

Given the importance of the checkpoint members to repair and genome stability as well as the involvement of unanticipated processes and pathways in the DDR, our goal is to discover novel processes that indirectly affect S. cerevisiae's response to MMS-induced damage.

To use $RNR3$ as a detector, we take advantage of reporter-based synthetic genetic array analysis (R-SGA) which combines fluorescence-based reporters with SGA analysis. We use a dual-colour reporter system with yeast enhanced green($yEGFP$) and blue fluorescent ($yEBFP$) proteins' genes under the control of the $RNR3$ promoter ($RNR3pr$) and actin promoter ($ACT1pr$), respectively. For our purposes, $yEGFP$ serves as a reporter of checkpoint activation; $yEBFP$ serves as an indicator of cellular expression capacity. Abnormalities in $yEBFP$ expression are interpreted as inherent abnormality in growth or expression without distinction made as to where the dysfunction lies.

## 4.2 Goals

The goals of this thesis are to:

1. develop a reproducible systematic screen protocol;
2. screen and identify gene mutations that confer differential abnormalities in $RNR3pr-GFP$ reporter expression at two MMS doses;

3. screen a subset of the outliers found in the first screen that are additionally deleted for $RFX1$ to identify differential genetic interactions.

## 4.3 Benchmark Expectations

As a way to consolidate the quality of our results, we expected to observe the following:

1. Non-essential genes involved in DNA damage sensing — e.g. $DDC2$, $RAD9$, $MRC1$, $DDC1$, $RAD17$, $MEC3$, and $RAD24$ — will lower reporter expression when deleted;

2. Deletion of $IES4$ will reduce reporter expression because of its role in en- hancing Rad53p phosphorylation [44];

3. Synergy between $ROX1 - RFX1$ and $MOT3 - RFX1$ as previously reported by Klinkenberg *et al.* [46].

4. Deletion of the $RFX1$, $ROX1$, and $MOT3$ would derepress $RNR3pr$ further leading to an enhanced phenotype [45, 46].

These expectations serve as a measure of agreement between our results and pertinent literature.

## 4.4 Hypotheses

Given the exploratory nature of this study, we can only hypothesize with some directionality the behaviour of a subset of strains based on previous published work.

Broadly, strains deleted for nonessential genes involved in DNA damage sensing and repair are expected to exhibit low fluorescence relative to neutral mutants. Similarly, because of the role of Ies4p and Ino80 in enhancing Rad53p phosphorylation in response to Mec1p/Tel1p activation, we also expect that gene deletions of Ino80 complex members will emerge as hits that reduce reporter expression. Other chromatin remodeling complexes, particularly histone acetylation complexes (HATs), which decondense for expression DNA

may also non-specifically decrease reporter expression if their deletion interferes with the derepression and/or tran- scription of $RNR3$.

Regarding the epistasis miniarray with $rfx1\Delta$, we expect that deletion of HATs will reduce the high fluorescence levels conferred by $rfx1\Delta$.

# 5 | Methods

## 5.1 Library Generation

### 5.1.1 Synthetic Genetic Array Analysis

We used the query strain as a $MAT\ \alpha$ (BY4742; Table 5.1) containing the dual- reporter construct depicted in Figure 5.1 at the $MET15$ site.

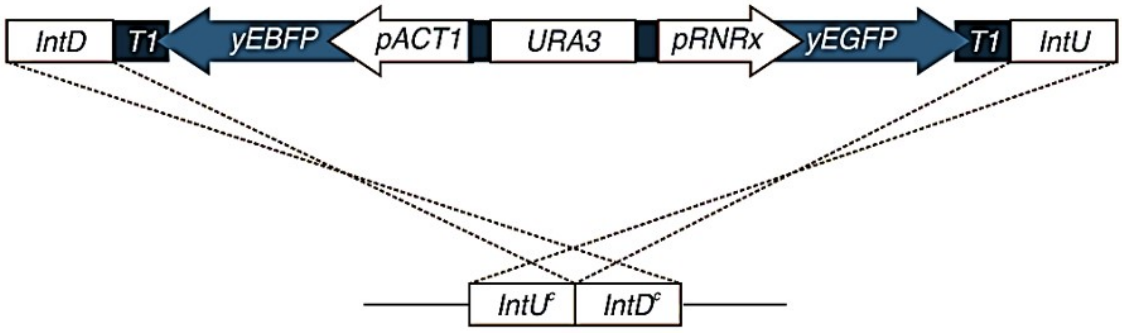
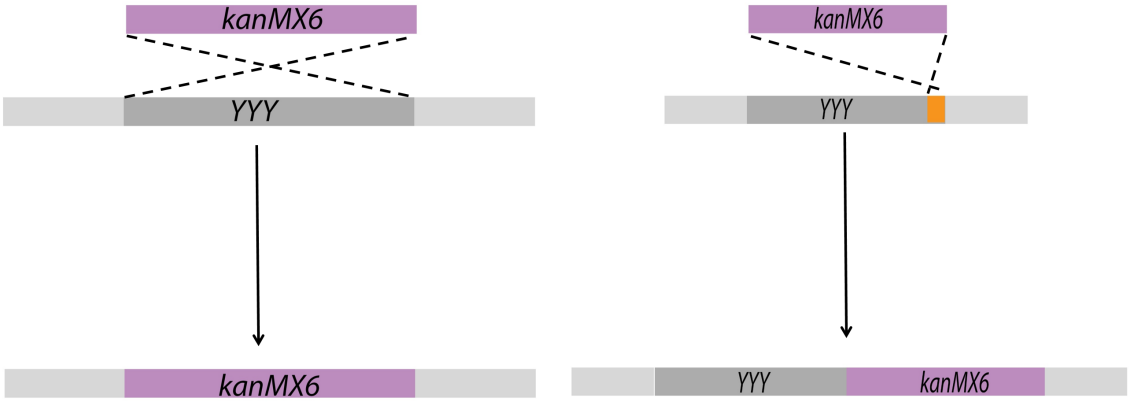



**Figure 5.1:** Dual-reporter insert structure with $URA3$ to facilitate auxotrophic selection, $RNR3pr$-driven yeast enhanced GFP ($yEGFP$), and an $ACT1pr$-driven yeast enhanced internal control, $yEBFP$. $T1$ refers to $TEF1$ terminators; $IntU$ and $IntD$ refer to complementary sequences upstream and downstream, respectively, of internal $MET15$ sequence. [68]

The MAT A array contains 5,850 strains, each of which contains the *BY4741* background genotype and a kanMX -replaced gene (Table 5.1). The strains are originally from the commercial deletion mutant array (DMA) — a collection of strains with nonessential gene knock-outs — and the decreased abundance by mRNA disruption (DAmP) array — a collection of essential gene knock-downs (Fig. 5.2).




**(a)** **(b)**

**Figure 5.2:** Deletion by homologous recombination of a selection cassette ($kanMX6$). (a) DMA: the cassette replaces the gene completely. (b) DAmP: cassette disrupts gene by integration at the 3' end. Disrupted gene is knocked down — i.e. its expression is reduced by reducing mRNA levels.

Before mating, strains from the two collections were shuffled and condensed from 96-colony to 384-colony formats. To each 384 density plate, four WTs were added: $YOR202W$ ($his3\Delta$), $YEL021W$ ($ura3\Delta$), $YDL227C$ ($ho\Delta$), and $YCL018W$($leu2\Delta$). Hereafter, the neutral mutants will be referred to as wild- types (WTs).

**Table 5.1:** Parent strains and plasmids used in this study

| Name | Genotype | Source |
|---|---|---|
| pFA6A-kanMX6 | $AMPpr - AMP\ TEFpr - KANR - TEFter$ | Bahler *et al.*, 1998 |
| pFA6$\alpha$-natMX6 | $AMPpr - AMP\ TEFpr - NRSR - TEFter$ | Hentges *et al.*, 2005 |
| pHPA03 | $ACT1pr - yEBFP\ TEF1pr - URA3 - TEFter$ | In house |
| | $RNR3pr - yEGFP$; Tet$^{r}$ | |
| BY4741 | $MatAcan1\Delta :: STE2pr - his5^{Sp}$ | Brachmann *et al.*, |
| | $lyp1\Delta :: STE3pr - LEU2his3\Delta1$ | 1998 derivative |
| | $leu2\Delta0ura3\Delta0$ | |
| BY4742 | $Mat\alpha can1\Delta :: STE2pr - his5^{Sp}$ | Brachmann *et al.*, |
| | $lyp1\Delta :: STE3pr - LEU2his3\Delta1$ | 1998 derivative |
| | $leu2\Delta0ura3\Delta0$ | |

Facilitated by the Singer RoToR automated pinning robot, the query and the array strains (plated on yeast extract peptone dextrose (YPD media) + 200 mg/L G418 agar plates) were transferred and mated on YPD agar plates. After incubation for 48 h at 30 deg $C$, diploids that have successfully mated were selected for using synthetic media (SM) complete agar plates supplemented with 200 mg/L G418 agar plates and uracil drop-out. Colonies that are viable after overnight incubation at 30 deg $C$ (Gyromax TM 737 Orbital Incubator Shaker, Amerex Instruments Inc.) on the selection plates are then forced to sporulate on enriched sporulation (SPO) media. The SPO plates were incubated at 25 deg $C$ for 5 days, with routine microscope checks of spore formation every day after the first 48 h.

Sporulated colonies were then transferred to SM media without histidine and arginine, and supplemented with 100 mg/L L-canavanine sulfate salt (Sigma, C9758) for $MATA$ haploid selection. Viable cells after 2-day incubation at 30 deg $C$ were again selected on the same medium plus 200 mg/L G418 and without uracil. After the second selection process, cells were transferred to liquid SM media without histidine, uracil, and arginine, and supplemented with 100 mg/L L-canavanine and 200 mg/L G418 for growth that minimizes plate-center effects on growth. Strains were preserved at -80 deg $C$ (Forma -86C ULT Freezer, Thermo Scientific) as well as back-crossed to a $MAT\alpha$ BY4742 derivative to ascertain that the former are $MATA$ haploids.

All in all, 194/5850 strains from the DMA and DAmP collections did not survive the SGA process.

### 5.1.2 Haploid Test

A $MAT\alpha$ BY4742 variant was grown to confluence overnight before pinning into a grid of 384 colonies on yeast extract peptone dextrose (YPD) agar medium. The grid of $MAT\alpha$ tester strain were incubated at 30 deg $C$ until colonies were visibly large (maximum 48 h). These colonies were then transferred to a replica of the library colonies on YPD agar

plates, and incubated overnight at 30 deg $C$. Mated colonies were subsequently transferred to SC-6 agar medium for selection and incubated prior to visual inspection of growth. Of the 5,656 strains that survived the SGA protocol, 183 failed the haploid test, suggesting that they may be either diploid or $MAT\alpha$.

## 5.2 PCR-based Strain Validation

Polymerase chain reaction (PCR)-based validation involves the amplification of a sequence internal to the gene deleted. If the gene is not present within the cell, then no amplicon will be detected through DNA gel electrophoresis.

Quick genome extraction was done using a colony of the strain of interest exposed to 20 $\mu L$ of breaking buffer, acid-washed glass beads, and 20 $\mu L$ of 25:24:1 phenol:chloroform:isoamyl alcohol. The sample was vortexed using VortexGenie for 3 min prior to addition of 20 $\mu L$ Tris-EDTA (TE) buffer (pH 8.0) to chelate and solubilize nucleic acids. Samples were vortexed at 13000g for 5 min., and their supernatant was transferred to 100 $\mu L$ of 100% ethanol. The ethanol-supernatant mixture was vortexed for 5 min at 13000g. After discarding the supernatant, the remaining pellet was air-dried for 30 min. before addition of 20 $\mu L$ TE buffer.

PCR-validation was done in 25-$\mu L$ aliquots using Phusion polymerase and internal primers. PCR products were run on 1% agar-TE (+ 2 $\mu L$ RedSafe dye) gels in TE buffer for 45 min. at 80 V. See Appendix 2 for list of primers.

## 5.3 Flow Cytometry General

We use CyAn ADP 9.0 Analyzer (Beckman Coulter) with an auxiliary IntelliCyt plate sampling platform (Beckman Coulter) for high-throughput analysis of our cells. CyAn ADP provides the two scatter plus nine fluorescence parameters that allow the detection of a range of fluorophores. The analyzer used here is capable of collecting up to 70,000

events/s and has a resolution of 65,536 channels for each fluorescence parameter. Selected parameter values and their time stamps are then saved in a FCS 3.0 file that can be read by commercial software for well parsing and data analysis. HyperCyt™, for example, enables the investigator to separate (parse) signals into wells by inference from the temporal gaps between signal peaks. Parsed wells can then be analyzed using a commercial analytic workstation, e.g. Kaluza™or FlowJo™, or statistical software, e.g. MATLAB™. We use MATLAB™for both well-parsing, thanks to an in-house graphical user interface (GUI) that uses the `fcs_read` function, and fluorescence analysis because it allows for greater freedom and transparency in structuring, cleaning, and analyzing data.

## 5.4 Primary and Secondary Chemogenetic Screen

The following subsections detail the protocols used for the primary and secondary chemogenetic screens. Given the intriguing dose-response behaviour of the reporter (Fig. 5.3 inset) in BY4741, methyl methanesulfonate (MMS) concentrations were chosen so that they provide snapshots of the cellular landscape at peak reporter expression (low MMS) and closer to basal expression (intermediate MMS). Our interest in reporter behaviour at intermediate MMS doses is two-fold:

1. cellular behaviour at intermediate MMS doses is relatively unexplored com- pared to low MMS;

2. we want to know whether the cellular landscape resembles that of the untreated group given that the distribution of $yEGFP$ is bimodal (Fig. 5.3).

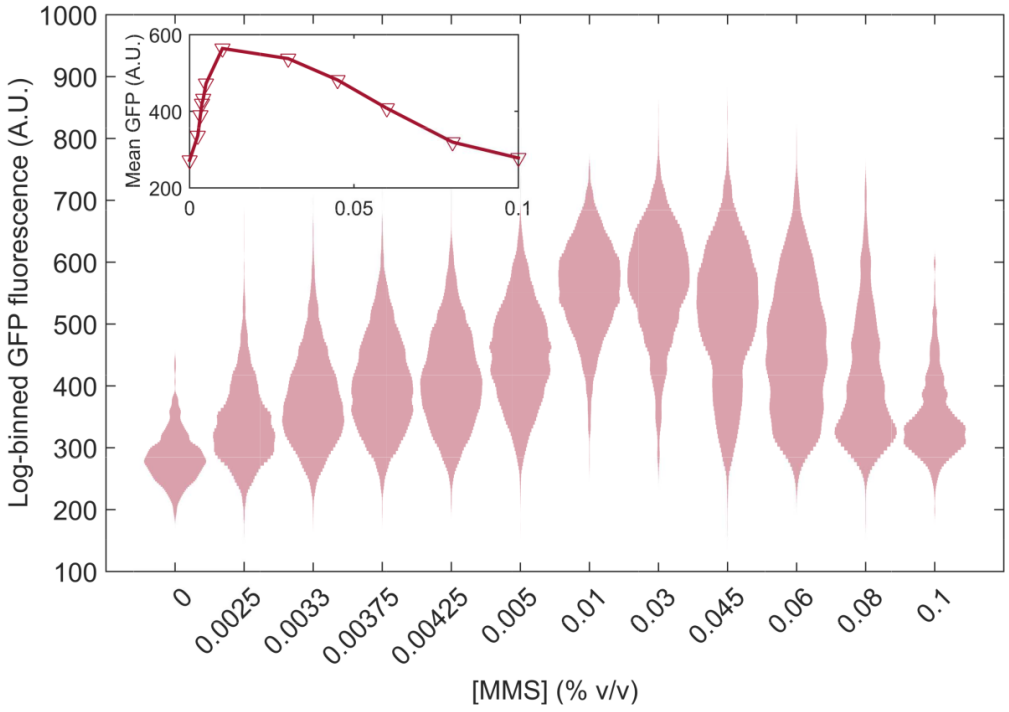



**Figure 5.3:** Violin plot of the $BY7471$ derivative $yor202w\Delta$'s MMS dose-response curve. Logarithmically-amplified $yEGFP$ fluorescence was rescaled from $2^16 to 2^10$ to better visualize the shape of the distribution. Cells were grown for 24 h in SM complete + glucose and adenine, before release in the same media. Released cultures were incubated for 3 h at 30 deg $C$, treated with MMS, and incubated again for 5 h prior to flow cytometry screening.

### 5.4.1 Primary Screen

**Sample Preparation.** A plate from the library of *S. cerevisiae MAT A* mutants was thawed at room temperature after preservation at -80 deg $C$ in 15% (v/v) glycerol. After complete thawing, the 384-well plate was de-condensed using the SGA robot (Singer RoToR) into four agar plates containing YPD supplemented with 200 mg/L G418 (Wisent 400-130IG) media. The agar plates were then incubated for 48 h at 30.0 deg $C$.

Forty-eight hours later, cells from the incubated culture were released into fresh, sterile 200 $\mu L$ SM complete supplemented with glucose (2% final concentration) in sterile, round-bottom, shallow 96-well plates (Singer) using the SGA robot. The liquid culture plates were then sealed in a plastic container lined with sterile damp towels to minimize

evaporation, and incubated for 48 h at 30.0 deg $C$.

**Drug Exposure.** On screen day, the liquid culture plates were mixed and 2 $\mu$L/well were transferred to eight sterile deep 96-well plate with glass beads con- taining 400 $\mu$ L/well SM complete at 20 min intervals. Out of the two sets of replicates per plate, four are designated to be exposed to low (0.015% (v/v)) methyl methanesulfonate (MMS; Sigma, 129925) while the remaining set would receive an intermediate dose (0.055% (v/v)) of MMS.

Prior to drug introduction, the set of eight release plates were incubated at 30.0 deg $C$ with 200 rpm shaking for 3 h. Stock solutions of MMS (99%, Sigma Aldrich) doses were diluted in SM complete to obtain 0.675% for the low dose, and 2.48% for the intermediate dose. After the 3h-incubation period of each plate, 10 $\mu$L of the MMS stock was mixed with the mid-log cells in each well, beginning with the plates in the 0.015% MMS-condition set, and followed by the 0.055% MMS-condition replicates. Each plate was incubated for 5 h at 30.0 deg $C$ with 200 rpm shaking post-drug introduction.

**Flow Cytometry.** Immediately before flow cytometry, the cells were diluted by 1:4 in sterile 50 mM sodium citrate buffer (pH 7.4) in sterile, shallow, round-bottom 96-well plates.

Screening was done using the IntelliCyt and CyAn ADP analyzer systems (both Beckman Coulter) with the following parameter settings: 0.56% threshold; side scatter 400 V; FITC 480 V; and Violet 1 650 V. All parameter gains were set to 1.0.

The plate reader's sampling was set to the following protocol: sipping for eight seconds/well, with two seconds uptime between wells, and shaking every 12 wells for four seconds. No inter-well rinsing was used, however, with cleaning mainly done in isotone sheath fluid (IsoFlow™, Beckman Coulter) for 30 s after each sub-plate was completely

screened.

### 5.4.2 Secondary Screen

**Sample Preparation and Drug Exposure.** Cells were diluted 1:100 in SM complete with 200 mg/L G418 without ammonium sulphate and incubated for 45 h at 30 deg $C$ and 250 rpm shaking. Absorbance at 600 nm ($OD_{600}$) was measured after dilution of wells by 1:9 in SM complete with 200 mg/L G418 without ammonium sulphate. Samples from the original plate were released using a volume that yields $OD_{600} = 0.03$ in 400 $\mu$L of SM complete with 200 mg/L G418 without ammonium sulphate. Released samples were grown for 3 h with shaking prior to being split into the three 100-$\mu$L treatment groups: untreated, 0.015% MMS, or 0.060% MMS. Both drug receiving groups were treated from 2× MMS in SM + G418 without ammonium sulphate stocks prepared on the same day, and all plates were prepared at 25-min. intervals.

**Flow Cytometry.** Cells were not diluted for flow cytometry to maximize sample size. Three drops of FlowPro fluorescent beads (Beckman Coulter) were added to the first well of each plate. Screening was done using the IntelliCyt and CyAn ADP analyzer systems with the following parameter settings: 0.1% threshold; side scatter 400 V; FITC 550 V; and Violet 1 550 V. Parameter gains were set to 8.0, 1.0, and 1.0, respectively; FS gain was set to 5.0. The plate reader's sampling was set to the following protocol: sipping for 12 s/well, with 2 s uptime between wells, and shaking every 12 wells for 12 s. No inter-well rinsing was used, however, with cleaning done in isotone sheath fluid (IsoFlow™, Beckman Coulter), 10% bleach, and deionized distilled water for 30 s each after each screening.

## 5.5 $rfx1\Delta$ Chemogenetic Epistasis Miniarray

### 5.5.1 Library Generation

For the creation of the RFX1 -deleted miniarray, SGA steps were done in liquid phase —- i.e. all media were liquid —- except for the sporulation step. Strains of interest were

individually diluted 1:60 in YPD + 200 mg/L G418 + ade- nine, followed by incubation for 2 days at 30 deg $C$ and 200 rpm shaking. $BY4742rfx1\Delta :: natMX$ (query), created using high-efficiency LiAc yeast transformation [99], was grown overnight at the same incubation settings. Each and query were mixed in YPD + adenine at a ratio of 1:1:2 (strain:query:media) before overnight incubation at 30 deg $C$ and 200 rpm shaking.

Diploid selection was done by inoculating 15 $\mu$L in 185 $\mu$L YPD + 200 mg/L G418+ 100 mg/L ClonNAT (Werner Bio Agents) and incubating the mixture for 2 days at 30 deg $C$ and 200 rpm shaking. Codón *et al.* reported that sporulation in liquid media is inefficient [20]. Consequently, SPO agar plates with 96 indentations/plate (made by sterile RoToR pin pads) were filled with 3 muL of the selected diploids. SPO plates were air-dried and then incubated and checked as described in section 5.1.1. Haploid selection was done in liquid media with one part sporulated cells and 19 parts SM –HIS, –URA, and –ARG supplemented with 100 mg/L L-canavanine, 200 mg/L G418, and 100 mg/L ClonNAT. Cells were incubated for 3 days with shaking before incubation on SM -HIS, -URA, and -ARG supplemented with 100 mg/L L-canavanine, 200 mg/L G418, and 100 mg/L ClonNAT agar plate for 2 days. Finally, 300 $\mu$L of YPD + 200 mg/L G418 + adenine was inoculated with the colonies and incubated overnight with shaking. At this point, colonies were preserved in 15% glycerol at -80 deg $C$, and a haploid test was performed.

The haploid test was done in liquid-phase with assessment of success based on an $OD_{600}$ threshold of 0.4 read using the Viktor 3 V 1420 multilabel counter (PerkinElmer). Of the 570 strains included in this miniarray, 96.8% were viable and, of those, 94.4% passed the haploid test — i.e. $MAT\ A$ haploids.

### 5.5.2 Screen Protocol

The double mutant miniarray was screened using the protocol detailed in section 5.4.2.

### 5.5.3 Bimodality Test

Strains that displayed bimodal fluorescence distributions in the $rfx1\Delta$ chemogenetic EMAP were streaked onto YPD agar plates, followed by incubation for 24 h at 30 deg $C$. Three colonies of each sample were then patched onto YPD agar plates, and incubated at 30 deg $C$ for 48 h. The patches were then replica plated onto YPD agar plates containing 100 mg/L ClonNat, 200 mg/L G418, URA drop-out, or 100 mg/L L-Canavanine. Replica plates were then incubated at 30 deg $C$ for 3 d.

## 5.6 Outlier Identification

Because we have access to single-cell data, we looked for measures of outlier (or hit) identification that incorporates the distribution shape. For hit selection, we used two scores that rely indirectly on the distribution of fluorescence profiles via the cumulative distribution function (CDF). The first section describes the first score used for the primary screen and choice of strains to be included in the secondary screen. The second section describes the second score used for ranking outliers in the secondary and epistasis miniarray, as well as a measure of fitness for calculating epistasis.

### 5.6.1 Root Mean Square Error (RMSE)

Filtration and gating begin by randomly selecting 1,000 $yEBFP$ events from each strain. These events were then used to construct a total $yEBFP$ distribution, from which the upper and lower limits that subtend 98% of the population are determined using Eq. 5.1 (Fig. 5.4).

$$f(x,\sigma) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad x \in [0,\infty) \tag{5.1}$$

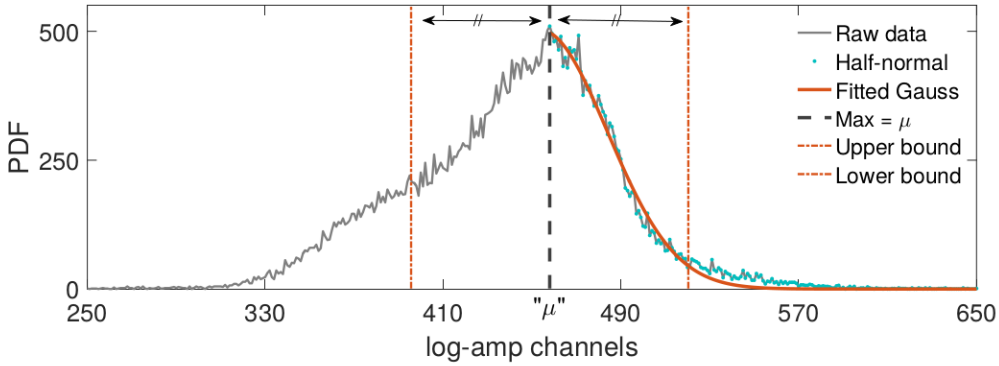



**Figure 5.4:** Strict *yEBFP* cut-off based on half-Normal distribution. The upper-tail of the dis- tribution following the mode is assumed to be half-Normal and fitted as such using MATLAB™. Upon finding the intensity within which 98% of the events fall in the half-Normal curve, the deviation from the mode is used to mirror the threshold.

The *yEBFP* limits are then used to filter out what a cellular *yEBFP* profile is, by applying these limits as thresholds in all samples. The filtered event are then gated using a density-based function on FS and SS. The gate captures 80% of the events around the densest region. Events within this gate are now considered to be cells with high confidence, and their *yEGFP* responses are normalized based on the MMS-induced shift in *yEBFP* (Fig. 5.5), before obtaining the samples' cumulative distributions function (CDF).

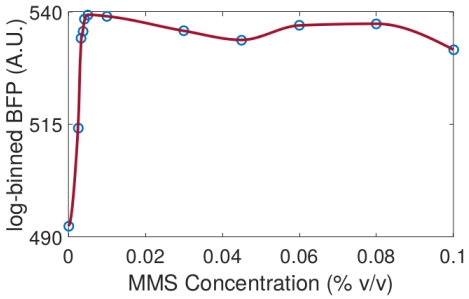



**Figure 5.5:** MMS-dose response of logarithmic-binned *yEBFP* in *BY4741*.

Using the CDF instead of single-value summaries is advantageous as it eliminates the assumption that the yEGFP distributions are parametric. Cumulative distribution functions were then generated for individual samples, as well as for average plate profiles when samples from all sub-plates screened on the same day and at the same drug condition were aggregated.

To score the difference between both the two curves, we used the root-mean- square-error, a measure of accuracy between a theoretical CDF and an empirical one. We sought to take advantage of the measure by designating plate average CDFs as theoretical CDFs, and calculating RMSE as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(CDF\left(GFP_{mut_i} - \overline{CDF\left(GFP\text{PPaverage}_i\right)}\right)^2\right.}{n}} \tag{5.2}$$

where $i$ is a point and $n$ is the total number of points on the CDF of every sample. The RMSE is, therefore, a single-value measure of deviance from the average, with the added advantage that it does not presume that there is an underlying probability distribution to our data.

The individual scores were then normalized by the mean and standard error of WT RMSEs.

### 5.6.2 Area Under Curve (AUC)

One of the main challenges with RMSE as a hit score is that it does not provide information on effect directionality, which limits outlier detection reproducibility (Appendix 1) when MMS is introduced. We looked to a close measure based on CDF that evaluates deviations in general shape and position relative to WTs.

Weighing the cumulative probabilities by their channel values gives a measure that incorporates position of the distribution, with signals on the lower end of fluorescence weighing less than the more intense counterparts. Thus, the area- under-curve (AUC) of a weighted-CDF (wCDF) — i.e. its definite integral — incorporates information about the location and spread of a sample's distribution, and deviation from the mean WT profiles on the same plate (Fig. 5.6).

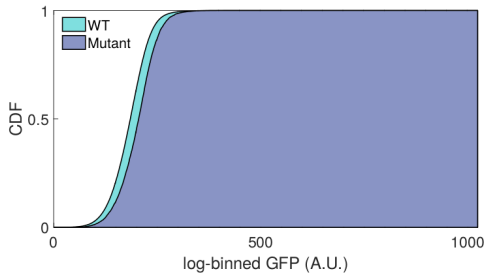



**Figure 5.6:** Area under wCDF curves of WT and prototypical mutant. AUC score subtracts WT AUC from mutant AUC yielding green region. This area is normalized by standard error in WT AUCs.

To evaluate the AUC of CDFs, we first linearized fluorescence data using

$$GFP_{lin} = 0.1024e^{0.008995GFP_{log}} \tag{5.3}$$

Trapezoidal integration of $\text{wCDF}_{\text{lin}}$ was done using

$$\int_a^b f(x)dx = \frac{1}{2}\sum_{n=1}^{N}\left[f\left(x_{n+1}\right) + f\left(x_n\right)\right]\left(x_{n+1} - x_n\right) \tag{5.4}$$

where $f(x)$ is the $CDF(linearGFP)$. Trapezoidal integration essentially divides the AUC of wCDF into trapezoids, calculates their areas, and adds them.

Scores were normalized by the mean and standard error of same-plate WT AUCs.

### 5.6.3 Epistasis Score Analysis

Phenix et al. describe a multiplicative model of epistasis in a haploid system of *S. cerevisiae*. They defined genetic interaction as the fold-change in combinatorial deletion growth fitness relative to those of the individual single knock-outs [68]. Multiplicative epistasis was evaluated using:

$$\epsilon = \log_{10}\left(\frac{F_{dka} \times F_{wt}}{F_{sko} \times F_{rfz1\Delta}}\right) \tag{5.5}$$

where F is the fitness measure (raw AUC), dko is the double knock-out and sko is the single knock-out raw AUC from the secondary screens [68]. Interpretation of the score range is as follows:

$$\epsilon = \begin{cases} > 1 & \text{strong antagonism} \\ 0 < \epsilon \leq 1 & \text{weak antagonism} \\ \approx 0 & \text{independence} \\ -1 \leq \epsilon < 0 & \text{weak synergy} \\ < -1 & \text{strong synergy} \end{cases} \tag{5.6}$$

We used Eq 5.5 to evaluate epistasis in RFX1 double mutants because of the score's success and consistency in predicting the hierarchy of *GAL* genes in their products' pathway [68]. MMS-dependent changes in interaction were then evaluated by subtracting $\epsilon$-scores of samples in the untreated group from their $\epsilon$-scores at 0.015% MMS.

### 5.6.4 Functional Enrichment

Due to observed discrepancies in GO term allocation between multiple Gene Ontology databases, functional enrichment was done using MATLAB™. The curated Saccharomyces Genome Database (SGD) GO database was used as the background against which enrichment was evaluated using the modified Fisher's exact test:

$$P = (x|M, K, N) = 1 - \sum_{i=0}^{x} \frac{\binom{K}{i}\binom{M-K}{N-i}}{\binom{M}{N}} \tag{5.7}$$

The cumulative hypergeometric distribution (Eq. 5.6) describes the probability that one randomly gets more than or equal to $x$ query genes matching to a GO category (desired success rate) if one sampled the genome $N$ times (sample size), given the relative size of the GO category ($K$) to the genome size ($M$).

# 6 | Results

## 6.1 Chemogenetic screens of DMA and DAmP libraries reveal differential effectors of *RNR3pr* activity

The sequence and transcriptional dynamics of *RNR3* and its promoter are relatively well-characterized. Previous studies reported the gene's nucleosome map, transcriptional repressors (and interactions amongst which), repression sites, as well as its product's structure, regulation, and kinetics. *RNR3*, and specifically its promoter, is an attractive detector system because it is expressed preferentially during cellular damage, and is downstream of the replication and DNA damage checkpoints.

Transcription at the promoter site correlates with the activation of the replication and DNA damage checkpoints, and Phenix *et al.* found that *RNR3* mRNA levels increase in a dose-dependent manner, peaking at ≈50-fold increase when cells are exposed to the potent alkylating agent, MMS. Lundin et al. explored the effect of MMS and how the drug might be mediating genotoxicity. They postulate that the mechanism of MMS-induced genotoxicity is mainly due to the sensitization of the damaged sequence to heat, and replication fork stalling [54]. Both outcomes increase the risk of SSBs and DSBs, which should activate the replication and DNA damage checkpoints, respectively. Because these checkpoints act as highly conserved hubs of signal coordination relaying DNA health status to cell-cycle progression, we linked the *RNR3pr* to *yEGFP* to serve

as a reporter of checkpoint activation. This way, we measure the significance of certain genes to checkpoint activation by proxy.

We screened a collection of 5,850 strains in the DMA and DAmP libraries, each of which harbours the genotype signature of *BY4741* (Table 3) and the reporter at the *MET15* locus(Fig. 6.1).

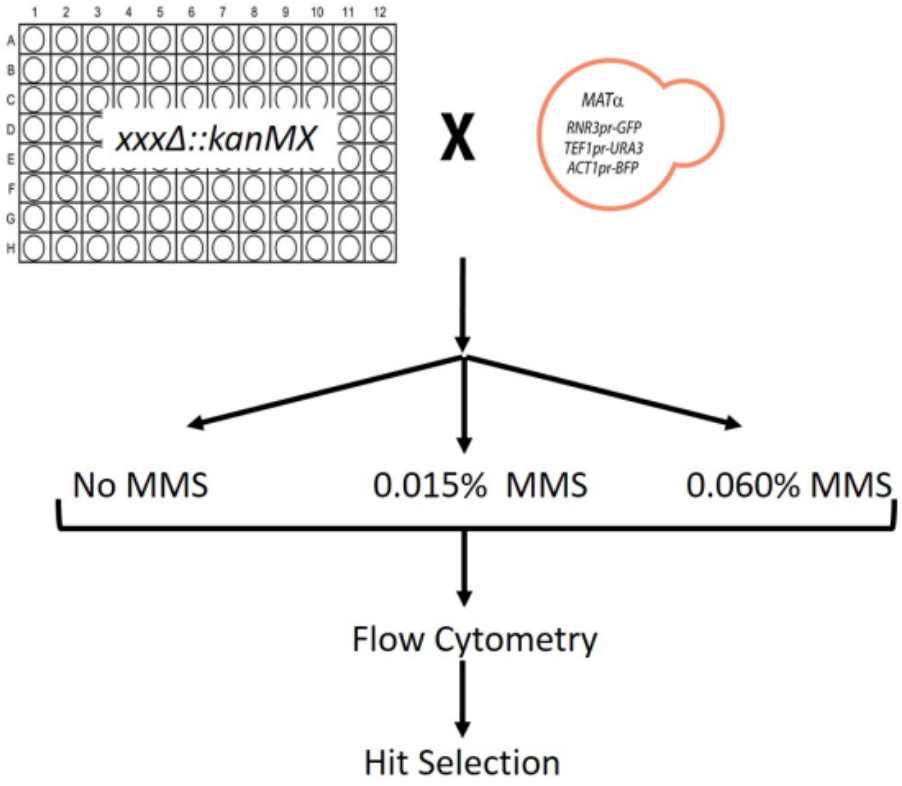



**Figure 6.1:** Schematic of secondary protocol flow. The reporter-carrying query was mated with mutants in the commercial DMA and DAmP libraries to generate the library used for screening. The library was first screened at two MMS conditions, 0.025% and 0.055%. A subset of the library was then re-screened more rigorously at three conditions: no drug, 0.015%, and 0.060% (v/v).

Analysis of primary screen data was challenging for three reasons. First, Cyan ADP is sensitive, detecting debris and random perturbations as events. The events observed are therefore not solely cellular, and we need to find a way to distinguish what a 'cell' is from a 'non-cell'. Second, during protocol development, it appeared that de-condensation

of the 384-well format of the yeast library is required. One reason being that there was an observed increase in both the mean *yEGFP* and mean *yEBFP* intensities over the 1.5 hours required to screen a 384-well plate (master plate). This is probably due to the accumulation of both fluorescent protein over time within the cells, leading to a shift in the fraction of cells expressing at high intensities. De-condensation into 96 strains/plate (called sub-plates) allowed the minimization of this shift; unfortunately, it also introduced the added challenge of not having the WTs on the sub-plates, because of the way the library was originally constructed. Finally, commonly-used screen scores are usually single-value summaries, through which we hope to describe the shape of the graph. These summaries include the coefficient of variation (CV; standard deviation divided by the mean), mean, median, interquartile dispersion (QCD; interquartile range divided by the median), and fold-change. After attempting to analyze and rank (Z-score) the primary screen's strains based on these measures, we observed that they do not adequately capture visibly extreme outliers. This is probably because these measures are based on an assumed underlying probability distribution function to the *yEGFP* response — i.e. the assumption that the distributions are parametric.

Solutions to all three challenges were incorporated in filtering, gating, and ranking MATLAB™scripts developed by Kærn (2014). The scripts apply gates based on FS, SS, and BFP followed by root-square-mean-error (RMSE) scoring and ranking (see 5.6). The resulting list of potential hits were re-screened more rigorously.

Strains selected for the secondary screen potential false negatives that belong to the same complex as, or are known physical interactors or regulators of the potential hits set. In total, 819 potential outliers were screened using the secondary screen protocol described in section 5.4. We screened the strains at three MMS conditions: intermediate (0.060%), low (0.015%), and none. The inclusion of no treatment here allows us to assess basal (MMS-independent) and differential (MMS-dependent) effects.

### 6.1.1 Benchmarks

Untreated replicates ($n = 86$) showed little correlation with both RMSE and AUC scores(Fig. 6.2a). Scores based on AUC of linear *yEGFP* CDFs showed strong reproducibility($r^2 = 0.9216$ and 0.7181) at low and intermediate MMS doses, respectively (Fig. 6.2b & c ). It is worth noting that at the lower end of fluorescence decades, the flow cytometer is sensitive and susceptible to fluctuations that may register as significant. To see if the low correlation is a result of score processing, Fig. 6.2d shows the fluorescence correlation of replicate samples. Plate-to-plate variability in the four WT distributions is minor at 0%, and increases with addition of MMS (Fig. 6.2e)

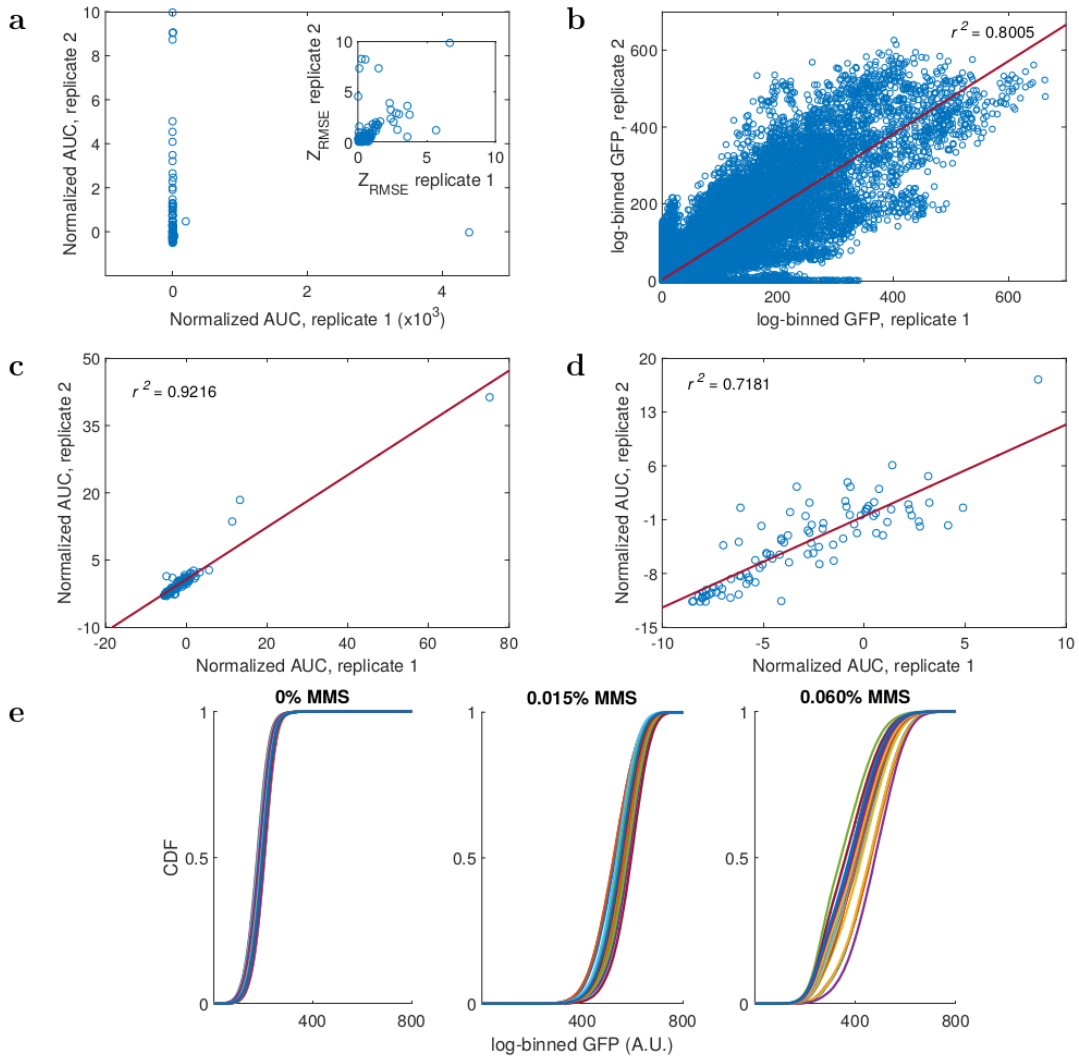



**Figure 6.2:** Reproducibility of secondary screen samples ($n = 86$) at three MMS conditions. (a—d) Simple linear regression of independent replicates at (a) 0%, (b) 0.015%, and (c) 0.060% MMS (v/v). (d) Fluorescence intensities and linear regression of replicate samples (e) Pooled WT *Saccharomyces cerevisiae* cumulative distribution functions($n = 36$) at the three drug conditions.

The screens are primarily exploratory. We expected the following to appear as outliers in the primary screen results:

1. Genes involved in DNA damage sensing — e.g. *DDC2*, *RAD9*, *MRC1*, *DDC1*, *RAD17*, *MEC3*, and *RAD24* — will lower reporter expression when deleted;

2. Disruption of the Ino80 complex will lower reporter expression, because of its role in enhancing Rad53p phosphorylation;

3. Deletion of the *RFX1*, *ROX1*, and *MOT3* would derepress *RNR3pr* further leading to an enhanced phenotype.

These expectations serve as a benchmark of the agreement of our results with the literature.

*RAD17, MEC3,* and *DDC1* encode members of the PCNA-like clamp (9-1-1 in humans) involved in checkpoint activation and anchoring of checkpoint and repair machinery; *RAD24* encodes a subunit of a loader that loads the PCNA-like clamp onto DNA. The deletion of the four sensors therefore should inhibit the ability of the PCNA-like clamp to activate the replication checkpoint. The current model of the checkpoint puts the PCNA-like clamp upstream of Mec1p, with the former activating the kinases via Dpb1p (rev. in [64]). Our results, however, display residual reporter expression suggesting that the damage checkpoint(s) is activated (Fig. 6.3). We have not evaluated Rad53p phosphorylation in these mutants, but it is likely that the DDC branch (Tel1p-mediated) of DDR is acting as a redundant effector of *RNR3*, and thereby *yEGFP*, expression.

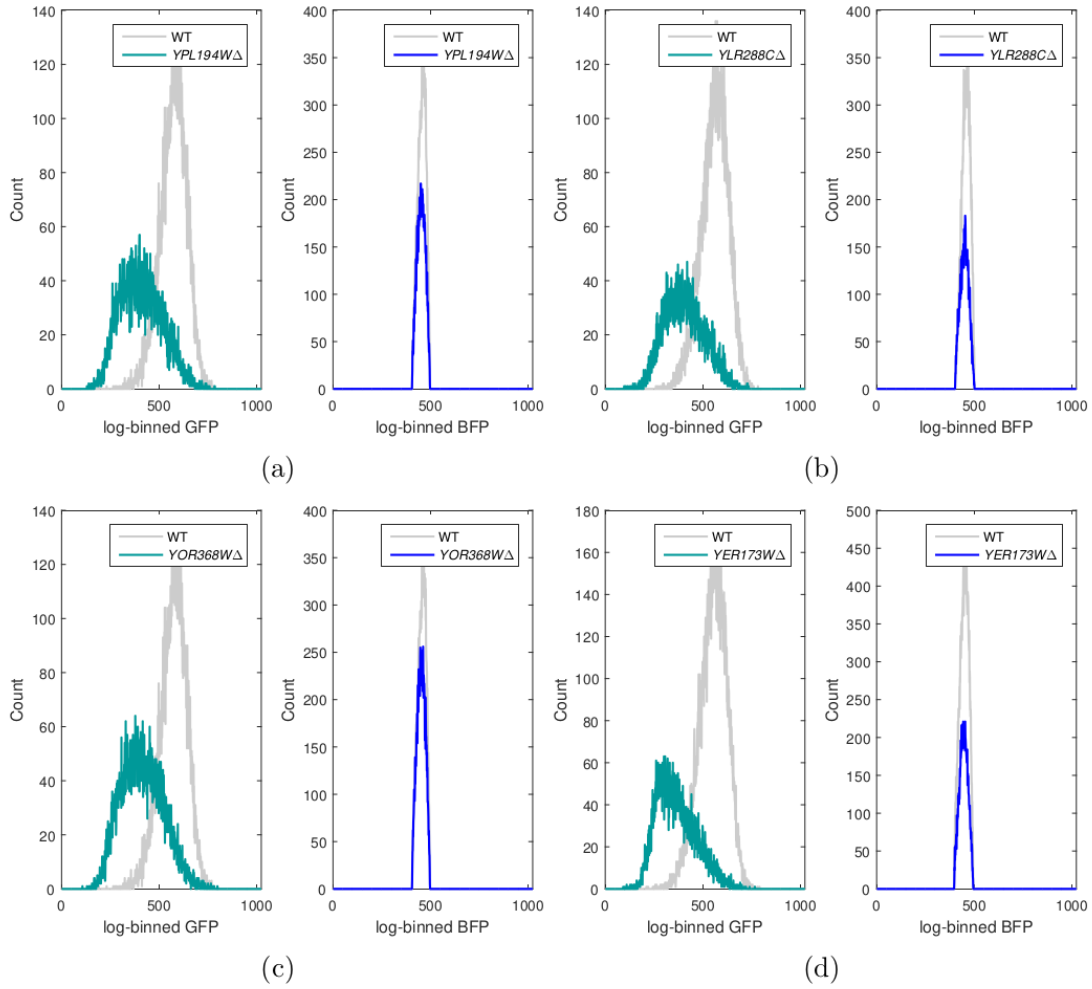



**Figure 6.3:** Fluorescence profiles of the DNA damage response sensors mutants (a) *ddc*1Δ, (b) *mec*3Δ, (c) *rad*17Δ, and (d) *rad*24Δ at 0.015% MMS. WT denotes *ydl*227*c*Δ. Samples were gated based on FS, SS, and *yEBFP* parameters of aggregated WTs per plate.

Surprisingly, reporter expression of each of the four sensor mutants at 0.060% resembles that of WTs (Fig. 6.4), suggesting that more extensive damage affects *RNR3* expression equally between the mutants and WTs.

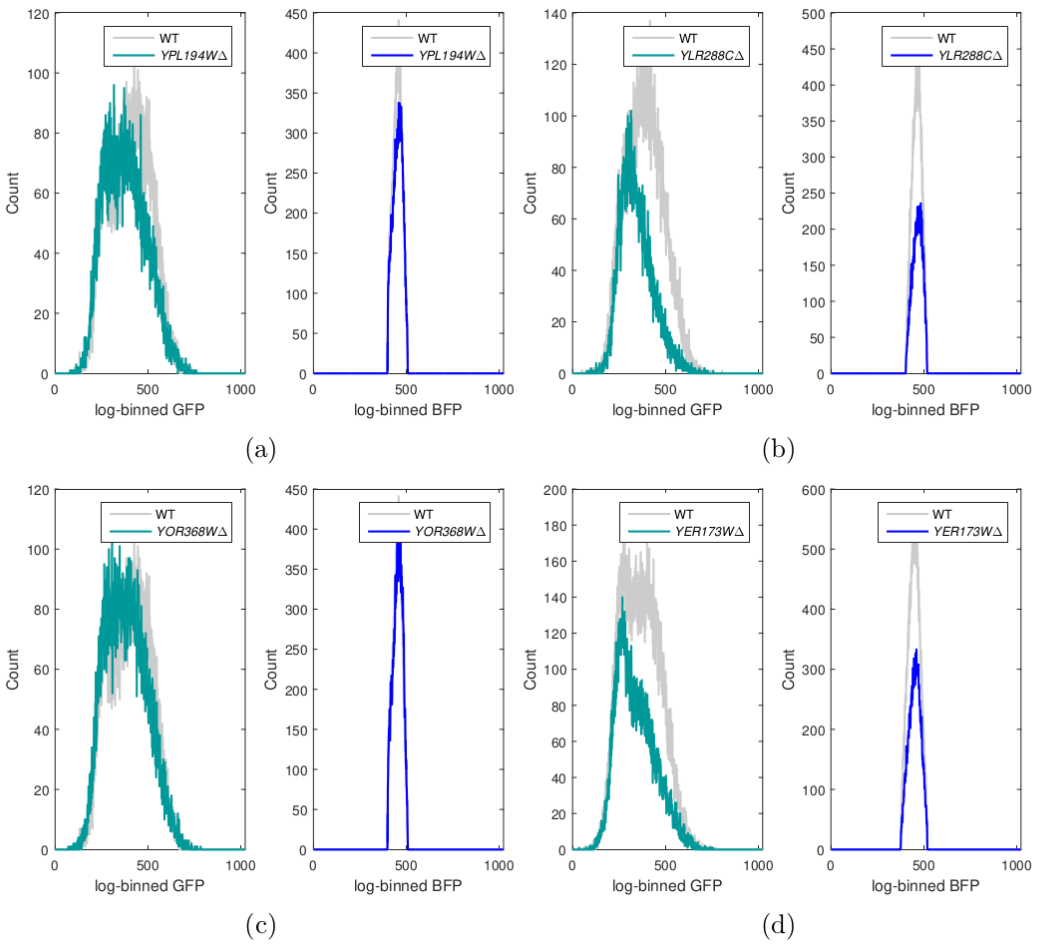



**Figure 6.4:** Fluorescence profiles of the DNA damage response sensors mutants (a) *ddc*1Δ, (b) *mec*3Δ, (c) *rad*17Δ, and (d) *rad*24Δ at 0.060% MMS. WT denotes *ydl*227*c*Δ.

Recall that *RNR3* expression is cell-cycle dependent [29]. In the presence of excessive damage, a greater proportion of cells may be arrested at a cell-cycle stage in which *RNR3* expression ceases. If gene deletions do not fundamentally alter promoter responsiveness to the stimulus or cell-cycle progression, it is reasonable to expect that fluorescence intensity levels at arrested stages resemble those of WTs.

We also expected that disruptions to the Ino80 complex will reduce *RNR3* reporter expression due to impaired Rad53p phosphorylation or the lack of HDAC activity by Ino80 — i.e. *RNR3* specific or non-specific effects. Our results support the expectation (Fig. 6.5). *IES1* encodes Ies1p, a member of the Nhp10 body module of the Ino80 complex. The Nhp10 module is the least conserved component of the Ino80 complex:

its deletion leads to minor impairment in Ino80 affinity to nucleosomes, as well as in remodeling [80]. Our results agree with the nonessential role that Ies1p plays in the Ino80 complex (Fig 6.5a). Similarly, the deletion of *IES2* and *ARP5* appear to be nonessential, with only *ies*2$\Delta$ conferring a minor reduction in *RNR3* reporter expression (Fig. 6.5b & c). Deletion of *IES4*, enhancer of Rad53p hyperphosphorylation in response to Mec1p/Tel1p activation, did not adversely affect *RNR3*-reporter expression (Fig. 6.5d), suggesting that Rad53p hyperphosphorylation can be predominantly mediated without Ino80 involvement. The presence of gated event counts implies that *IES4* is not essential to Ino80 function. Belonging to the same Arp8 foot module as Ies4p, Taf14p (encoded by *TAF14*) appears to play a more essential role in survival (Fig. 6.5e). Its deletion clearly compromises gated event count, supporting observations that the loss of the Arp8 module confers global remodeling impairment by the Ino80 complex [80].

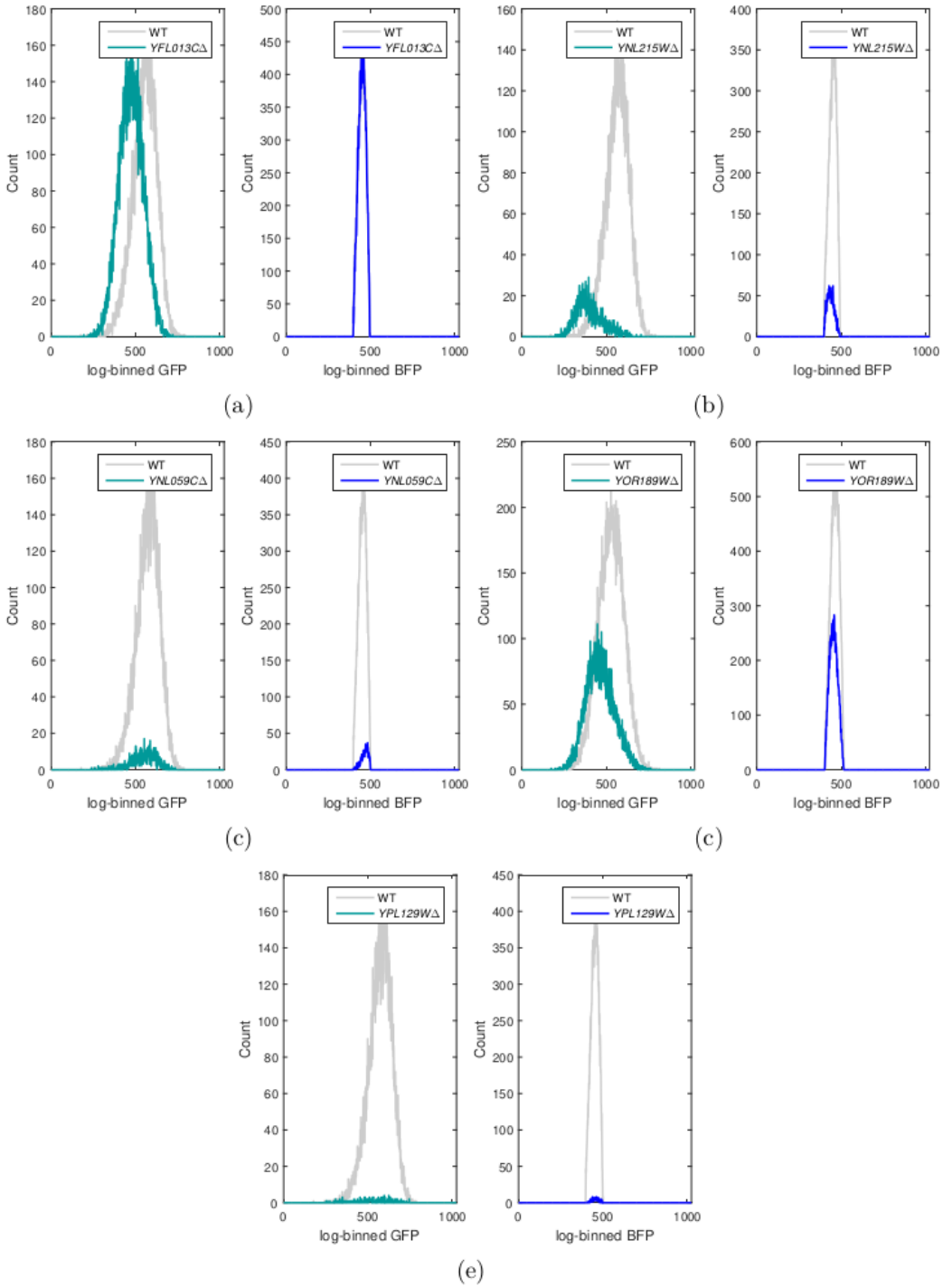



**Figure 6.5:** Fluorescence profiles of the Ino80 complex mutants (a) $ies1\Delta$, (b) $ies2\Delta$, (c) $arp5\Delta$ and (d) $taf14\Delta$ at 0.015% MMS.

Finally, we expected that deletion of the synergistic repressors of the *RNR3pr*, Rfx1p, Rox1p, or Mot3p, to emerge as enhancers of *RNR3* expression at all drug conditions [46]. In the absence of MMS, however, only $rfx1\Delta$ and $rox1\Delta$ appear to erepress RNR3

reporter expression (Fig. 6.6a & b). The strength of derepression obeserved in our screens is in line with the trend reported by Klinkenberg *et al.* [46], wherein Rfx1p is the strongest repressor and Mot3p is the weakest. The derepression associated with $rox1\Delta$ also supports the contribution of Rox1p in recruiting the global repressor Ssn6p-Tup1p [60].

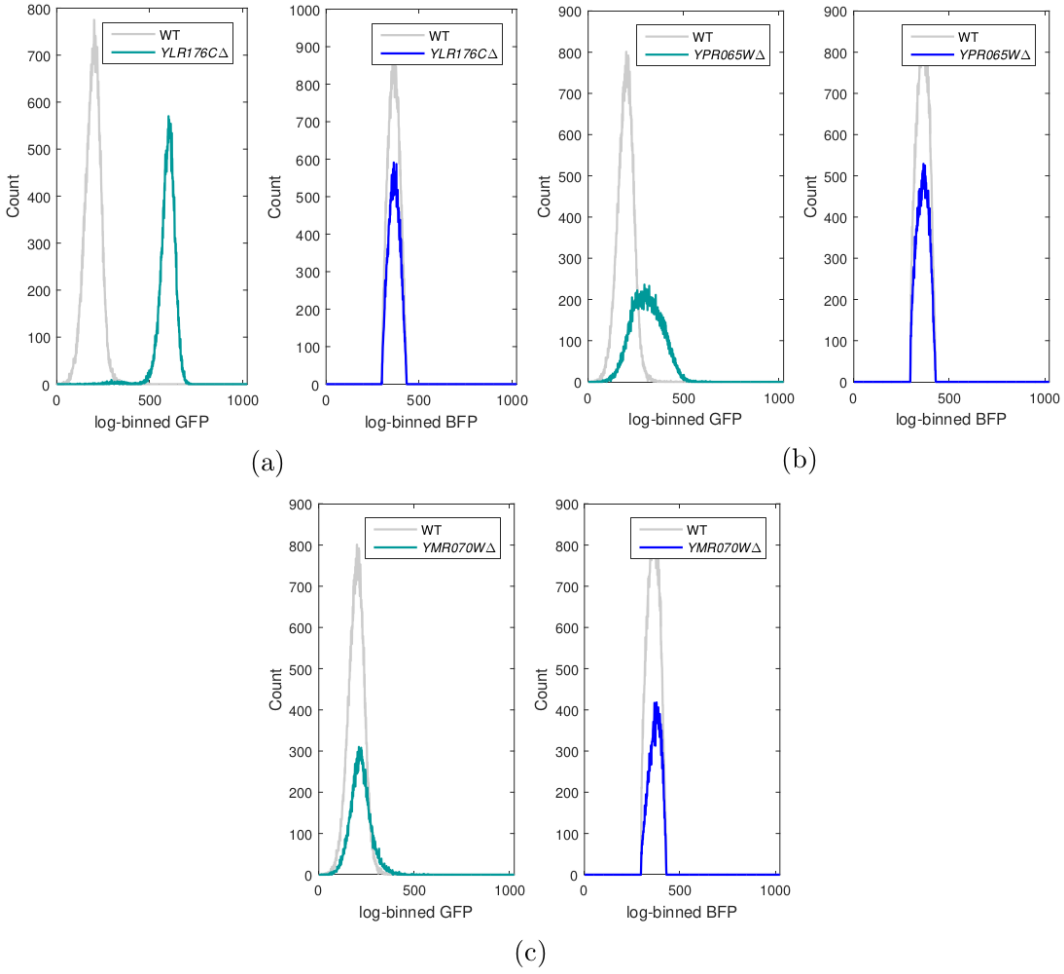



**Figure 6.6:** Fluorescence profiles of RNR3pr repressor mutants (a) $rfx1\Delta$, (b) $rox1\Delta$, and (c) $mot3\Delta$ in the absence of MMS.

Evaluated at low and intermediate MMS, we observe dose-dependence in repression, especially by $rox1\Delta$ at 0.015% and $mot3\Delta$ at 0.060% MMS (Fig. 6.7).

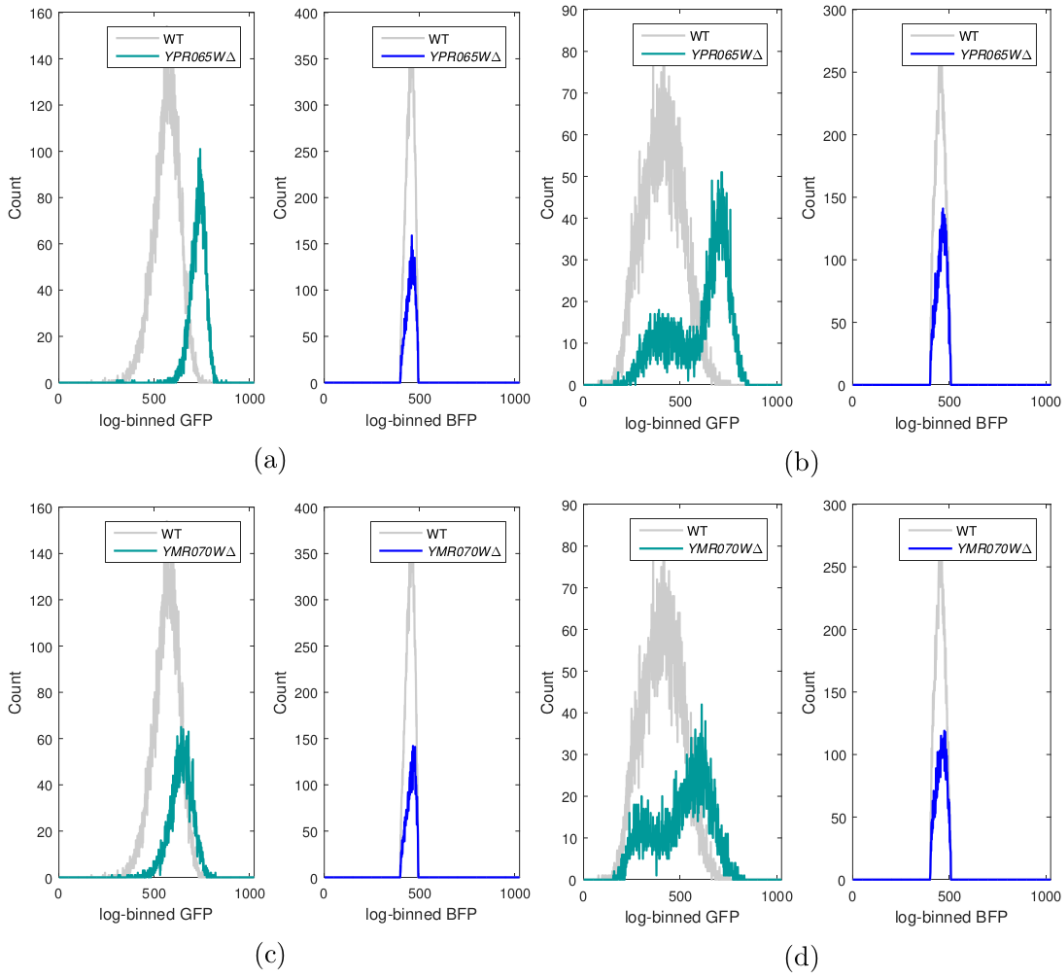



**Figure 6.7:** Derepression of *RNR3pr* by (a—b) *ROX1* and (c—d) *MOT3* deletion is MMS dose- dependent. (a & c) 0.015% MMS; (b & d) 0.060% MMS.

Of interest is the behaviour of *mot*3Δ in response to MMS. While exposure to low MMS clearly shifted the distribution to the right, there appears to be a low- expressing subpopulation that persists with exposure to intermediate MMS (Fig. 6.7b & d). PCR-based validation of *mot*3Δ shows no amplicon upon extracting the strain's genome (Appendix 3). This suggests that the bimodality observed here is not a result of a mixed population. Klinkenberg *et al.* reported the potency of the three repressors in an *RNR3pr-lacZ* system previously. According to the results of their epistasis analysis, the repressors rank as *RFX1*, *ROX1*, followed by *MOT3* in terms of their ability to repress *RNR3pr*. Our results at no and low MMS agree with their ranking; at intermediate MMS, *ROX1* slightly overtakes *RFX1*.

Given the reproducibility of our results, inferred from replicate sample and WT distributions (Fig. 6.2), as well as agreement with pertinent literature, we are confident in the quality of our screen data.

### 6.1.2 Chemogenetic Screens Results

Ranking based on normalized AUC ($Z_{AUC}$) yielded persistent and differential outliers that affect reporter phenotype but not viability (inferred as gated event count $>$ 500 events) (Fig. 6.8 and 6.9).

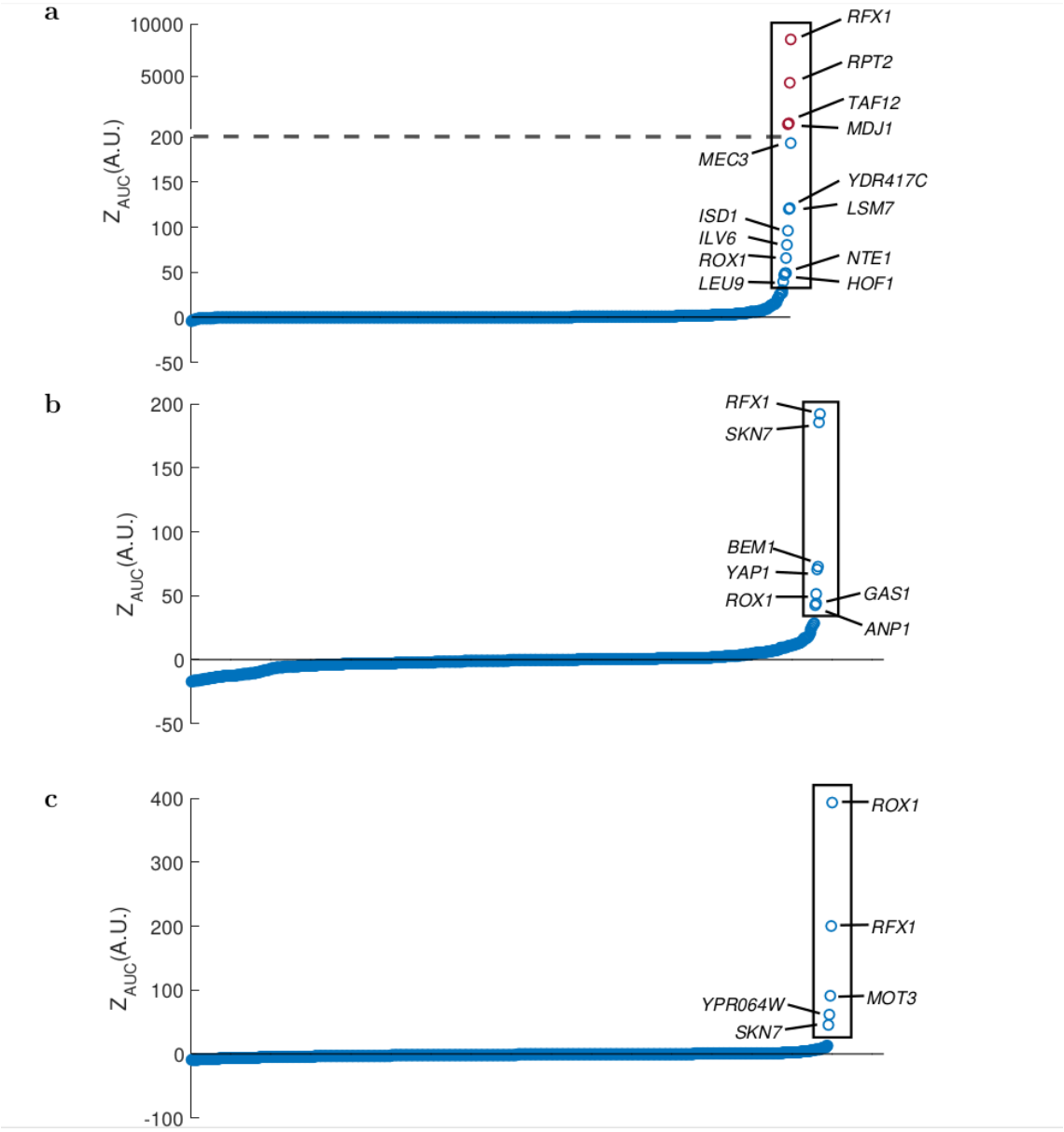



**Figure 6.8:** $Z_{AUC}$ -ranked strains at (a) 0%, (b) 0.015%, and (c) 0.060% MMS. Channel-weighted scores normalized to mean raw AUC of WTs at 0% MMS yielded very high scorers; break in y-axis (dashed line) used to visualize outliers whose effects would be masked by *RFX1*, *RPT2*, *TAF1*, and *MDJ1* scores.

Out of the 812 strains in the secondary screen, three are detected as outliers that enhance reporter expression in the absence of MMS-induced damage (Fig. 6.9).

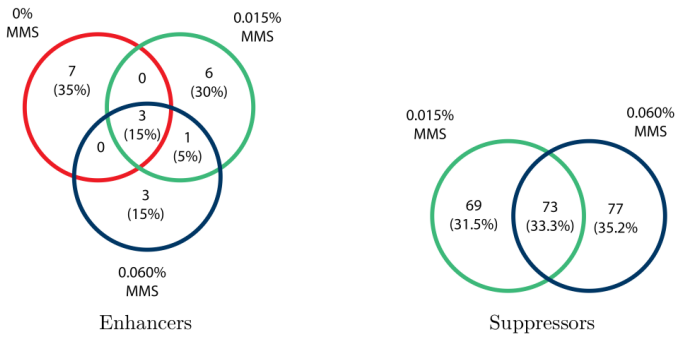



**Figure 6.9:** Overlap of outlier sets across the different MMS conditions.

As expected two of the *RNR3* (as well as *RNR2* and *4*) repressors, *RFX1* and *ROX1*, are among the enhancers. The remaining enhancer, *MDJ1*, encodes a chaperone protein that mediates proteasomal degradation of misfolded proteins [72, 84] (Fig. 6.10).

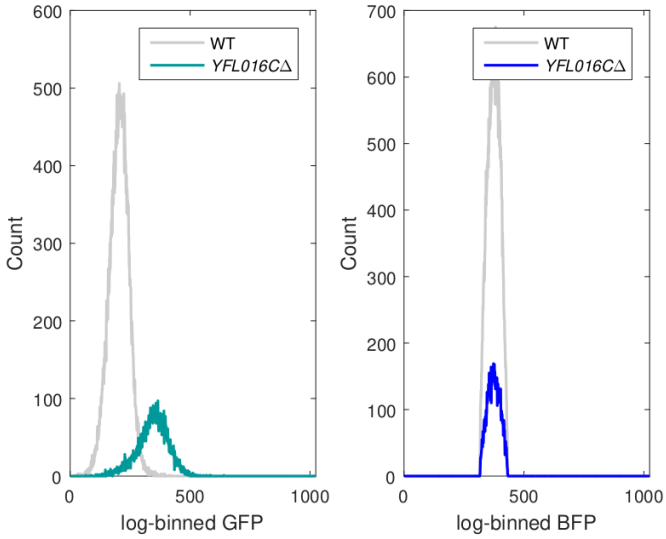



**Figure 6.10:** Fluorescence profile of *mdj*1Δ at 0% MMS.

The remaining enhancers (according to the AUC score) are involved in ribosome biogenesis (*ILV6*), proteasomal degradation (*RPT2*), chromatin remodeling (*TAF12*), oxidative stress response (*ISD11*), mRNA decay (*LSM7*), replication and DNA damage sensing and repair (*MEC3*), and dubious open reading frame (ORF) that overlaps with *RPL12B*. The effects of the remaining enhancers are minimal upon examining their fluorescence profiles, however. They appear as hits most likely because most fluorescence distributions are at the lower end of the flow cytometer's detection limit, which is sensitive

to minor fluctuations. For this reason, no significant phenotype suppressors are detected at 0% MMS.

The SGA score of *mdj*1Δ, *ilv*6Δ, and *mec*3Δ suggests that they enhance *yEGFP* expression when untreated but suppress the phenotype when treated with low MMS (Fig. 6.11). Their suppressive effects are moderate when exposed to intermediate MMS.

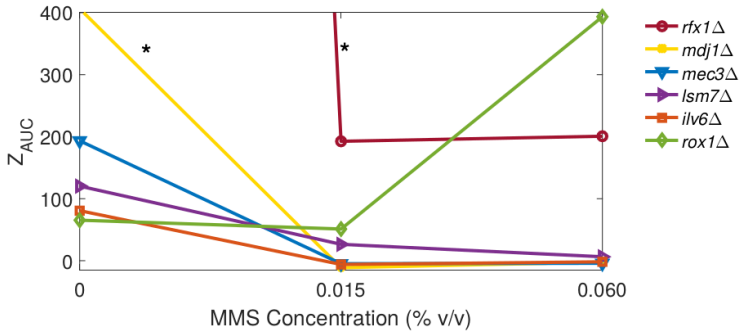



**Figure 6.11:** $Z_{AUC}$-ranked phenotype enhancers' dose-dependent change in rank. Asterisks indicate that scores at 0% MMS were too high to adequately present them without masking the rank trends.

### 6.1.3 The Hendry Dataset

Hendry *et al.* published a paper with the same objectives [35]. The main differences between our experimental designs are:

**The high-throughput screening devices:** we used flow cytometry to gain information on reporter expression at the single-cell level. Their study is centered around two microscopy platforms: Typhoon Trio Variable Mode Imager [35], and EVOTEC Opera confocal microscope system [34]. Both microscopy systems measure population-level fluorescences.

**Screening phases:** our protocol was developed to use liquid-phase growth and drug treatment; they explored both solid and liquid phases. Their liquid protocol resulted in more similarities to our hits set [34].

**Drug conditions:** we treat cells with two doses of MMS that represent the peak and one of the descent points of the dose response curve. In their published article, cells were treated with one dose that confers close to maximal *RNR3* reporter expression (Fig. 5.2).

**Scoring:** AUC is used for ranking after normalizing to WTs' scores. Strains with less than 500 gated event counts were discarded from the analysis. Their score is a binary logarithm of the fold-change in *RNR3* reporter fluorescence relative to their internal control (*tdTomato* driven by *RPL39pr*). Colonies with areas $< 500$ pixels were discarded.

**Temperature-sensitive alleles:** they have access to these alleles enabling them to study a portion of essential genes.

Because *RNR3pr-GFP* expression enhancement is more easily explained as derepression, we were interested in the hits reported to increase *RNR3* expression ($n = 160$). Although their hits were assayed at the three MMS conditions of our secondary screen, the low dose is of interest because it facilitates comparison as both 0.015% and 0.03% MMS lead to maximal *yEGFP* (Fig. 5.2). We found that maximum reporter enhancement was achieved by $rad55\Delta$, $psy3\Delta$, $sae2\Delta$ and $ssk1\Delta$ at low MMS (Fig. 6.12).

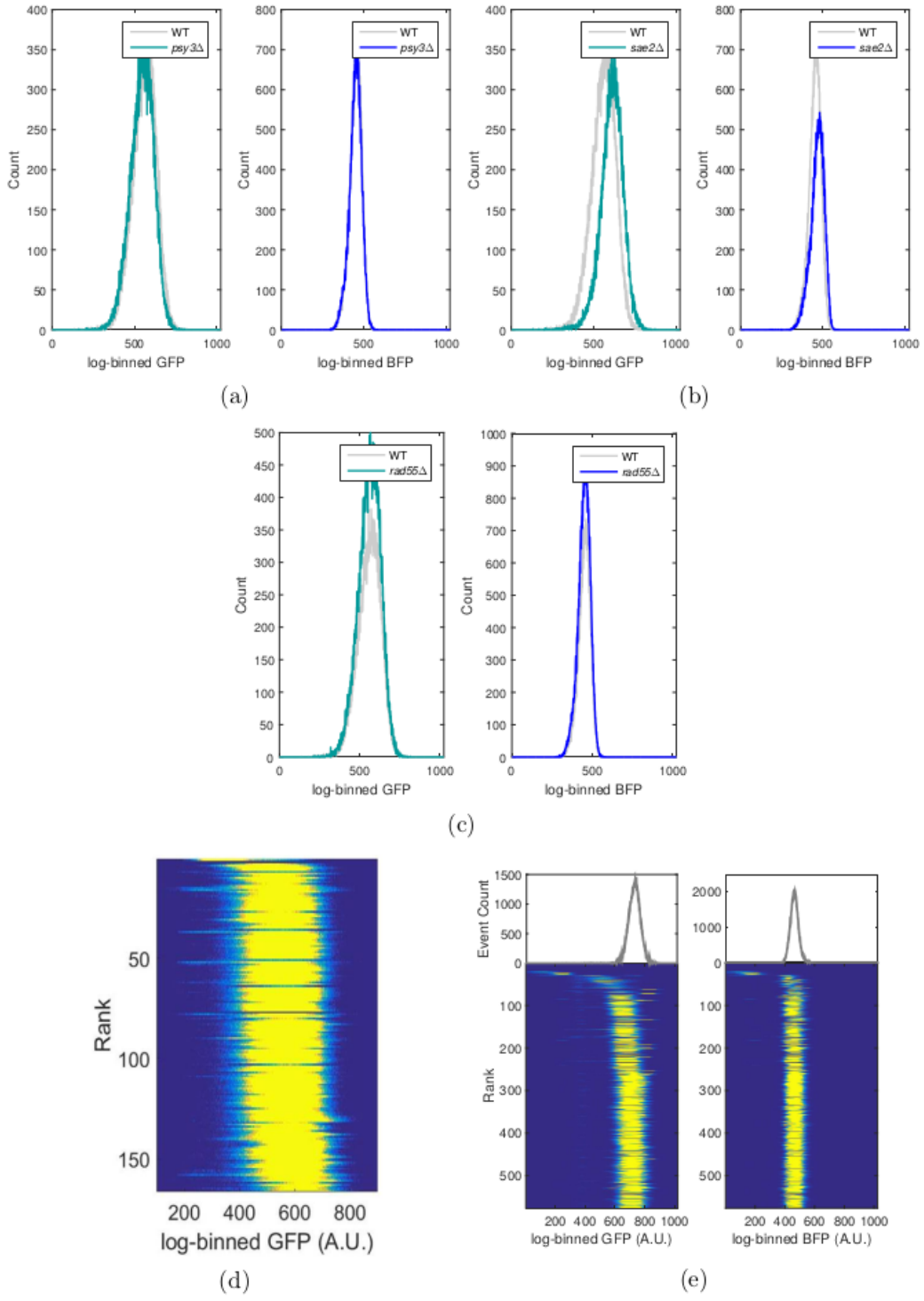



**Figure 6.12:** Phenotype enhancers in the Hendry data set screened at 0.015% MMS using our secondary screen protocol. All *yEGFP* profiles were normalized by subtracting their respective *yEBFP* shifts (**top left**). Distributions of the hits that enhanced RNR3 the most, (a) *psy*3Δ, (b) *sae*2Δ, and (c) *rad*55Δ, respectively. (d) & (e): A comparison of the maximum deviation achieved in the Hendry dataset outliers (d) and those found in our data set (e) at 0.015% MMS.

The remaining strains did not enhance *RNR3* expression relative to WT. Two gene deletions, *srv*2Δ and *dot*1Δ, reduced *RNR3* expression (Fig. 6.13); the remaining strains did not differ in their distributions from WT.

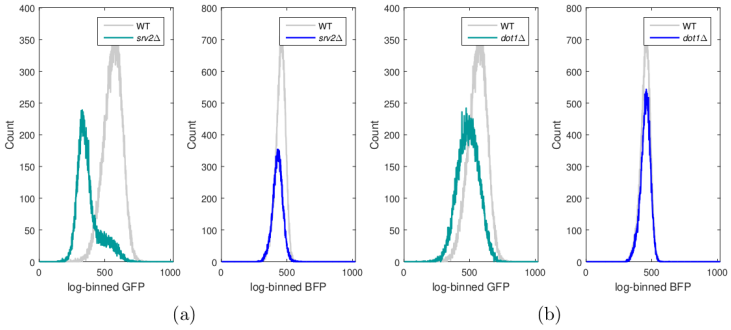



**Figure 6.13:** Two phenotype enhancers in the Hendry data set appear as suppressors at 0.015% MMS. (a) *srv*2Δ; (b) *dot*1Δ.

We expect that the sources of differences between our results are a consequence of all of the following:

- OD control on our part: Population density appears to have an effect on the fluorescence achieved at different MMS doses (Appendix 7). We controlled population density at release of overnight cultures, prior to MMS addition, and by assaying 96 samples at a time to minimize doubling effects during flow-cytometry screening;
- Choice of internal control: The *ACT1pr* is a more constitutive promoter than *RPL39pr*, which interestingly enough is inducible during oxidative stress. *ACT1pr*, however, carries the risk of being silenced during growth defects/slow growth as *ACT1* encodes actin;
- Solid vs. liquid media: Uneven exposure of colonies to nutrients on solid media (plate-edge effects) will further contribute to the population density effect on fluorescence capacity. Moreover, the solid drug plates carry the risk of uneven MMS exposure as well, with cells at the top of the colony exposed less to the drug than those at the bottom. In fact, a colony may not be a synchronized collection of yeast cells for the very reason that there is a difference in exposure to plate contents between different colony zones.

- Statistical analysis: normalization based on our respective internal controls, sensitivity of the instruments, and reproducibility of hit list given a partic- ular method of scoring all play an important role here. Case in point, the difference between the reproducibility obtained in replicates from an RMSE or an AUC score is different (compare Fig. 6.2a—d to Appendix 1).

### 6.1.4 Functional Enrichment

With MMS exposure, the majority of outliers — 150/160 at 0.015% and 150/173 at 0.060% MMS — were phenotype suppressors (Fig. 6.9). Separating the outliers per drug condition and effect direction (increase or decrease), we explored the enriched functional modules in each snapshot. Due to inconsistencies in gene ontology (GO) term allocation between different GO engines, functional enrich- ment was done using MATLAB™.

The GO SLIM database was chosen as an alternative to the complete GO association database, because the latter produces many highly-specific GO categories composed of few members ($< 10$). A result of this specificity is categories with significant enrichment that do not provide meaning in terms of identifying the global modules/processes affecting *RNR3pr*. Functional enrichment was evaluated for all three GO aspects: molecular function (MF), biological process (BP), and cellular component (CC) Categories with more than 1700 elements were removed from the analysis as they are too broad to be insightful.

There are not many enrichment GO terms at 0% MMS, primarily because the list of outliers (all enhancers) comprised of ten ORFs. The most enriched biological process ($P = 7.71 \times 10^{-3}$) was RNA catabolic process (GO:0006401) followed by nucleic acid binding transcription factor activity (GO:0001071) (Fig. 6.14a). Addition of 0.015% MMS resulted in both enhancer and suppressor phenotype relative to WTs. Enhancers were enriched in genes implicated in three broad processes: transcription, stress responses, and morphogenesis (Fig. 6.14b & c). Transcription GO categories are the most enriched with transcription from RNA Pol II promoter, nucleic acid binding transcription factor activity,

and DNA binding categories. Enriched stress response categories include response to oxidative stress, chemical, and heat, as well as signal transduction activity. The majority of enriched modules in the phenotypic suppressors' set falls broadly into two categories:

1. Chromatin dynamics: remodeling, histone modification, chromatin organization;

2. Translation: structural constituent of ribosome, ribosomal biogenesis and assembly, tRNA processing, rRNA processing.

Comparison of low and intermediate MMS enrichment results reveals global changes in modules. At 0.060% MMS, the list of enhancers is enriched in chromatin remodeling GO terms, e.g. histone binding, modification, chromatin organization (Fig. 6.14d & e).

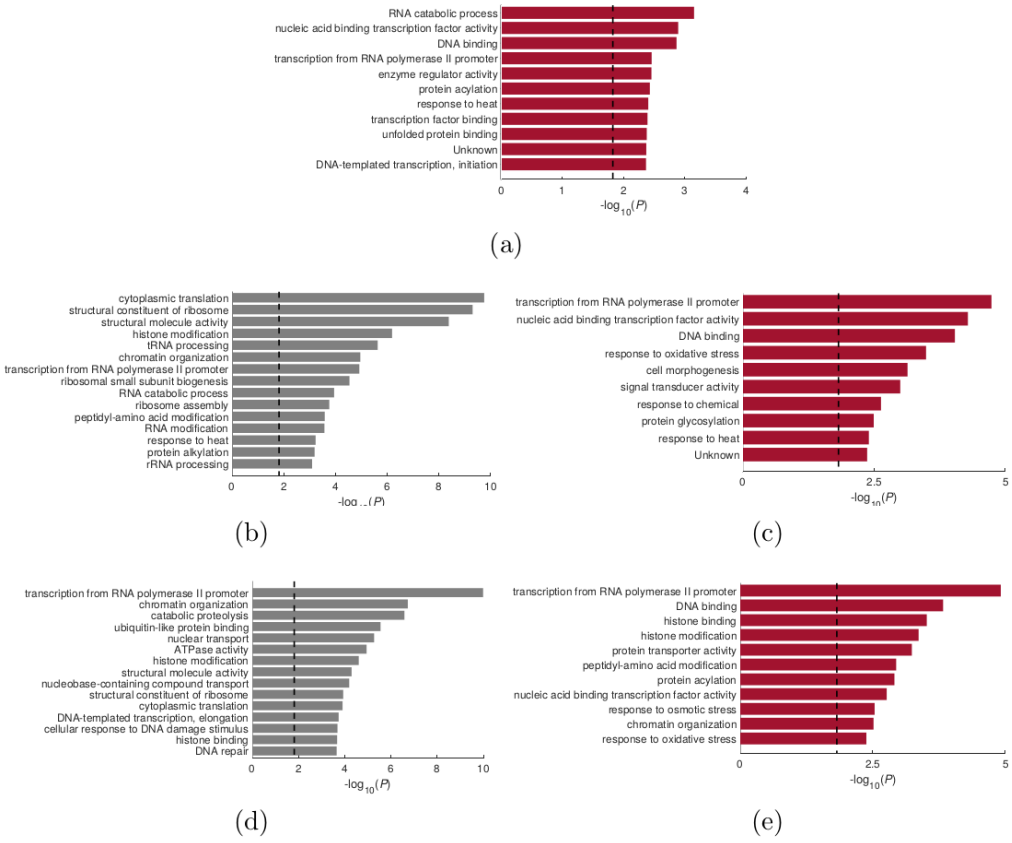



**Figure 6.14:** Functional enrichment analysis of enhancers and suppressors at the three drug conditions.(a) Only enhancers (n = 10)were evaluated for 0% MMS. (b and c) Suppressors (b) and enhancers (c) at 0.015% MMS were analyzed for the top 160 outliers.(d and e) Top 173 suppressors (d) and enhancers (e) at 0.060% MMS analyzed. Dashed line indicates $-\log_{10}(P_{hypergeometric})$ at which $P_{hypergeometric} = 0.05$.

Enriched processes and complexes are involved in translation (ribosome biogenesis and tRNA processing), chromatin remodeling, and histone modification (Appendix 5 Fig. 3). Four members of the histone acetyltransferase (HAT) SAGA were detected as suppressors at low doses of MMS. The Spt-Ada-Gcn5-Acetylatransferase (SAGA) complex is responsible for histone H3 acetylation that mediates unravelling of silenced regions for expression [19, 59]. *RNR3pr*-driven ex- pression, for example, is regulated by SAGA-mediated recruitment of TATA-box binding protein (TBP) [13, 92]. Therefore, deletion of SAGA members, Ngg1p, Spt8p, Sgf73p, or Gcn5p, is expected to suppress both the reporter and the internal control — i.e. non-specific effects (Fig. 6.15). While all four mutants reduced *RNR3* expression, only *ngg*1$\Delta$ has an effect on the internal control (Fig. 6.15b).

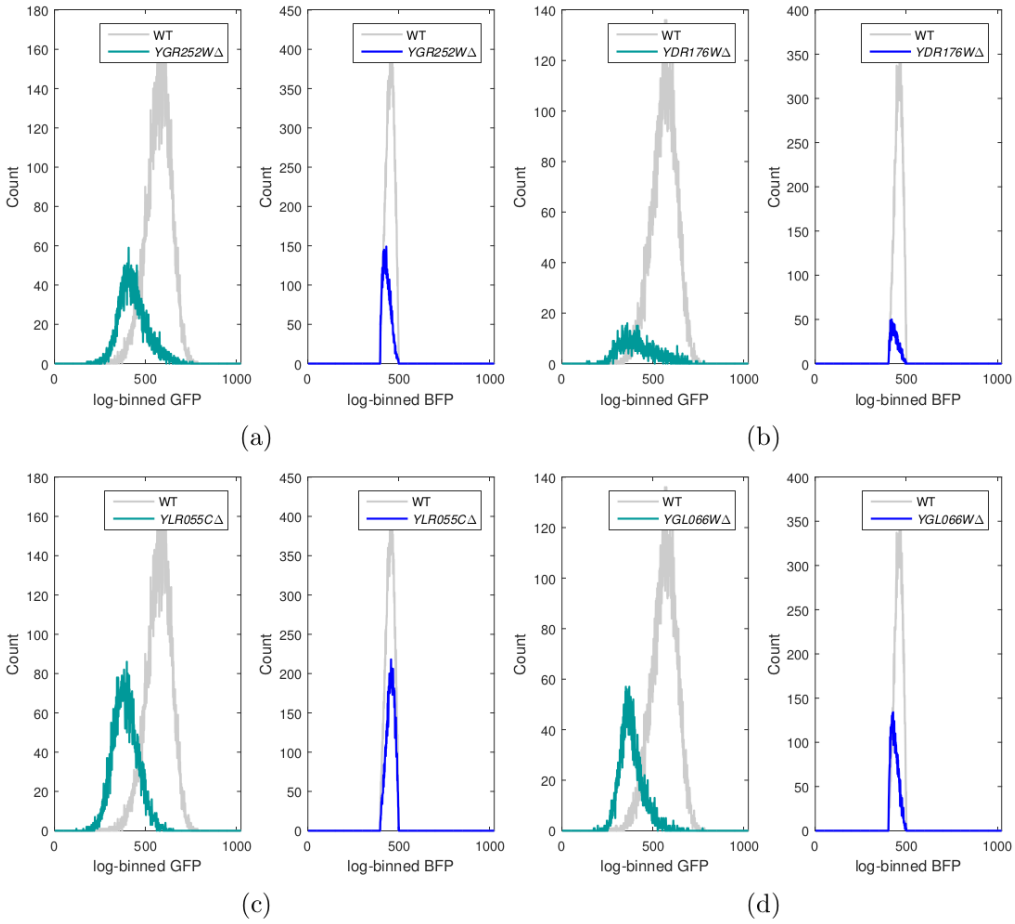



**Figure 6.15:** Fluorescence profiles of mutants of SAGA genes (a) *GCN5*, (b) *NGG1*, (c) *SPT8*, and (d) *SGF73* at 0.015% MMS.

Deletion of *SPT8* appears to differentially suppress the reporter (Fig. 6.15a), suggesting that Spt8p is not necessary for global SAGA function. Spt8p has been shown to impart differential recruitment of SAGA to SAGA-dependent promoters [13]. While *SPT8* is required for the recruitment of SAGA to *ADH1pr*, *PHO84pr*, and *BDF2pr*, it is not necessary for SAGA recruitment to *GAL1pr* [13].

The case of chromatin remodeling and histone modification factors is more complex than what meets the eye. The *ACT1pr* used to drive *yEBFP* as an inter- nal control exhibits few H3 acetylation events [24], suggesting that relatively little remodeling occurs within it. In fact, *ACT1pr* is highly constitutive; its expression may be mediated predominantly by $TF_{II}D$ as is the case with many housekeeping genes [41].

Since acetylation is generally associated with gene expression, it follows that deacetylation by histone deacetylase complexes (HDACs) promotes expression silencing. Disruption of HDACs is thus expected to increase reporter expression. Oddly, this is not the case, with the suppressor outlier subset enriched in members of Rpd3L and Rpd3L-expanded complex at low and intermediate MMS, respectively (Appendix Fig. 3). Granted that again some of these members impart abnormalities to our internal control, there are components that differentially affect the *RNR3pr* reporter, such as *SIN3* (Fig. 6.16).

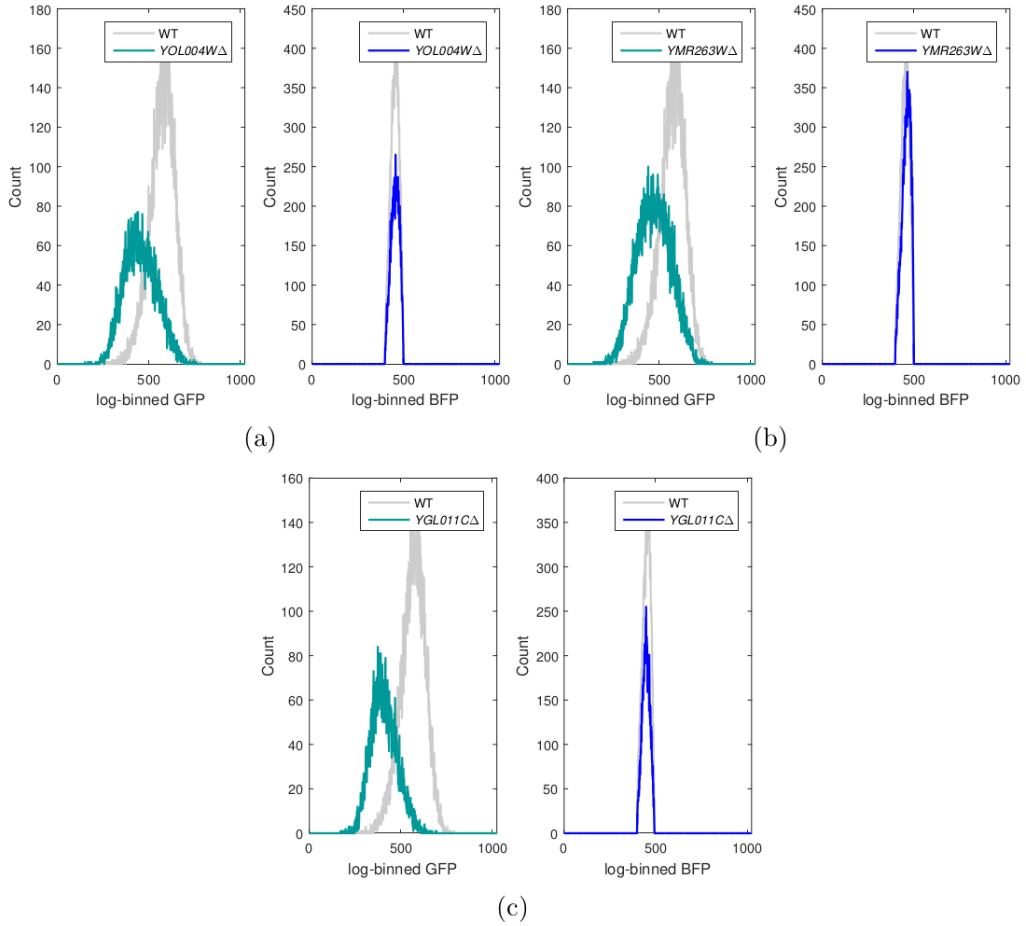



**Figure 6.16:** Fluorescence profiles of mutants of histone deacetylation (a) *SIN3*, (b) *SAP30*, and (c) *SCL1* at 0.015% MMS.

Complicating matters, Sin3p and Rpd3p interact to both repress and activate gene expression via chromatin remodeling [22, 43, 83, 90]. Unfortunately, at the current level of our experimental design, we cannot discern whether *sin*3Δ is truly differentially affecting *RNR3pr* over *ACT1pr* as the latter undergoes little gene silencing. What is more surprising, however, is that translation and ribosomal biogenesis also resulted in differential reporter effects as opposed to the expected non-specificity (Fig. 6.17 & 6.18).

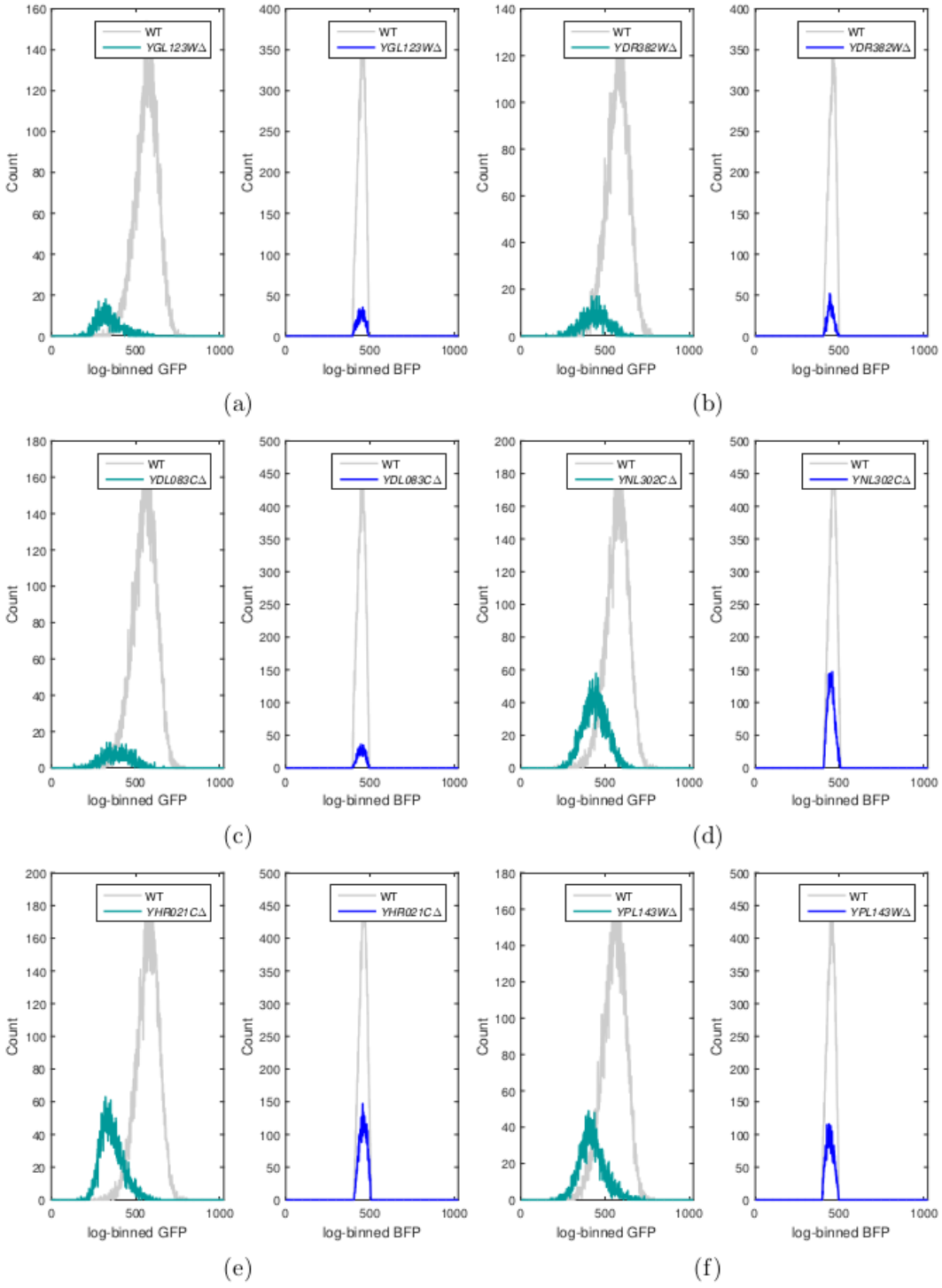



**Figure 6.17:** Fluorescence profiles of mutants of ribosomal biogenesis (a) *RPS2*, (b) *RPP2B*, (c) *RPS16B*, and (d) *RPS19B*, (e) *RPS27B*, and (f) *RPL33A* at 0.015% MMS.

One of the recent developments in proteomics is the recognition of specialized ribosomes (rev. in []). These intricate complexes of rRNA and proteins appear to modulate

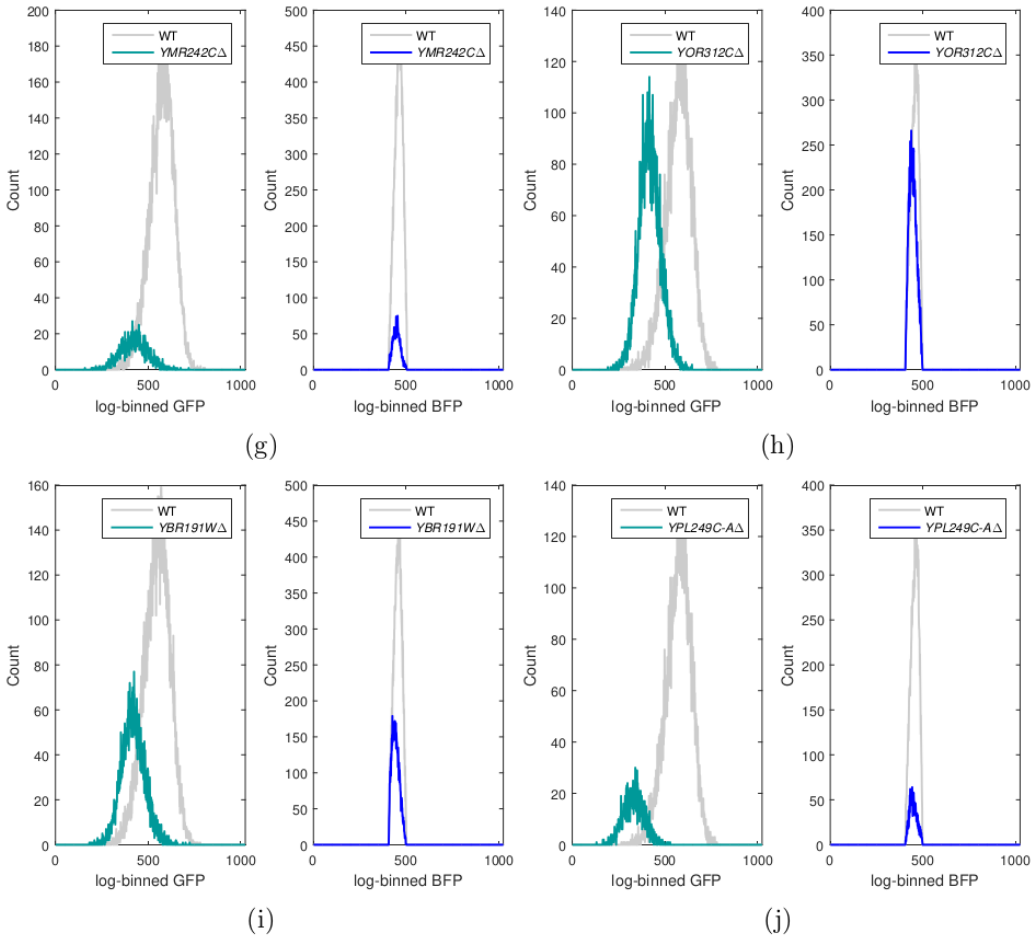



**Figure 6.17:** Fluorescence profiles of mutants of ribosomal biogenesis (g) *RPL20A*, (h) *RPL20B*, (i) *RPL21A* and (j) *RPL36B* 0.015% MMS.

their composition and thereby their behaviours in response to stress signals. Ribosomal specialization establishes the translational machinery as a level of rapid response in the event of stress. While the strategies of recognition and regulation are yet to be discovered, Chan *et al.* implicated tRNA modification. They reported that oxidative stress increases tRNA methylation at the wobble site (the 'loose' site in codons that promotes degeneracy) via the methyltransferase Trm4p [18]. Moreover, methylation of the leucine tRNA leads to translational selectivity of *RPL22A* mRNA because it contains more TTG than its paralog [12, 18]. Our results identify several ribosomal paralog genes (Fig. 6.17) and genes involved in tRNA modification (Fig. 6.18) that appear to differentially affect *RNR3* reporter expression.

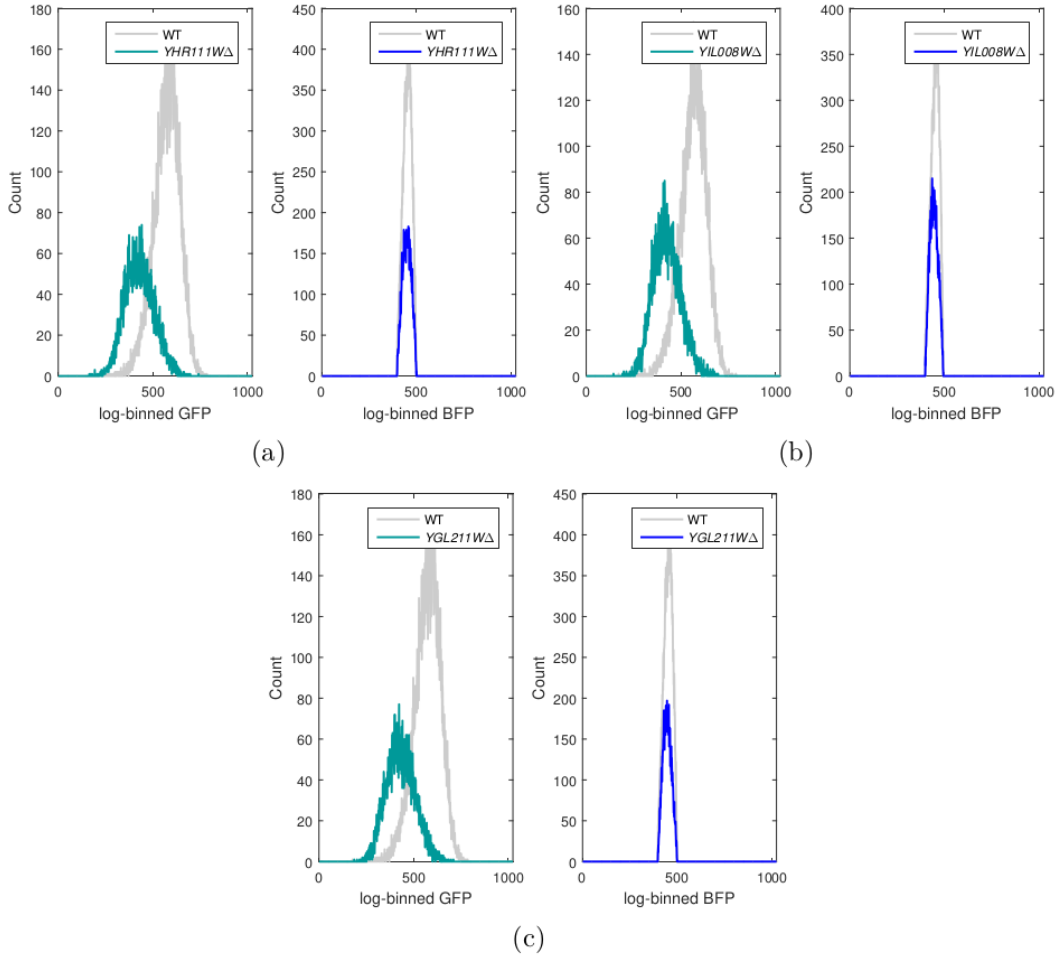



**Figure 6.18:** Fluorescence profiles of mutants of tRNA modification (a) *UBA4*, (b) *URM1*, and (c) *NCS6* at 0.015% MMS.

Notably, with exposure to MMS, deletion of genes required for ROS clearance, such as *YAP1*, *SKN7*, and *GSH1*, emerge as outliers, albeit with contradictory effects. *yap*1Δ and *skn*7Δ enhance reporter expression (Fig. 6.19a—d), suggesting that they may be *RNR3pr* repressors; *gsh*1Δ suppresses reporter expression (Fig. 6.19e).

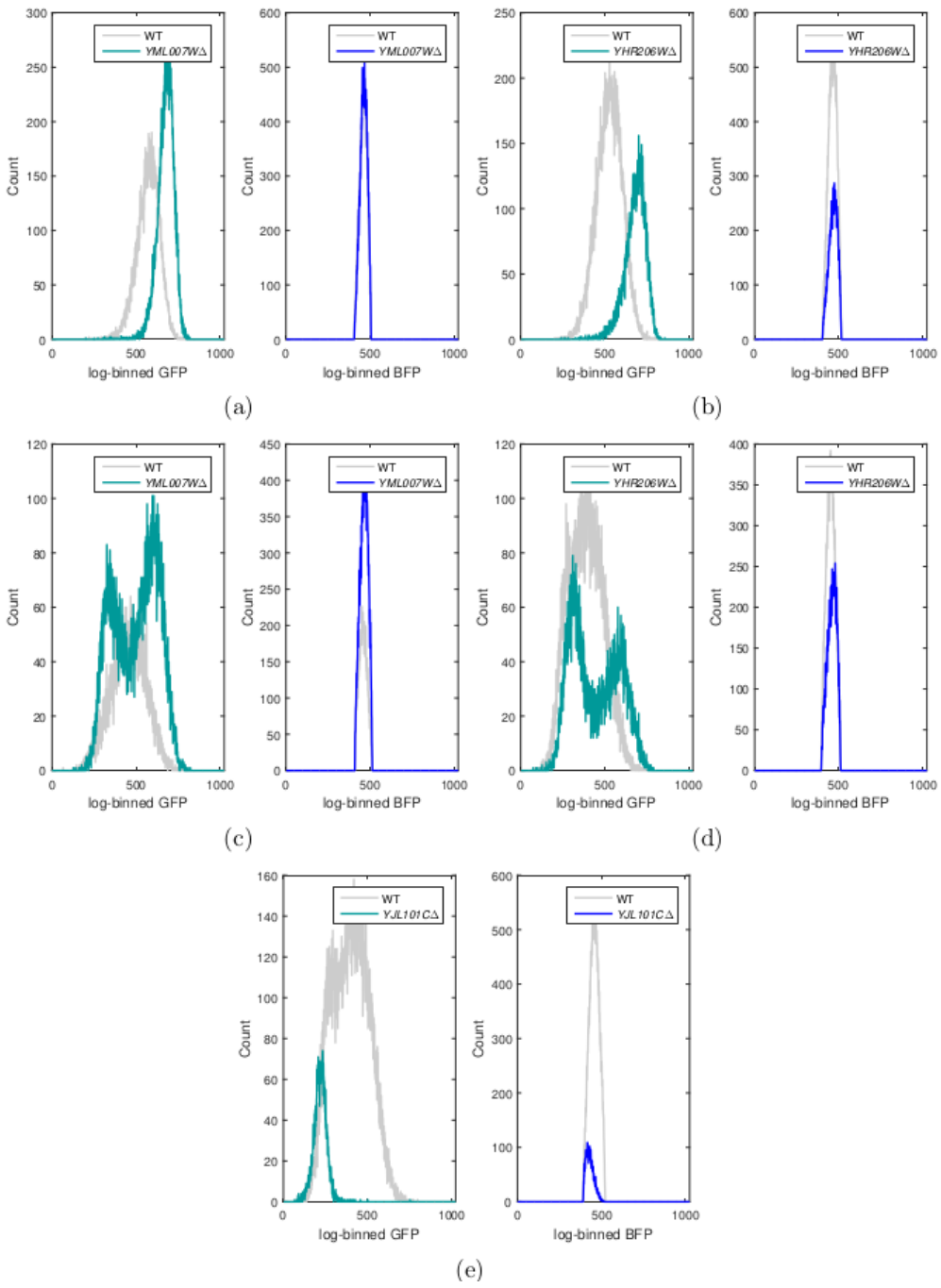



**Figure 6.19:** Fluorescence profiles of ROS clearance mutants at 0.015% and 0.060% MMS. (a) *yap*1Δ and (b) *skn*7Δ at 0.015% MMS; (c) *yap*1Δ, (d) *skn*7Δ, and (e) *gsh*1Δ at 0.060% MMS.

Our results agree with findings that *yap*1 deletion enhances the sensitivity of an *RNR3pr-lacZ* system to MMS [93]. With further implication of the oxidative response through enrichment of stress-responsive ribosomal proteins, and tRNA processing, the

involvement of oxidative stress genes in regulating our reporter is an intriguing theme, especially since only Sod1p so far has been linked to *RNR3* and the DDR [78, 81]. Moreover, although *ROX1* and *MOT3* are involved in oxygen dynamics, they are predominantly involved in response to hypoxia; the link, if existent, between hypoxia and oxidative stress response is still elusive.

Given the ability of $yap1\Delta$ and $skn7\Delta$ to derepress *RNR3* reporter expression, and that both Sod1p and the Fe-S mitochondrial clusters can affect Dun1p independent of Rad53p, we were curious to see if the outliers from our secondary screen integrate their effect on *RNR3* in a checkpoint-dependent or -independent manner.

## 6.2 Chemogenetic screen of $rfx1\Delta$ double mutants reveals MMS-dependent interactions

The aim of this screen is to further refine the effects of the outliers by identifying where they integrate their effect with respect to the checkpoint pathway. Since strains included in the following screen were detected outliers, deletion of *RFX1* (*CRT1*) should indicate whether there effects were mediated by *RFX1* (and therefore potentially the checkpoint) or not. We chose to use *RFX1* since its role in *RNR3pr* repression is clear. Screening a miniarray of double mutants allows us to glean two pieces of information:

1. checkpoint-dependence and -independence: assuming that the checkpoint pathway is as depicted in Fig. 2.1, deletion of *RFX1* creates a break be- tween the replication checkpoint and the reporter. If the deletion of gene $X$ suppresses the $rfx1\Delta$ phenotype, then gene $X$ is involved in either derepression or activation of $RNR3$ transcription. If deletion of $X$, however, enhances the $rfx1\Delta$ phenotype, then we infer that gene $X$ is involved in repression of $RNR3$ transcription;

2. how interactions with *RFX1* change with MMS.

Given the structure of *RNR3pr* inferred from DNAse I sensitivity assay, there are twelve

currently-known repression sites beginning at approximately 85–610 bp upstream of the codon encoding ATG (translation start) [46]. Mapping shows that the most abundant repression sites belong to Mot3p and are located at the beginning and end of the stretch of repressor sites [51, 75]. In the middle lie Rox1p- and Rfx1p-recognized repressor sites. Repression at the *RNR3pr* involves the interaction between the three repressors in recruiting Ssn6p-Tup1p, which in turn promotes chromatin silencing of the region by stabilizing histone deacetylation [45,46,60]. Therefore, *RNR3pr* is highly repressed. Moreover, the presence of many repressor sites allows the expression of *RNR3* to be graded, with the magnitude of repres- sion possibly dependent on both the number and location of these sites with respect to nucleosome-favouring regions. The activation of the checkpoint, therefore, does not guarantee that the RNR3pr is completely derepressed. Klinkenberg's study confirms this, as combinatorial deletion of the repressors yielding the highest $\beta$-galactosidase activity with the triple mutant, followed by the $rfx1\Delta rox1\Delta$, $rfx1\Delta mot3\Delta$, and finally $rox1\Delta mot3\Delta$ (least synergy) double mutants [46].

Viewed differently, enhancement of *RNR3pr-GFP* expression suggests that repression at the *RNR3pr* is relaxed more than the usual in response to DNA damage — i.e. gene deletions conferring phenotypic enhancement are potential repressors of the *RNR3pr*. We wondered, then, whether they can act synergistically or antagonistically with the major *RNR3pr* repressor, *RFX1*. We also wondered whether MMS-induced damage can change the genetic interaction landscape of *RFX1*. We expect that deletion of genes whose products are involved in unravelling nucleosomes are expected to suppress the $rfx1\Delta$ phenotype. In contrast, deletion of the, *ROX1* and *MOT3*, *RNR3pr* repressors should increase *yEGFP* expression [46].

### 6.2.1 $rfx1\Delta$ Phenotype Effectors

Taking the 552 outliers from the secondary screen (including those detected from the data set published by Hendry *et al.*), double mutants were generated using a liquid-phase SGA protocol and $BY4742\ Mat\alpha rfx1\Delta :: natMX$ as query strain (Fig. 6.20).

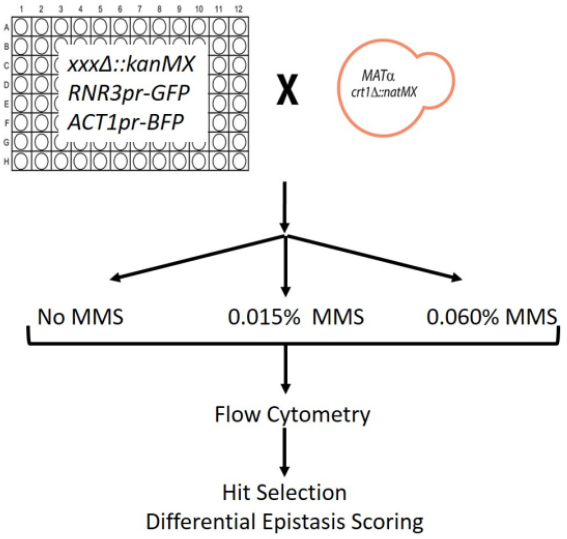



**Figure 6.20:** Schematic of epistasis miniarray protocol.

The assayed and scored double mutants again demonstrate strong reproducibility in identification: $r^2 = 0.8697$, 0.9205, and 0.8334 for strains assayed at 0%, 0.015%, and 0.060% MMS, respectively (Fig. 6.21a–c). WT distributions showed little variability in response MMS (Fig. 6.21d).

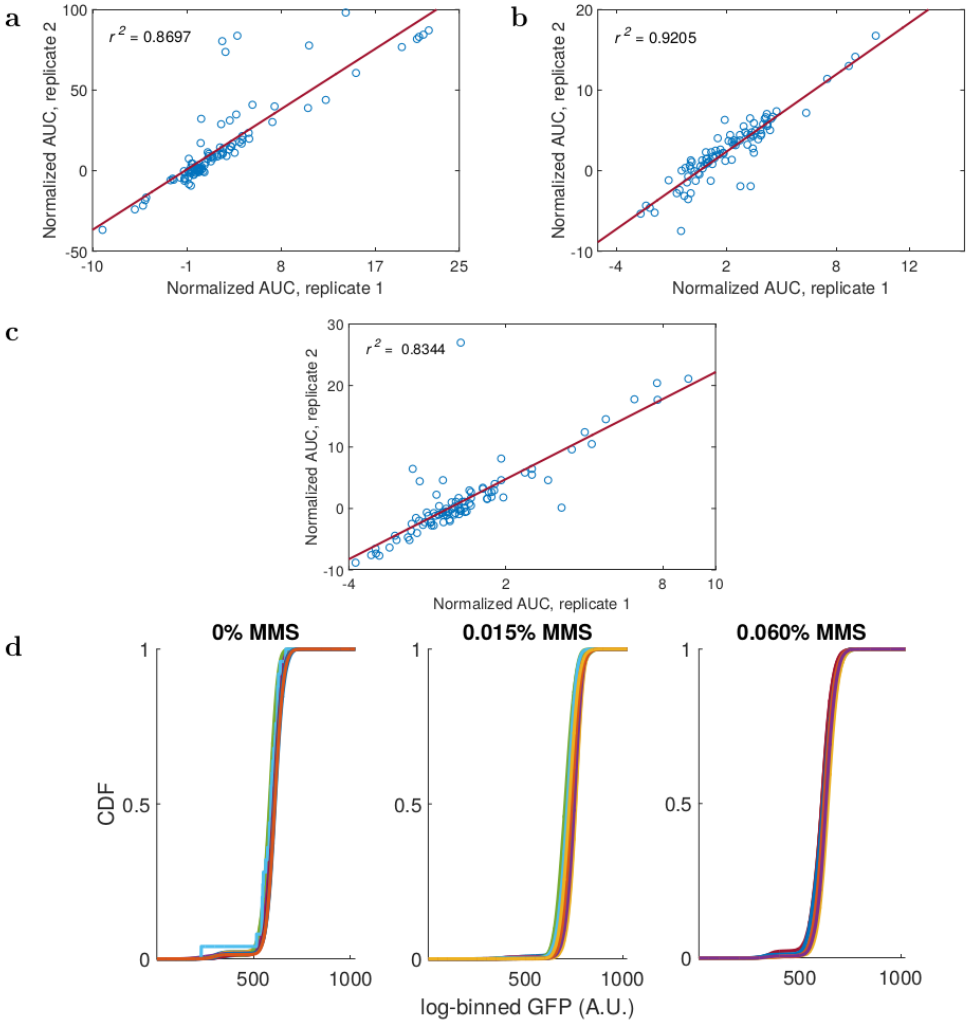



**Figure 6.21:** Reproducibility of epistasis miniarray screen at the three MMS conditions. (a– c) Simple linear regression of normalized AUC scores of independent replicates at (a) 0%, (b) 0.015%, and (c) 0.060% MMS (v/v).(d) Variability of WT CDFs between plates assayed on different days.

The $Z_{AUC}$ scores of the double mutants display $rfx1\Delta$ phenotype enhancement and suppression in the absence of MMS (Fig. 6.22).

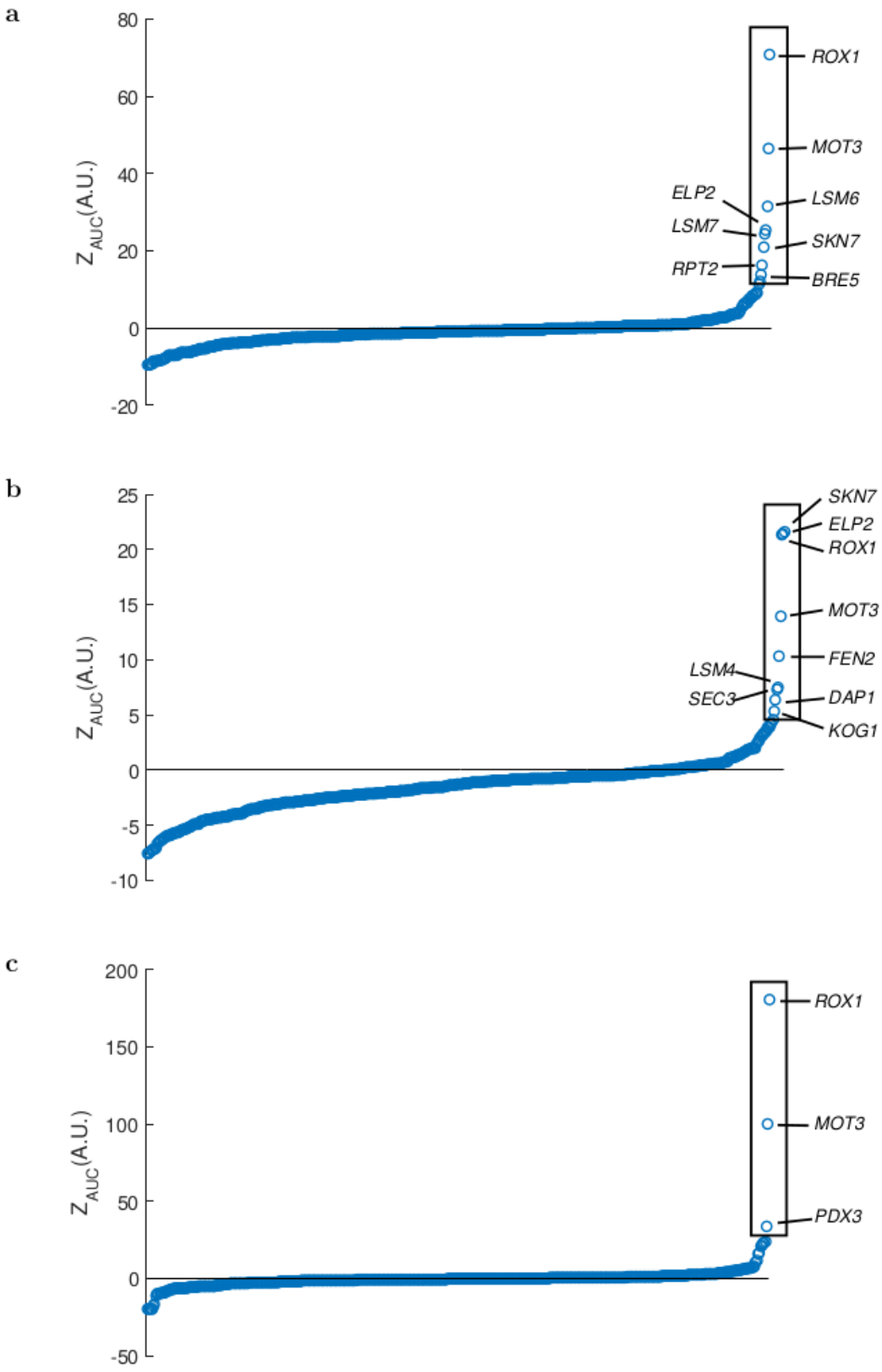



**Figure 6.22:** $Z_{AUC}$-ranked $rfx1\Delta$ double mutant strains at (a) 0%, (b) 0.015%, and (c) 0.060% MMS.

Fluorescence profiles of the top outliers fall into one of three types: enhancement, suppression, and bimodal distributions (See Fig. 6.23 for an example of each).

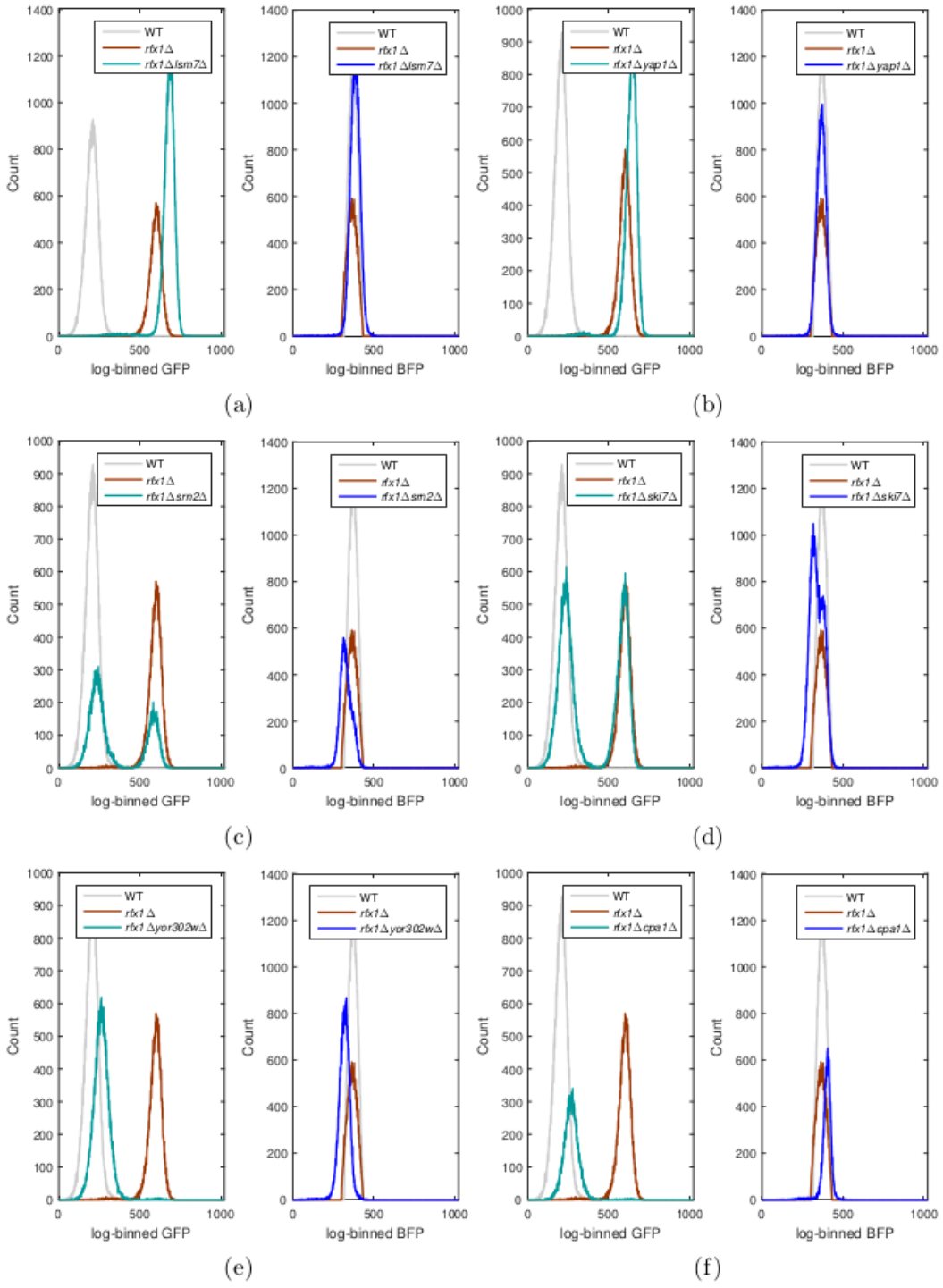



**Figure 6.23:** Double mutants show basal enhancement (a,b), suppression (e,f), or both (c,d) of the $rfx1\Delta$ phenotype.

When untreated, *rox*1Δ*rfx*1Δ and *mot*3Δ*rfx*1Δ top the enhancer lists (Fig. 6.22a & 6.24).

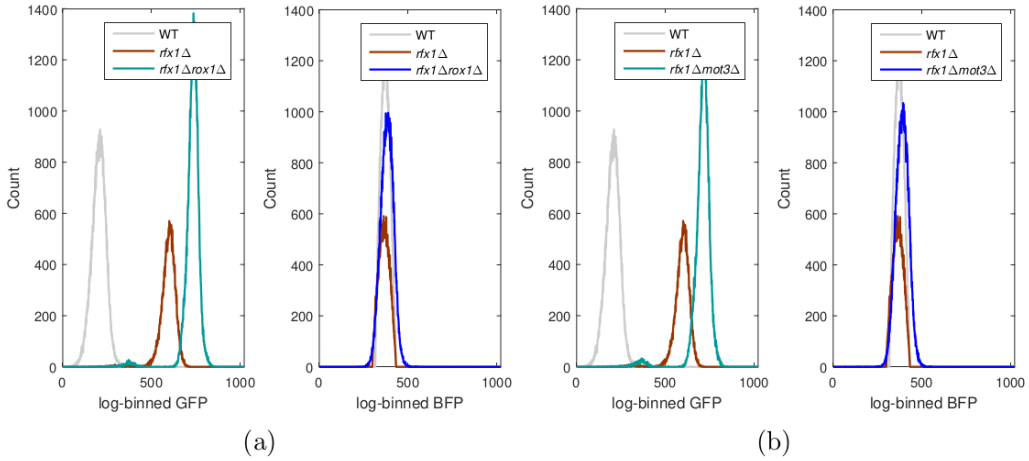



**Figure 6.24:** Fluorescence profiles of known *RFX1* synergists (a) *ROX1* and (b) *MOT3* 0% MMS.

Strains deleted for *SKN7*, and to a lesser extent *YAP1* (Fig. 6.22b), again enhance reporter expression in the absence of *RFX1* with and without MMS (Fig. 6.25). Other enhancers are involved in mRNA decay and export, and tRNA wob- ble modification via the Elongator holoenzyme complex (Appendix 5 Fig. 4 & 5).

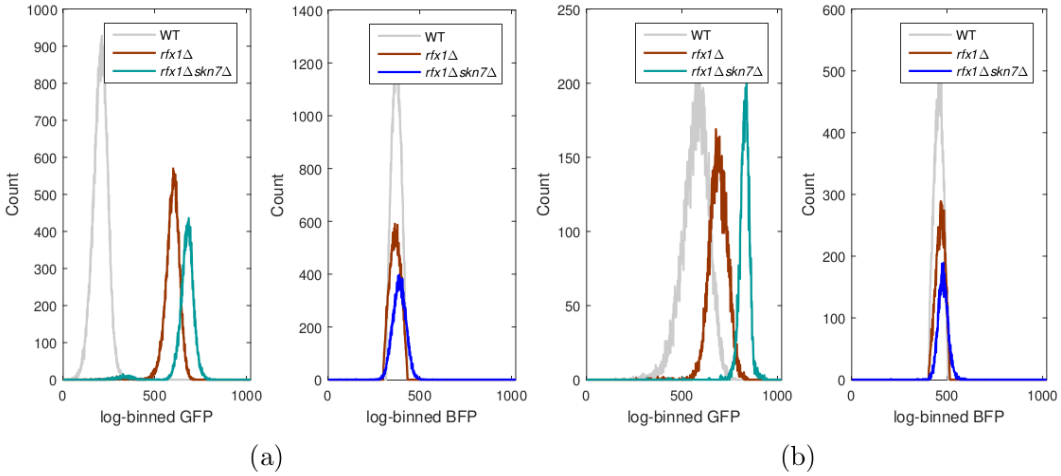



**Figure 6.25:** Dose-dependent *rfx*1Δ-phenotype enhancement by *skn*7Δ at (a) 0% and (b) 0.015% MMS.

On the other end, suppressors of *rfx*1Δ include genes involved in chromatin remodeling (*IES1*, *IES2*, *RTT109*, *SPT8*, *HDA2*, *TAF14*), translation (*ASC1*, *RPS4A*, *RPS20*,

*RPB9*, *TMA23*, *RPL39*), oxidative stress (*HFI1*, *CCS1*), transcription and repair (*CTK3, MMS1, KAP123, RAD52*) (Appendix 5 Fig. 4). *ELP6*, which encodes a member of Elongator holoenzyme, suppresses the $rfx1\Delta$ phenotype unlike the deletion of *ELP2*, a core member (Fig. 6.26).

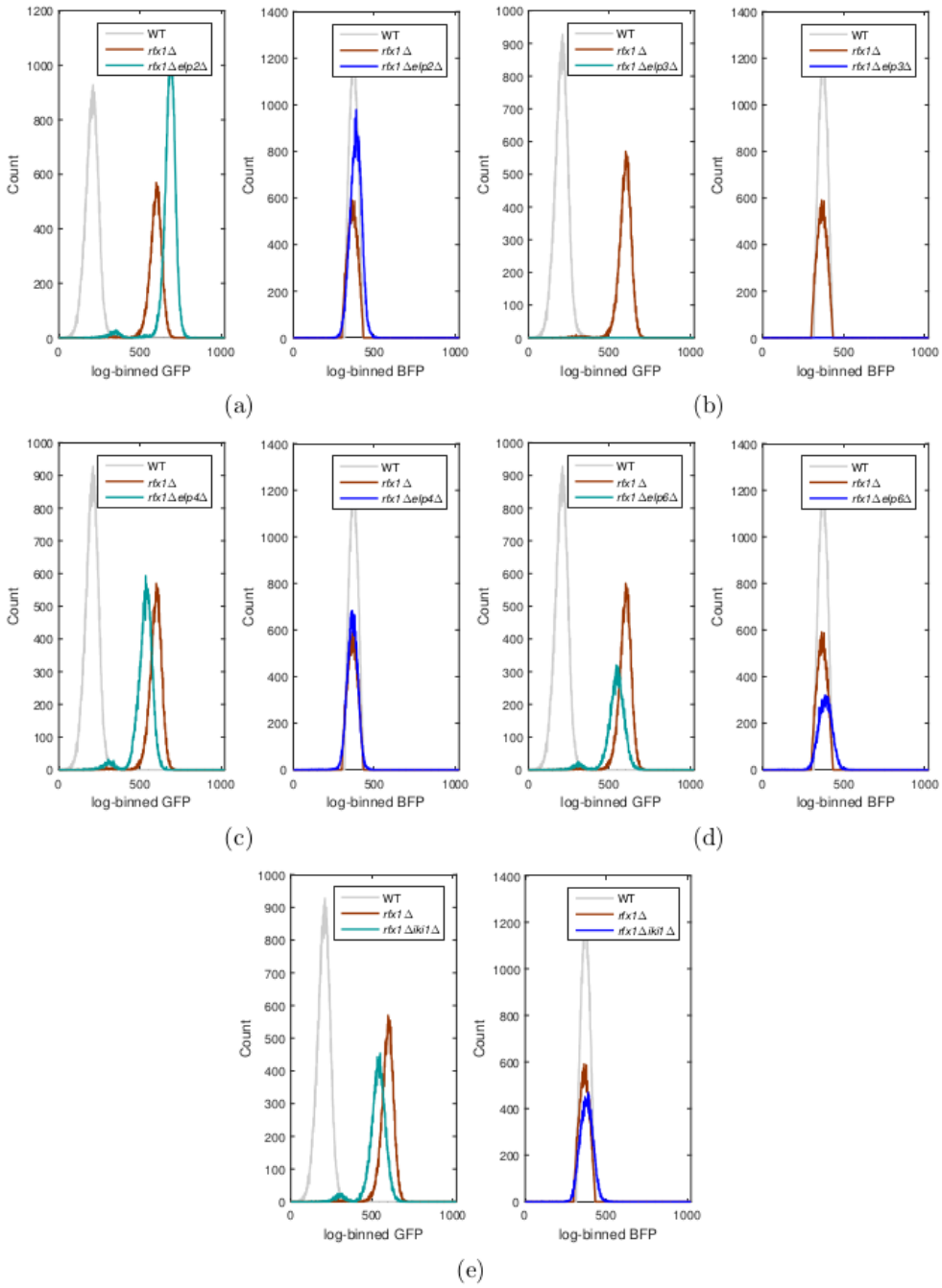



**Figure 6.26:** Fluorescence profiles of mutants of the Elongator holoenzyme complex in the absence of *RFX1* and MMS. (a) $rfx1\Delta elp2\Delta$; (b) $rfx1\Delta elp3\Delta$; (c) $rfx1\Delta elp4\Delta$; (d) $rfx1\Delta elp6\Delta$; (e) $rfx1\Delta iki1\Delta$

Anion-exchange chromatography by Mono Q exposed that Elongator is made up of two heterotrimeric subcomplexes:

1. Iki3p-Elp2p-Elp3p core subcomplex;

2. Iki1p-Elp4p-Elp6p accessory subcomplex [48, 87].

We observe that disruption of the accessory subcomplex led to a minor decrease in reporter expression (Fig. 6.26b–d) consistent with loss of Elongator's transcriptional and acetylation functions [86]. Disruption of the core subcomplex (Fig. 6.26a), on the other hand, is inconsistent with Elongator's function despite it being essential for the complex's assembly [87]. Again, we cannot assess whether the effect of *elp2*Δ is global because of limitations with the *ACT1pr*-driven internal control. While phenotype suppression was not significant, the absence of *RFX1* rescues lethality of *iki1*Δ; whether this is due to increased dNTP levels or derepression of target genes involved in translation is unknown (Fig. 6.26e).

Addition of MMS results in dose-dependent $rfx1\Delta$ phenotypes (Fig. 6.27). Outliers in these sets were not separated based on their yEBFP profiles, thereby representing unique and common mutants that confer abnormality in reporter and expression capacity.

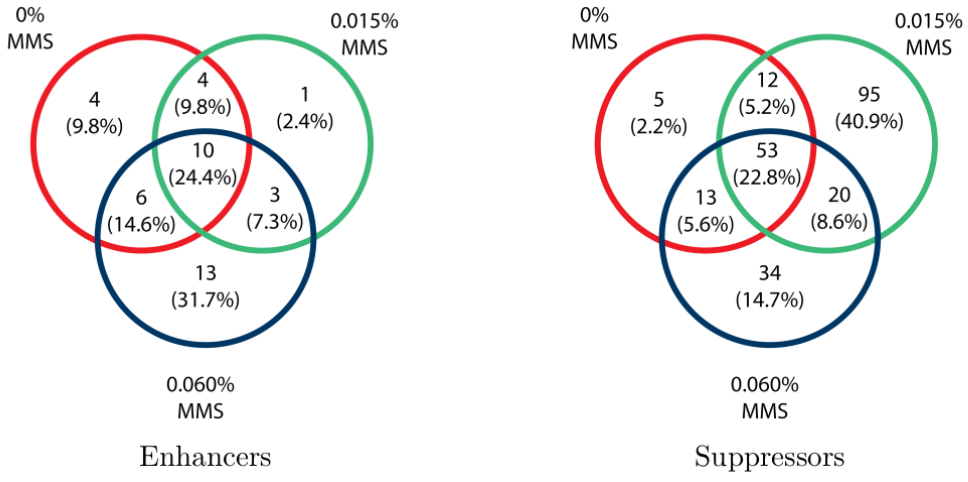



**Figure 6.27:** Overlap of outlier sets across the different MMS conditions. Percentages refer to the proportion of subset in the size of the outlier list across all three MMS conditions.

Functional and complex enrichment of $rfx1\Delta$ outliers at 0% ($n = 107$), 0.015% ($n = 198$), and 0.060% MMS ($n = 182$) show that the most enriched terms in suppressors at the three drug conditions relate to chromatin remodeling, transcription, or translation (Appendix 5 Fig. 4a,c,e & 5a,c,e). Significant enrichment ($P < 10^{-7}$) in DNA repair (GO:0006281) and cellular response to DNA damage stimulus (GO:0006974) is also also observed in the set of suppressors at 0.015% MMS (Appendix 5 Fig. 4b).

Enhancers of the $rfx1\Delta$,on the other hand, phenotype are enriched in members of complexes involved in degradation:

1. protein degradation: Ubp3-Bre5 deubiquitination complex; proteasome regulatory particle, base subcomplex; proteasome regulatory particle

2. mRNA degradation: U6 snRNP; Lsm1-7-Pat1; Elongator holoenzyme; and cytoplasmic mRNA processing body.

### 6.2.2 *RFX1* Genetic Interactions

To assess whether the genes in the outlier set are interacting with *RFX1*, we evaluated genetic interactions by accounting for fold-change in expression between the single knock-outs, double knock-outs, and neutral mutants. With the deletion of *RFX1* in the single-knock-outs, we have crudely divided their effect into two groups:

1. Outliers that funnel their effects on reporter expression via *RFX1*, and, potentially, the checkpoint;

2. Outliers that do not affect reporter expression via *RFX1* and the checkpoint pathway, e.g. checkpoint independence or effect downstream of transcription (Fig. 6.28).

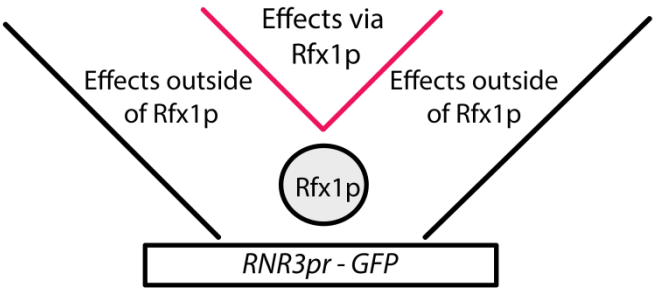



**Figure 6.28:** Deletion of *RFX1* should refine the effects of the outliers to those that act outside of the checkpoints and those who potentially do. Assuming *RNR3* is regulated as depicted in Fig. 2.1, outliers that lose their effect in the absence of *RFX1* — i.e. their effect is different from secondary screen's — are potentially acting through the checkpoint. Their point of influence along the DDR would not be clear. Those that retain their effect may be independent effectors of *RNR3pr*.

$\epsilon$-scores of the 552 strains in the miniarray were evaluated relative to their profiles in the secondary screen, exposing modest *RFX1* interactions ($10^{-\epsilon}$-fold) when exposed to no and intermediate MMS, and a cluster of strong interactors ($n = 44$) in low MMS. $\epsilon$ scores of double mutants assayed at 0% and 0.060% MMS were very similar and small (Fig. 6.29a & b). In contrast, $\epsilon$ scores at 0.015% vs. 0% MMS groups display two clusters (Fig. 6.29c).

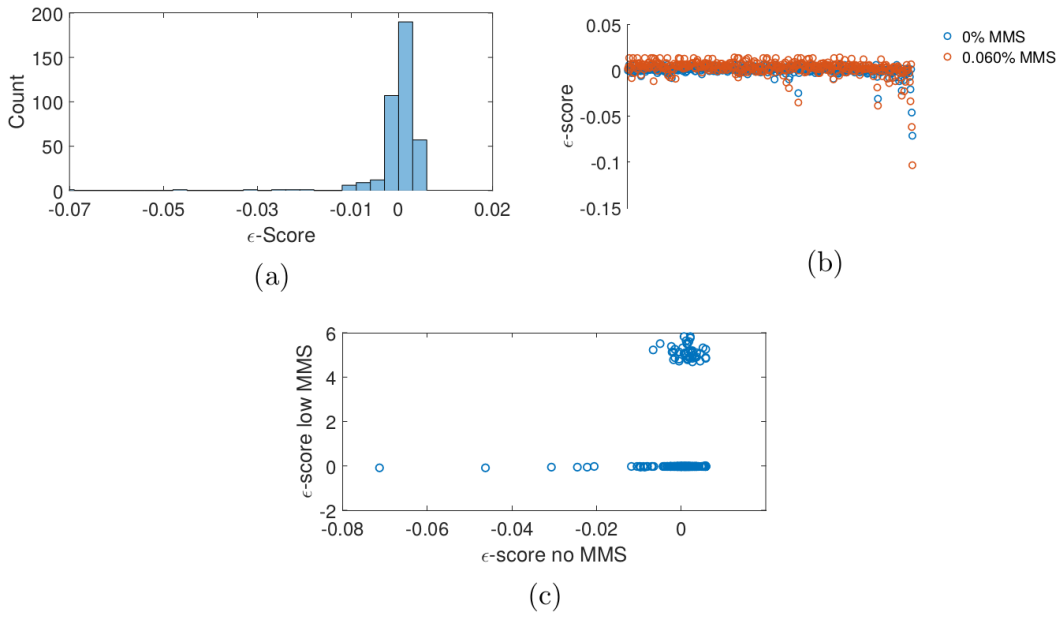



**Figure 6.29:** Distribution of multiplicative $\epsilon$ scores across the different MMS conditions. (a) Distribution of $\epsilon$ score at 0% MMS. (b) Extensive overlap of $\epsilon$ scores at 0% and 0.060% MMS. (c) $\epsilon$ scores at 0.015% MMS plotted against their $\epsilon$ scores when untreated shows a cluster of synergists.

Due to overlapping scores at 0% and 0.060% MMS, differential genetic inter- actions were assessed between 0.015% MMS and 0% groups (Fig. 6.30). The resultant genetic network displays that the strongest interactors are antagonists and are involved in ribosomal biogenesis (*RPS16B, RPL33A, RPB9, RPL43A, RPL13A*), chromatin remodeling (*IES1, ESA1, BRE1, CHD1*), negative regulation of chromatin silencing (*RXT2, SIF2*), translation (*TPD3, DHH1, SLH1*), DNA damage repair (*RAD52, RTT109, KAP123*).

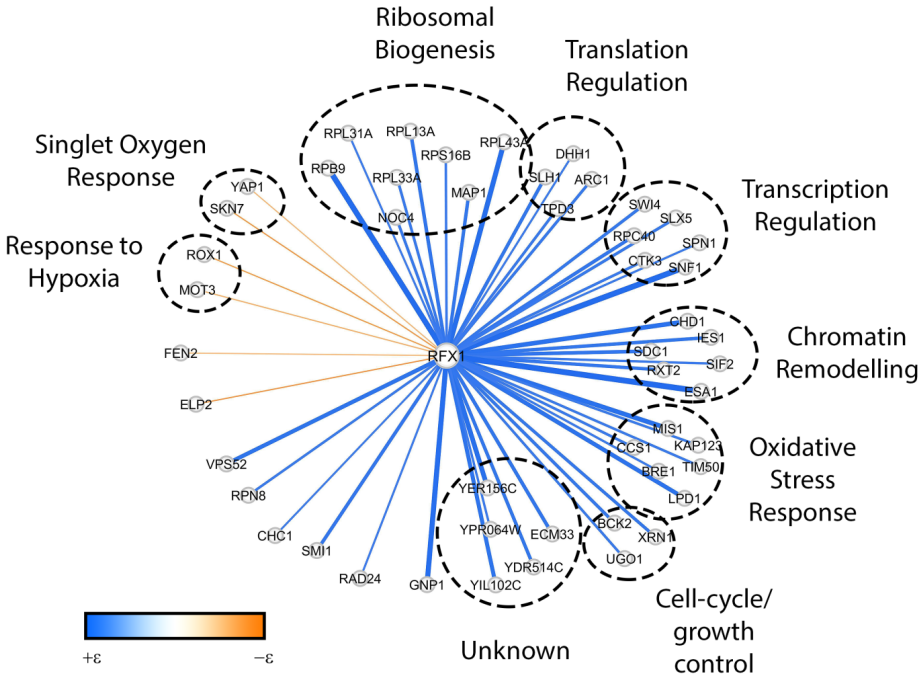



**Figure 6.30:** Genetic interaction network of *RFX1* displaying the differential epistatic interactions at 0.015% MMS. Edge thickness positively correlates with $|\epsilon|$ score; orange edges denote differential synergy and blue edges denote differential antagonism.

Synergists conferred modest differential interactions with *RFX1*, with the strongest interactors *ELP2*, *SKN7*, and *ROX1* conferring 2.6–3-fold increases in reporter expression. The set of synergists includes previously identified enhancers in ad- dition to *FEN2*, an $H^+$-pantothenate (vitamin $B_5$) symporter [76]. Pantothenate is a precursor of cofactor A (CoA) involved in a range of metabolic processes, as well as tRNA wobble position and histone acetylation. Although no study was done on the mechanism and effect of pantothenate deficiency in *S. cerevisiae*, we hypothesize that if pantothenate is the major precursor of CoA and the symporter is the major source for pantothenate availability in the cell, deficiency in CoA will mimic the decrease in acetylation events observed in *Drosophila* and mammalian cells [69].

## 6.3 Bimodal Distributions

While generally the wild-type response to MMS in *BY4741* is bimodal (i.e. two distinct fluorescence subpopulations can be defined), the separation of the subpopulations is more

pronounced in the $rfx1\Delta$ double mutant miniarray. A total of twenty-seven strains that appear to have a bimodal distributions were streaked out on four selection media — L-canavanine, URA drop-out, G418, and *ClonNAT* — for two reasons:

1. To infer whether selection steps have failed during double mutant generation;
2. To isolate a patch of isogenic cells to be assayed.

All three patches of twenty-seven strains passed the selection tests (see A.4) and were assayed at the three MMS conditions. What emerged was a mixture of consistent and inconsistent profiles among the biological replicates and across the MMS conditions (Table 7). Within a drug condition, inconsistencies encompass various scenarios:

1. 2/3 replicates have bimodal profiles, and the remaining replicate's is unimodal;
2. 2/3 replicates have bimodal profiles, and the remaining is bimodal with a shifted peak;
3. 2/3 replicates have unimodal profiles, and the remaining replicate's is bimodal;
4. 2/3 replicates have unimodal profiles, and the remaining replicate's is unimodal with a shifted peak.

Between drug conditions,

- Strains that were consistent become inconsistent;
- Strains that were unimodal become bimodal and *v.v.*

# 7 | Discussion

High-throughput technologies pave the way to studying cellular responses at levels that encompass cell-wide dynamics. In addition to piecing together different phenomes from low-throughput studies, which can be time-consuming and narrow in scope, one gains a 'bird's-eye view' of how processes and pathways shift in response to stress. Our goal was to apply two of these techniques to explore the complexity of the replication and DNA damage checkpoints in budding yeast in three ways:

1. Identifying biological processes unexpected to play a role in checkpoint activation;
2. Refining the effector list to identify genes that may affect reporter expression independently of the checkpoint;
3. Dose-dependence of dominant biological processes and interactions.

The primary and secondary chemogenetic screens narrowed the list of potential effectors from 5,850 DMA and DAmP strains to 552 outliers that may be acting through the checkpoints. Challenges with the primary screen and data analysis facilitated refinement of our protocol to control for population density. We implemented these improvements by including OD 600 checks prior to screening and screening strains in batches of 96 at a time. Using flow cytometry, de-condensation into 96-well plates did not elongate the duration of screening.

While we cannot report the false positive rate of the primary screen, of the 141 WTs assayed, only 3 were identified as potential outliers after RMSE-scoring. Similarly, we cannot infer the false positive rate from the secondary screen's results, because of the inclusion of potential false negative strains. Of the 819 re-screened strains, we identify

552 as hits with high confidence.

Reporter expression is expected to fall under four archetypes:

1. Phenotype suppression effected by deletion of checkpoint-independent *RNR3* transcriptional activators, checkpoint activators, and proteins involved in derepression, chromatin remodeling, transcription or translation;

2. No change which would include genes that do not regulate checkpoint activation, and those whose effect on the checkpoint is masked;

3. Phenotype enhancement which would include genes involved in repression of *RNR3* transcription (regardless of dependence on checkpoint activation), or nucleosome deposition.

4. Bimodal distribution, from which we infer that the gene deletion stabilizes arrest of one of the subpopulations at a stage where *RNR3* is not expressed after mitosis — i.e. G1. Gene deletions that are expected to do so may include genes involved in cell-cycle restart after arrest, or those that effect dysregulation in cell-cycle rate.

While several biological processes emerge at the different MMS doses in the primary, secondary, or EMAP, a few do so across the board. Oxidative stress response, ribosomal biogenesis, and transcriptional regulation appear to affect *RNR3* expression without affecting the expression of housekeeping genes (inferred from BFP). Chromatin remodeling also affects *RNR3* expression; whether the effect extends to *yEBFP*, we cannot discern given the constraints of using *ACT1pr*. We suspect that the involvement of these four processes — especially members in each group involved in oxidative and hypoxic stress responses such as *UBA4*, *NOC4*, and *RXT2* — implicate oxidative damage in the regulation of *RNR3*. Recall that the primary and secondary screen results detect *YAP1* and *SKN7* as potential repressors, and *GSH1* as a potential activator of *RNR3*. When evaluated in the epistasis miniarray, *YAP1* and *SKN7* persisted among synergists, albeit weakly, suggesting that they might be affecting *RNR3* expression independent of *RFX1*

and the checkpoint(s).

The oxidative stress response is complex. Oxidative stress in *S. cerevisiase* takes the form of an imbalance in the generation of reactive oxygen species (ROS) by oxidative phosphorylation and their clearance by antioxidants/reductases, such as superoxide dismutases and thioredoxin peroxidase. Because the accumulation of ROS in the cell leads to oxidative damage of macromolecules, including genomic and mitochondrial DNA, oxidative stress responders may influence DNA damage repair signalling. An increase in ROS activates several pathways that lead to the transcription of genes involved in ROS clearance, including superoxide dismutase 1 (*SOD1*) [50].

Orp1p senses the increase of $H_2O_2$ by reducing the ROS [55]. Upon its own oxidation, Orp1p interacts Yap1p-Ybp1p to relay the message by getting reduced (and oxidizing Yap1p) (rev. in [21]). The now-oxidized Yap1p dissociates from its regulatory unit, Ybp1p, and accumulates in the nucleus [8, 71]. Yap1p, then, either mediates transcription of antioxidant genes [32]. Yap1p also appears to interact with chromatin remodeling complexes; whether this is solely to augment transcriptional activation is unknown.

Detection of superoxide radicals also activates Sod1p, which is involved in the (1) conversion of the superoxides into $H_2O_2$, and (2) transcription of genes involved in antioxidation, DNA repair, general stress resistance, Cu/Fe homeostasis, and DNA replication-stress responses. Tsang et al. demonstrated the ability of Sod1p to bind the RNR3pr, as well as the increase in Sod1p-RNR3pr binding with administration of $H_2O_2$ [81]. Their study placed Sod1p downstream of Dun1p, with the latter as a regulator of Sod1p's transcriptional activity. The net result of both Sod1p- and Yap1p(-Skn7p)-mediated transcriptional activation is an increase in the number of entities that can reduce ROS to $H_2O$.

How might ROS play a role in our system?

Endogenous and exogenous DNA damage can yield an increase in certain types of ROS, such as $H_2O_2$. Deletion of central components of BER, Ntg1p, Ntg2p, and Apn1p, and NER, Rad1p, expels higher levels of hyrodxyl radicals and H 2 O 2 in the absence and presence of MMS [71]. Deletion of genes that are not associated with repair pathways, such as *TSA1* and *DUN1*, also increase basal levels of ROS in the absence of drug-induced damage [79]. Alkylation damage, therefore, might not be the only source of damage the cells are experiencing in our screen, with oxidative damage affecting *RNR3* as a result of null mutations, MMS, or both.

While *sod*1Δ is in our library, PCR-mediated confirmation was inconclusive. Nevertheless, *yap*1Δ and *skn*7Δ mutants appear to derepress RNR3 expression at low MMS, and a bimodal distribution of the response at intermediate MMS. Their profiles suggest that either Yap1p and/or Skn7p are direct repressors or they mediate repression of *RNR3* via repressors of *RNR3* transcription. Both Yap1p and Skn7p are considered transcriptional activators that either jointly or independently promote expression of oxidative stress response genes. There is no evidence supporting a repressor function for either. While there is a putative Yap1p binding site upstream of the *RNR3* start signal [32, 93], there is no reported Skn7p binding site. If direct repression is at work, then Skn7p might be augmenting its effect via Yap1p and Ssn6p-Tup1p — i.e. Yap1p is capable of transcriptional repression.

Indirect derepression is also a very likely explanation of the increased *RNR3* expression. The deletion of *YAP1*, in particular, means that clearance of $H_2O_2$ is hampered by (1) lack of re-generation of Orp1p, and (2) lack of transcription of oxidative stress response genes. While its activation mechanism is still unknown, Skn7p nonetheless also activates the transcription of genes independently (e.g. *OLA1* and *DNM1*) and in conjunction with Yap1p (e.g. *SOD1*, *SOD2*, *TRX2*, and *TSA1*) [50]. Thus, deletion of *SKN7* is also expected to compromise the oxida- tive stress response. Consequently, the levels of intracellular ROS are expected to rise, leading to greater oxidative damage. How this

damage is implemented in the RNR3 response is unknown and may involve overlapping pathways. For example, high $H_2O_2$ -mediated DNA damage may activate the checkpoint pathway, with Mec1p transducing the signal all the way to increased *RNR3* expression. Alternatively, Yap1p/Skn7p may be exerting their effect via Rox1p/Mot3p. Yap1p is known to physically interact with Sro9p [74], a co-chaperone of Hap1p when haem is absent [37]. When Hap1p binds haem, it mediates the expression of two *RNR3* repressors, Mot3p and Rox1p [37]. Under oxidative stress, haem is not synthesized, and Hap1p does not activate the transcription of its genes. If Yap1p is required for Sro9p in the cell under stress conditions in order to mediate the activation of Hap1p by haem, it is expected that the deletion of *YAP1* would then not allow Hap1p to bind haem and become a transcriptional activator complex — i.e. *ROX1* and *MOT3* would not be synthesized, relieving some of the repression at the *RNR3pr*.

Epistasis analysis shows weak synergy between *YAP1* and *RFX1*, and *SKN7* and *RFX1*, suggesting that the effects of single mutants are primarily outside of Rfx1p and the checkpoint. Moreover, while *ROX1* and *MOT3* are involved in response to hypoxia, they may also represent points of checkpoint-independence. Their epistasis scores suggest this, and are in line with Klinkenberg *et al.*'s dis- covery that hypoxic conditions can increase *RNR3* expression in the absence of 4-nitroquinoline 1-oxide (4NQO), a radiomimetic drug [46].

## Bimodal profiles: self-preservation or cell-fate indecision?

Whether bimodality is a result of genetic-heterogeneity within colonies is unknown. Strains were originally streaked on YPD to isolate colonies that were patched again on YPD. These patches were then non-serially replica plated onto the four selection plates. While there is no systematic study on what makes colonies distinct entities, it is reasonable to assume that (1) the position of the mother cell's deposition, and (2) the fitness of the cells in nutrient uptake and metabolism are involved. Two genetically-different cells in close proximity then may proliferate within the same region and merge into one colony,

and the genetic heterogeneity is propagated.

In the chance that the cells are indeed isogenic, the inconsistency raises the question whether these specific deletions are dysregulating cell-cycle progression rate such that there are two distinct subpopulations that are out-of-sync and either cycling or arrested. If asynchrony is the reason behind this phenomenon, then we expect that there is an enrichment of slow-growth phenotype among the double mutants, as well as that combinatorial mutations preferentially elongate a few cell cycle stages. The two expectations can be assessed using BioscreenC, an OD plate reader that enables multi-day data collection in a temperature-controlled chamber, and centrifugal elutriation, which separates cells by size for volumetric measurements, respectively.

Regardless of the underlying causes of bimodal expression phenotypes, the inconsistency we observe here would be easily lost at a population-level measurements and analyses, increasing the rates of false positive and false negative errors depending on the profile. While characterization of bimodal distributions does add a level of complexity in deciding which of the peaks is the meaningful one (if not both) and interpreting the screen score, if the source of variability is understood, these challenges resolve themselves.

## Future Directions

Given the reproducibility of the screens as well as agreement with pertinent literature, we are confident in the quality of the results for follow-up studies.

Our results recapitulate previously published linkages between DNA damage repair, transcriptional regulation, and chromatin remodeling factors to checkpoint activation. What emerges as surprising is that there are chromatin remodeling components and ribosomal genes that might be affecting *RNR3pr* differentially. To ascertain their effect, chromatin remodeling and histone modification genes that conferred abnormality when deleted need to be regenerated with an internal control driven by a less constitutive pro-

moter. Ribosomal genes' effect on *RNR3*, on the other hand, would require assessment on the mRNA level in the presence and absence of MMS, cycloheximide, and ROS. Exploring the expression profiles in WT and mutants at these conditions could expose the changes to ribosomal composition in response to DNA damage. The mechanism through which different ribosomal variants are affecting expression of *RNR3* can then be investigated using Ribo-Seq.

Collectively, the biological processes enriched in all screens indicate that oxidative stress is a strong effector of *RNR3* transcription with both checkpoint-dependent and independent mechanisms. To expand on this, hierarchical epistasis analysis with pairwise mutants can be done to establish the relative order of the oxidative stress response genes in a hypothetical pathway. One of the main challenges of such a study lies in defining a meaningful measure of epistasis, given that bimodal profiles complicate interpretation. Foremost, however, is identifying the source of bimodality; if indeed the lower subpopulation is that of non-expressing cells, a probabilistic epistasis score should be defined and used.

Moreover, the translatability of the results found in our screens to RNR2 and *RNR4* would provide a link between checkpoint activation and dNTP biosynthesis. RNR is tightly regulated because it can easily unbalance the dNTP pool, increasing mutation rates. The enzyme is regulated in four ways:

1. transcription of *RNR* genes;
2. allosteric inhibition of large subunit by dATP (negative feedback);
3. competitive inhibition of Rnr1p by Sml1p;
4. sequestration of the large subunit and small subunits into the cytoplasm and nucleus, respectively.

Budding yeast naturally relaxes its negative feedback in response to DNA damage. Chabes *et al.* found that this relaxed feedback inhibition enhances damage resistance

when cells were exposed to 4NQO [17]. It also quadrupled the mutation rate at the $CAN^r$ ,with $A \bullet T \rightarrow G \bullet C$ substitutions doubling, and $G \bullet C \rightarrow A \bullet T$ substitutions increasing by $\approx$ 6.6-fold. Even without 4NQO-mediated damage, disruption of the allosteric site triples the mutation rate. Therefore, gene deletions that disrupt this balance will affect the rate of mutations. It would be interesting to see if the hits from the screens affect *RNR2* and *RNR4* product abundance (transcripts or proteins), and whether these changes induce gross morphological abnormalities. Translatability to *RNR2* and *4* would implicate dNTP pool regulation and its subsequent effects on mutation rates. Finally, while more challenging, the differential effects that some ribosomal and proteasomal degradation genes have on *RNR3* expression present another avenue for exploration. Many ribosomal genes appear to reduce *RNR3* expression and exert their effect upstream of *RFX1*, suggesting that they are involved in the expression of *RNR3* activators. Conversely, proteasomal degradation genes, while enriched in the secondary screen outlier set as phenotype suppressors, appear in the EMAP as enhancers of $rfx1\Delta$; the majority of proteasomal degradation genes do not affect $rfx1\Delta$ phenotype — i.e. they may be acting downstream/independent of the checkpoint pathway.

# Appendix 1: RMSE reproducibility

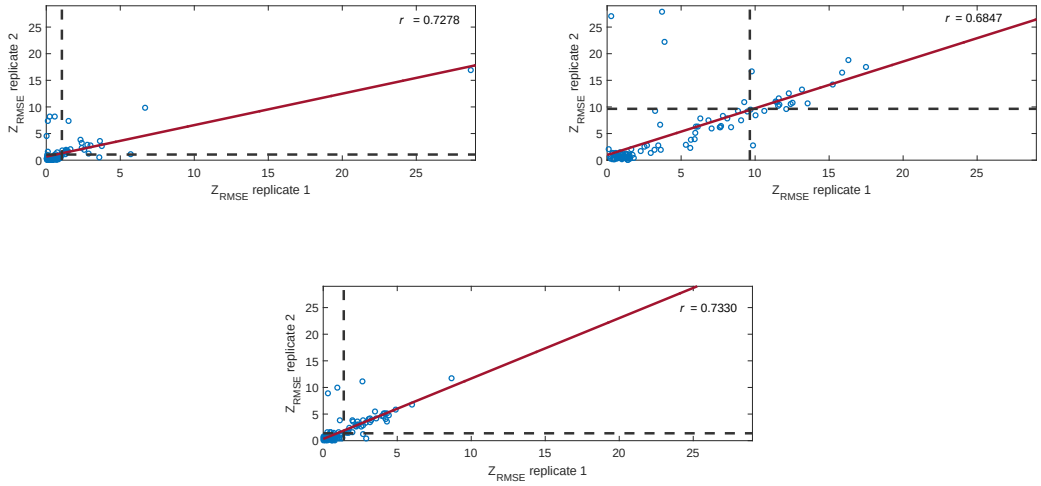



**Figure 1:** RMSE-score variability between 96 replicates at (a) 0%, (b) 0.015%, and (c) 0.060% MMS (v/v). Dashed grey lines denote $Z_{RMSE}$ at which $P \approx 0.05$.

# Appendix 2: Table of Primers

**Table 4:** List of primers used for PCR validation.

| ORF | | Sequence |
|---|---|---|
| *RFX1* | FP | CCTGAGCATGGACTCACAGA |
| *RFX1* | RP | AAGCGGATGTTCTGTGTTCG |
| *MOT3* | FP | AACAATATGTCTGCGTCGCC |
| *MOT3* | RP | TGGTTAGGTGCTGTCGTAGG |
| *ROX1* | FP | CAGACAGCACTACCACAGGA |
| *ROX1* | RP | GGTAGGTCTCTTGCCGTCTT |
| *IBA57* | FP | CTGAAGAAGGAACGACGAAGG |
| *IBA57* | RP | GAGAGATCATCCACTTGCTC |
| *TSA1* | FP | GAAGAACAAGGCGCTCAAGT |
| *TSA1* | RP | TTGTTGGCAGCTTCGAAGTA |
| *YAP1* | FP | CGAGCAACCGAAGAAGAAGG |
| *YAP1* | RP | GGAATCCAAAGTAGCGCTGG |
| *SKN7* | FP | CTGTCCAAGTTGTCAGCGAC |
| *SKN7* | RP | TCGGTATGGCTACGTCTGAC |
| *GSH1* | FP | GAATGATAAGGCGCCACTGG |
| *GSH1* | RP | GCAATTCACCGCTTGCTCTA |
| *BCK1* | FP | CTGGAACGCTTGAAAACCCA |
| *BCK1* | RP | CAATAATTGAGTGCGGCCGA |
| *SWI4* | FP | AATGCATCGATCAACCAGCC |
| *SWI4* | RP | TTGTTGCTGCCGTTACTGTC |
| *SWI6* | FP | AACCAACAACAACAGCACGT |
| *SWI6* | RP | AGTACCTGCCCCAAAATCCA |

# Appendix 3: PCR Validation Results

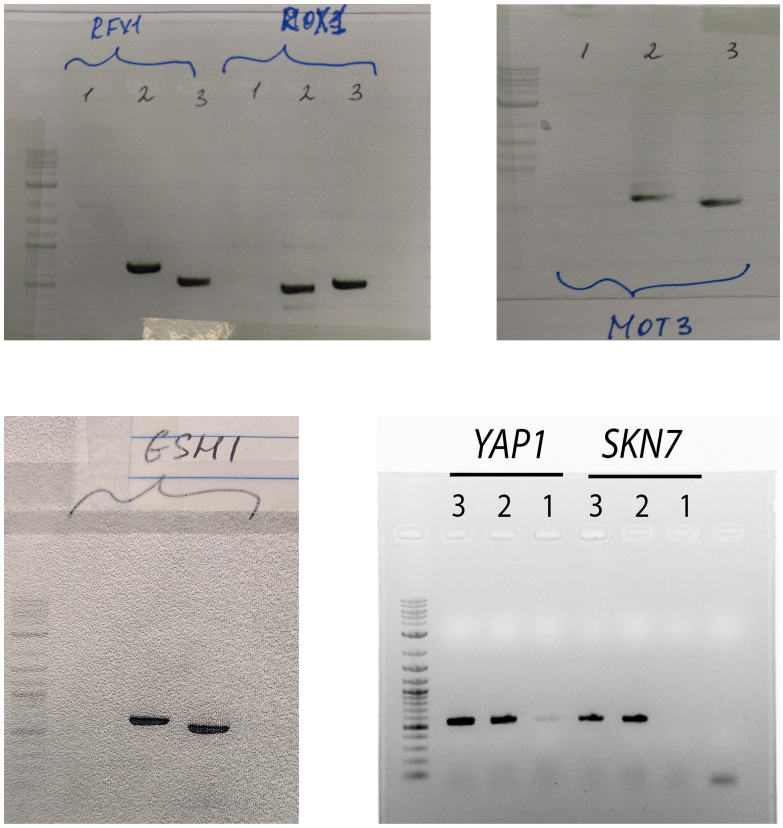



**Figure 2:** Gels of PCR-validated oxidative stress response and RNR3 repressor genes. 1) DMA or DAmP mutants' genome amplified with gene of interest's internal primers; 2) yor202w$\Delta$'s genome amplified with internal primers (technical positive control); 3) DMA or DAmP mutants' genome amplified with a different set of primers (genome extraction positive control).

## Appendix 4: Top Outliers and Reporter Categories

**Table 5:** Outliers with abnormal GFP and normal BFP at the three MMS conditions.

| 0% MMS | | 0.015% MMS | | 0.060% MMS | |
|---|---|---|---|---|---|
| ORF | Effect* | ORF | Effect | ORF | Effect |
| YPR065W | + | YGR057C | + | YBL071C-B | + |
| YCL009C | + | YPL205C | + | YBR171W | + |
| YER048W-A | + | YPR065W | + | YGL087C | + |
| YDR417C | + | YML007W | + | YOL087C | + |
| YLR288C | + | YBR200W | + | YKR074W | + |
| YFL016C | + | YHR206W | + | YHL015W | + |
| YDR145W | + | YLR176C | + | YPR068C | + |
| YDL007W | + | YGR104C | - | YBR134W | + |
| YLR176C | + | YNL296W | - | YML007W | + |
| | | YHR021C | - | YHR206W | + |
| | | YOR021C | - | YIL034C | - |
| | | YBL071W-A | - | YHR167W | - |
| | | YGR135W | - | YKR024C | - |
| | | YKL096C-B | - | YOL086C | - |
| | | YLR055C | - | YPR060C | - |
| | | YER074W | - | YBR109C | - |
| | | YML092C | - | YGL048C | - |
| | | YHR200W | - | YCL063W | - |
| | | YMR242C | - | YGR056W | - |
| | | YHR187W | - | YHR200W | - |
| | | YNL302C | - | YDL147W | - |
| | | YER155C | - | YPL061W | - |
| | | YMR032W | - | YKR072C | - |
| | | YHR010W | - | YIL090W | - |
| | | YKL110C | - | YDR253C | - |
| | | YKL009W | - | YDR394W | - |
| | | YEL044W | - | YGL076C | - |
| | | YOR293W | - | YDR156W | - |

* "+" refers to reporter expression greater than WT;
"-" denotes reporter expression less than WT.

**Table 6:** Top *RFX1* enhancers and suppressors at the three MMS conditions.

| 0% MMS | | 0.015% MMS | | 0.060% MMS | |
|---|---|---|---|---|---|
| ORF | Effect* | ORF | Effect | ORF | Effect |
| YJR074W | + | YGL048C | + | YFL016C | + |
| YCR028C | + | YOL136C | + | YLR150W | + |
| YGR036C | + | YDL007W | + | YGR229C | + |
| YER008C | + | YML007W | + | YFL025C | + |
| YGR229C | + | YLR150W | + | YMR060C | + |
| YGL048C | + | YDR418W | + | YER133W | + |
| YHR186C | + | YJR118C | + | YNL107W | + |
| YPR101W | + | YDR378C | + | YDL192W | + |
| YDL198C | + | YHR009C | + | YDR395W | + |
| YML007W | + | YHR186C | + | YDR176W | + |
| YMR060C | + | YPL170W | + | YBR171W | + |
| YHR009C | + | YER008C | + | YDR065W | + |
| YDR418W | + | YER112W | + | YOL136C | + |
| YER112W | + | YCR028C | + | YPR064W | + |
| YFL016C | + | YMR070W | + | YDR418W | + |
| YDR395W | + | YPR065W | + | YLR182W | + |
| YNR051C | + | YGR200C | + | YMR080C | + |
| YDL007W | + | YHR206W | + | YNL147W | + |
| YHR206W | + | YLR055C | - | YDR159W | + |
| YNL147W | + | YHR013C | - | YKR074W | + |
| YGR200C | + | YEL046C | - | YGL048C | + |
| YDR378C | + | YNR052C | - | YJR118C | + |
| YMR070W | + | YOR080W | - | YHR009C | + |
| YPR065W | + | YER068W | - | YPR101W | + |
| YNR032W | - | YPL129W | - | YER008C | + |
| YIL084C | - | YPL254W | - | YGR200C | + |
| YDR442W | - | YHR187W | - | YHR206W | + |
| YGL211W | - | YNL059C | - | YHR186C | + |
| YLR139C | - | YDR028C | - | YDR378C | + |
| YHL015W | - | YJL071W | - | YBR035C | + |
| YER111C | - | YEL051W | - | YMR070W | + |
| YKL144C | - | YKL009W | - | YPR065W | + |
| YKL016C | - | YHR100C | - | YMR116C | - |
| YKL110C | - | YKL110C | - | YJR109C | - |
| YGR157W | - | YGR159C | - | YJL071W | - |

| | | | | | |
|---|---|---|---|---|---|
| YPR133C | - | YMR116C | - | YDR028C | - |
| YNL162W | - | YOL145C | - | YMR062C | - |
| YHR111W | - | YER040W | - | YKR082W | - |
| YMR263W | - | YLL050C | - | YLL050C | - |
| YDR484W | - | YGL211W | - | YMR312W | - |
| YLR180W | - | YHR111W | - | YHR039C-B | - |
| YKR009C | - | YER110C | - | YJR007W | - |
| YBR058C-A | - | YER074W | - | YHR041C | - |
| YJL175W | - | YHR200W | - | YIL043C | - |
| YER156C | - | YKL119C | - | YER012W | - |
| YLR192C | - | YCR063W | - | YIL008W | - |
| YKR092C | - | YCL005W-A | - | YGL011C | - |

* "+" refers to reporter expression greater than WT;
"-" denotes reporter expression less than WT.

# Appendix 5: Functional Enrichment

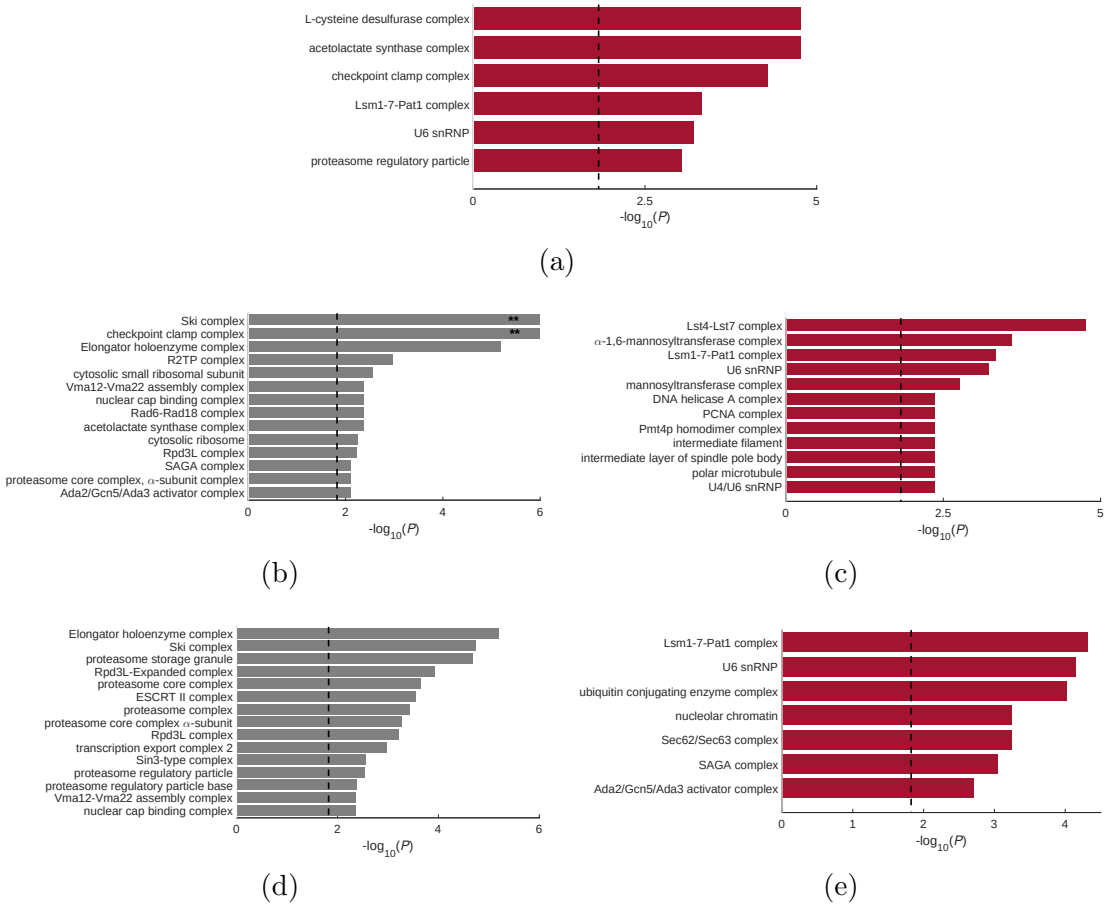



**Figure 3:** GO SLIM complex enrichment of enhancers and suppressors at (a) 0%, (b,c) 0.015%, and (d,e) 0.060% MMS. (a), (c), and (e) show enrichment of enhancers at their respective MMS doses, while (b) and (d) show enrichment of suppressors. Asterisks denote GO terms whose members are all detected outliers according to $Z_{AUC}$. Dashed line indicates $-log_{10}(P_{hypergeometric})$ at which $P_{hypergeometric} = 0.05$.

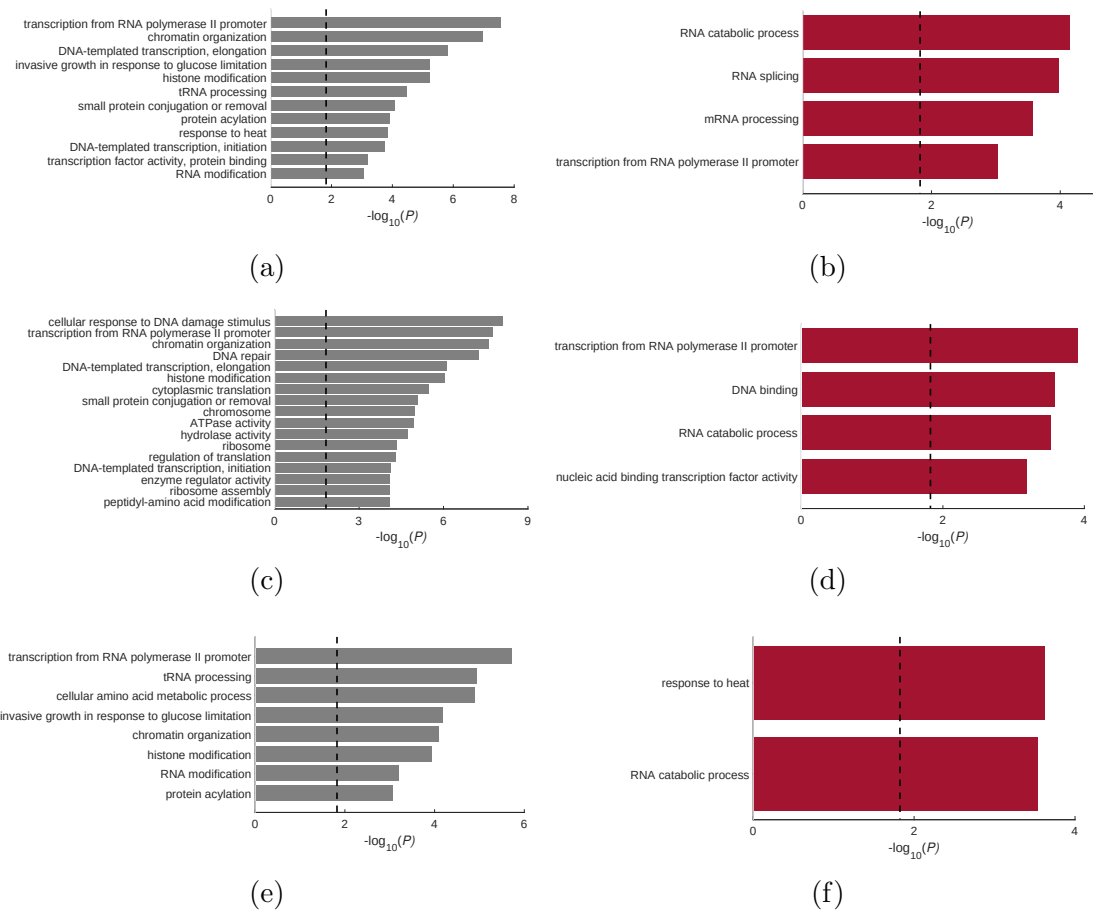



**Figure 4:** (a,b) GO SLIM functional enrichment in untreated double mutant outliers after separation into suppressors (a) or enhancers (b). (c,d) Enrichment in 0.015% MMS suppressors (c) or enhancers (d). (e,f) Enrichment in 0.060% MMS suppressors (e) or enhancers (f).

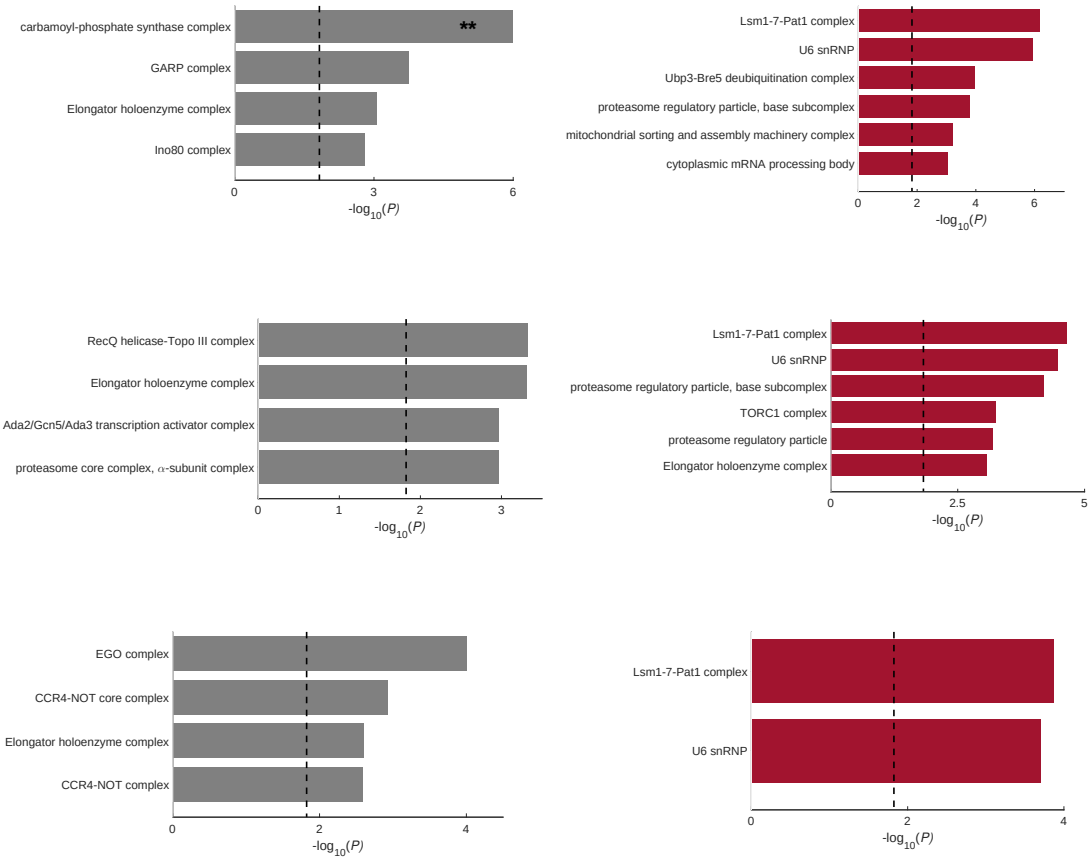



**Figure 5:** (a,b) GO SLIM complexes enrichment in untreated double mutant outliers after separation into suppressors (a) or enhancers (b). (c,d) Enrichment in 0.015% MMS suppressors (c) or enhancers (d). (e,f) Enrichment in 0.060% MMS suppressors (e) or enhancers (f).

# Appendix 6: Data Analysis Code

```
%% auc_score
% auc_score calculates the normalized AUCs and P-values of linearized GFP
% and BFP.
% Running script intializes window to enter the plate #. Make sure .MAT
% file names containing FCS data follow the same convention; change
% convention in filename; change Drug to concentration of interest
%
%
% created by:    Nada Elnour
% created on:    01.03.2016
% last Modified: 25.05.2016

clc; clear all

%% Load Data
prompt={'Enter plate number'};
name = 'Plate number';
options.Interpreter = 'tex';
defaultans = {'0'};
Plate = inputdlg(prompt,name,[1 40],defaultans,options);

filename = sprintf('gated_Plate %d_high_MMS.mat',str2num(Plate{1}));
load(filename)

Drug = 60;
channels = 1:200; % rescale from 2^16 channels to this range
ORFs = {samples.label};
event_count = cell2mat({samples.event_count});

control_idx = [1 2 95 96];

% Linearize Data
for p = 1:96
    for j = 1:size(samples(p).data,1) %event_count(p)
        BFP_samples(j,p) = { 0.1024*(exp(0.008995*samples(p).data(j,4))) };
        GFP_samples(j,p) = { 0.1024*(exp(0.008995*samples(p).data(j,3))) };
    end
end

BFP_samples(cellfun(@isempty,GFP_samples)) = {nan};
GFP_samples(cellfun(@isempty,GFP_samples)) = {nan};

gfp_samples = cell2mat(GFP_samples);
bfp_samples = cell2mat(BFP_samples);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% CDF profiles and AUC scores for outlier IDs

% Lin-binned Fluorescence Profiles

samples_BFP = histc(bfp_samples,channels);
samples_GFP = histc(gfp_samples,channels);

% CDFs of BFP and GFP
```

```
for s = 1: size(gfp_samples,2)
    samples_cBFP(:,s) = cumsum(samples_BFP(:,s)/sum(samples_BFP(:,s)));
    samples_cGFP(:,s) = cumsum(samples_GFP(:,s)/sum(samples_GFP(:,s)));
end

gfp_cont = samples_GFP(:,control_idx);
bfp_cont = bfp_samples(:,control_idx);

weighted_gfp = bsxfun(@times, samples_cGFP, channels');
auc_gfp = trapz(weighted_gfp);

auc_ctrls = auc_gfp(:,control_idx);
M = mean(auc_ctrls);
V = var(auc_ctrls);

auc_norm = (M-auc_gfp)/sqrt(V)/sqrt(4); % Z_{AUC} score

for i = 1:length(auc_norm)
    if auc_norm(i) < 0
        p_vals(i) = 1-erf(-auc_norm(i));
    else
        p_vals(i) = 1-erf(auc_norm(i));
    end
end

p_vals(isnan(p_vals)) = 1;

id = event_count > 500;
low = ORFs(~id);
ORFs = ORFs(id);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Save Variables
for x = 1:length(ORFs)
    ORFs{x} = regexprep(ORFs{x},'''','');
end

output = [ORFs; num2cell(auc_gfp(id)); num2cell(auc_norm(id)); num2cell(p_vals(id)
save(sprintf('AUC Plate %d %d%% MMS', str2num(Plate{1}), Drug),'output','channels'

display(sprintf('Plate %d done', str2num(Plate{1})))
```

```
%% go.m
% This script calculates the complement hypergeometric probability using
% the manually-curated Gene Ontology (GO) databases provided by SGD. Prior
% to intializing the script, ensure that curated databases are restructured
% by column. For example:
%       [Column 1]  [Column 2]   [Column 3]..... [Column n]
%       Systematic   Common        GO ID            GO
%         name        name
%
% created by:    Nada Elnour
% last modified: 25.05.2016 3:26 pm
%% Load curated database to workspace

drugs = {'0%', '0.015%', '0.060%'};
% goslimmapping = importdata('go_slim_mapping.xlsx');
goslimmapping = importdata('GO_Slim_Complexes.xlsx');

for i = 1: size(goslimmapping.textdata,1)
    ORFs{i} = goslimmapping.textdata{i,4};
    %ORFs{i} = goslimmapping{i,1}; % for go_slim_mapping
    ORFs{i} = regexprep(ORFs{i}, '\W','');

    database_GO{i} = goslimmapping.textdata{i,1};
    %database_GO{i} = goslimmapping{i,5}; % for go_slim_mapping
    database_GO{i} = regexprep(database_GO{i}, '''','');
end

database_GO_unique = unique(database_GO); % unique GO terms in the entire
                                          % database

for i = 1:length(database_GO_unique)
    tmp = {database_GO_unique{i}};
    membership = strcmp(database_GO, tmp);
    ORFs_matched{i} = ORFs(membership);
end

background_list = [database_GO_unique; ORFs_matched]';

% Load list of hits to workspace
condition = {'no','low','high'};
compiled_pvals = [];

for c = 1:length(condition)

    filename = sprintf('outliers_%s_mms_peak.xlsx',condition{c});
    ranked = importdata(filename);

    for i = 1:length(ranked)
        ranked{i} = regexprep(ranked{i}, '\W','');
    end

    for i = 1:length(ORFs_matched)
        temp_background = background_list{i,2} ;
        temp = ranked;

        [C ia ib] = intersect(temp, temp_background);
        query_hits{i} = C;
        number_of_hits(i) = numel(query_hits{i});
```

```
    end

    N = length(ranked); % total # of genes in query
    M = length(unique(ORFs)); % total # of S. cerevisiae ORFs

    for i = 1:length(database_GO_unique)
        X(i) = number_of_hits(i); % # of genes that match GOid from input
                                  % query
        K(i) = length(background_list{i,2});% total # of genes in GOid from
                                            % SGD
        p_vals(i) = hygecdf(X(i), M, K(i), N, 'upper')';
    end

    compiled_pvals = [compiled_pvals; p_vals];

    for i = 1:length(background_list)
        Output{i,1} = background_list{i,1};
        Output{i,2} = background_list{i,2};
        Output{i,3} = number_of_hits(i);
        Output{i,4} = p_vals(i);
    end
    %% Save

    fields = {'GOTerm' 'FullListofGenes' 'NumberofHits' 'P_vals'};
    Output = cell2struct(Output, fields, 2);

    save(sprintf('GO_Slim_lowpeak_complex_%smms', condition{c}), 'Output')

    clear Output
end
```

```
%% emap.m
% escore.m  calcuates the logarithmic multiplicative \epislon score used
% to infer statistical genetic interactions. It follows the form developed
% and published by Phenix et al. (2013).
%           - sko_orfs = ORFs of single-mutants
%           - dko_orfs = ORFs of double-mutants
%           - sko = vector of single-mutant fitness measures
%           - dko = vector of double-mutant fitness measures
%
% created by:    Nada Elnour
% last modified: 25.05.2016 3:17 pm

%%
function escore = emap(sko_orfs, sko, dko_orfs, dko)

for i = 1:length(dko_orfs)
    dko_orfs{i} = regexprep(dko_orfs{i},'\W','');
end

for i = 1:length(sko_orfs)
    sko_orfs{i} = regexprep(sko_orfs{i},'\W','');
end

[common isko idko] = intersect(sko_orfs, dko_orfs,'stable');
tmp_sko = sko(isko,:);
tmp_dko = dko(idko,:);
tmp = 'YLR176C';
rfx1 = sko(find(strcmp(sko_orfs,tmp)),:);

controls = {'YOR202W','YEL021W','YDL227C','YCL018W'};

tmp = ismember(sko_orfs,controls);
control_data = sko(tmp);
f_wt = mean(control_data);

escore = log10((f_wt*tmp_dko(:,1))./(tmp_sko(:,1)*rfx1(1)));
end
```

# Appendix 7: OD-dependent MMS-induced Fluorescence

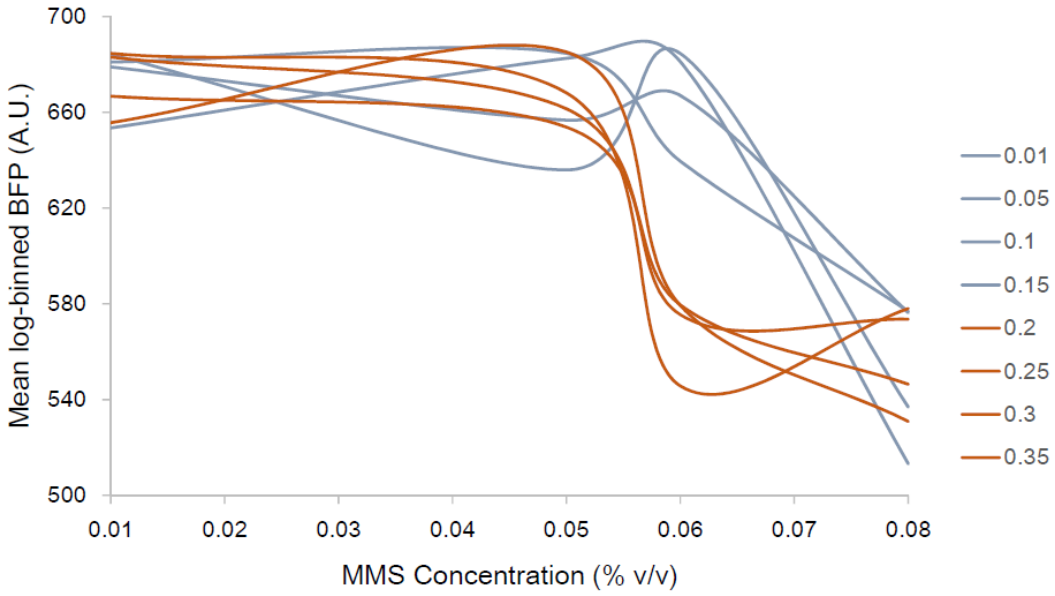



**Figure 6:** yEBFP fluorescence of a BY4741 variant containing the dual-reporter insert displays the effect of population density on the response to MMS. Population density is inferred via $OD_{600}$. An overnight culture was diluted in SM complete to the $OD_{600}$ concentrations depicted in the legend. Diluted samples were incubated for 3 h at 30 ˚C, prior to their transfer to a plate containing MMS. The plate was incubated for 5 h at 30 ˚C. Flow cytometry screening was done after diluting the samples 1:4 in NaCitrate (pH 8.0).

# Appendix 8: Bimodality Test Results

**Table 7:** Results of positive selection of 28 strains on agar media containing ClonNat, Canavanine, URA dropout, or G418, and their fluorescence distributions at the three MMS conditions.

| | | Selection | | | | 0% MMS | | 0.015% MMS | | 0.060% MMS | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ORF | Name | NAT | Canav | URA- | G418 | Bimodal * | Inconsistency * | Bimodal | Inconsistency | Bimodal | Inconsistency |
| YLR119W | SRN2 | + | + | + | + | 0 | 1 | 0 | 1 | 0 | 1 |
| YML081W | TDA9 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YOR076C | SKI7 | + | + | + | + | 0 | 0 | 1 | 1 | 1 | 1 |
| YOR293W | RPS10A | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YOR303W | CPA1 | + | + | + | + | 0 | 0 | 1 | 0 | 1 | 0 |
| YGL186C | TPN1 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YMR269W | TMA23 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YOL145C | CTR9 | + | + | + | + | 0 | 1 | 1 | 0 | 1 | 0 |
| YKL009W | MRT4 | + | + | + | + | 0 | 0 | 1 | 0 | 0 | 0 |
| YKL016C | ATP7 | + | + | + | + | 0 | 0 | 1 | 0 | 0 | 0 |
| YKL143W | LTV1 | + | + | + | + | 1 | 1 | 0 | 0 | 0 | 0 |
| YLR420W ** | URA4 | + | + | - | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YGL048C | RPT6 | + | + | + | + | 0 | 0 | 0 | 1 | 0 | 0 |
| YKR082W | NUP133 | + | + | + | + | 0 | 0 | 1 | 0 | 0 | 0 |
| YHR186C | KOG1 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YKR092C | SRP40 | + | + | + | + | 0 | 1 | 1 | 1 | 0 | 1 |
| YJR118C | ILM1 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YJR145C | RPS4A | + | + | + | + | 0 | 1 | 1 | 1 | 1 | 1 |
| YJR074W | MOG1 | + | + | + | + | 0 | 1 | 0 | 1 | 1 | 1 |
| YLR192C | HCR1 | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YNL071W | LAT1 | + | + | + | + | 0 | 1 | 0 | 1 | 0 | 1 |
| YDR442W | YDR442W | + | + | + | + | 0 | 0 | 1 | 0 | 1 | 0 |
| YDR484W | VPS52 | + | + | + | + | 0 | 0 | 1 | 0 | 1 | 0 |
| YDL083C | RPS16B | + | + | + | + | 0 | 1 | 0 | 1 | 0 | 1 |
| YGL127C | SOH1 | + | + | + | + | 0 | 0 | 0 | 0 | 1 | 0 |
| YLR376C | PSY3 | + | + | + | + | 0 | 1 | 1 | 0 | 0 | 1 |
| YLR235C | YLR235C | + | + | + | + | 0 | 0 | 0 | 0 | 0 | 0 |
| YJL124C | LSM1 | + | + | + | + | 0 | 1 | 1 | 1 | 1 | 1 |

* "0" refers to "False"; "1" refers to "True"

** URA4 did not display bimodality, but was included because both GFP and BFP displayed low fluorescence consistently

(1) Mendel’s Principles of Heredity. Cambridge University Press, 1909.

(2) Flow Cytometry: A Practical Approach, 3rd Edition. (Practical Approach Series) Oxford University Press, 2000.

(3) Flow Cytometry: First Principles, 2 ed. John Wiley & Sons, Inc., 2001.

(4) Introduction to Flow Cytometry, 1 ed. Cambridge University Press, 2004.

(5) Yeast Functional Genomics and Proteomics: Methods and Protocols. Humana Press, 2009, ch. Reporter-Based Synthetic Genetic Array Analysis: A Functional Genomics Approach for Investigating the Cell Cycle in *Saccharomyces cerevisiae*.

(6) Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J., and Stratton, M. R. Deciphering signatures of mutational processes operative in human cancer. *Cell Reports* **3**, 1 (2013), 246—59.

(7) Allen, J., Zhou, Z., Siede, W., Friedberg, E., and Elledge, S. The *SAD1/RAD53* protein kinase controls multiple checkpoints and DNA damage-induced transcription in yeast. *Genes Dev.* **8** (1994), 2416–2428.

(8) Rowe, L. A., Degtyareva, N., and Doetsch, P. W. Yap1: a DNA damage responder in *Saccharomyces cerevisiae. Mech. Age. Dev*. **133**, 4 (2012), 147–156.

(9) Baliga, N. S., Bjork, S. J., Bonneau, R., Pan, M., Iloanusi, C., Kottemann, M. C., Hood, L., and DiRuggiero, J. Systems level insights into the stress response to UV radiation in the halophilic archaeon *Halobacterium NRC-1. Gen. Res.* **14**, 2 (2004), 1025–1035.

(10) Bandyopadhyay, S., Mehta, M., Kuo, D., Sung, M.-K., Chuang, R., Jaehnig, E. J., Bodenmiller, B., Licon, K., Copeland, W., Shales, M., Fiedler,

D., Janusz, Dutkowski, Guénolé, A., van Attikum, H., M.Shokat, K., Kolodner, R. D., Huh, W.K., Aebersold, R., Keogh, M.C., Krogan, N. J., and Ideker, T. Rewiring of genetic networks in response to DNA damage. *DNA Repair* **330**, 6009 (2010), 1385–1389.

(11) Bartek, J., and Lukas, J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell* **3**, 5 (2003), 421–429.

(12) Begley, U., Dyavaiah, M., Patil, A., Rooney, J. P., Renzo, D. D., Young, C. M., Conklin, D. S., Zitomer, R. S., and Begley, T. J. Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol. Cell* **28** (2007), 860–870.

(13) Bhaumik, S. R., and Green, M. R. Differential requirement of SAGA components for recruitment of TATA-box-binding protein to promoters *in vivo*. *Mol. Cell Biol.* **22**, 21 (2002), 7365–7371.

(14) Bird, A. W., Yu, D. Y., Pray-Grant, M. G., Qiu, Q., Harmon, K. E., Megee, P. C., Grant, P. A., Smith, M. M., and Christman, M. F. Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* **419**, 6905 (2002), 411–415.

(15) Burkholder, A. C., and Hartwell, L. H. The yeast alpha-factor receptor: structural properties deduced from the sequence of the *STE2* gene. *Nuc. Acid. Res.* **13**, 23 (1985), 8463––8475.

(16) Caldecott, K. W. Single-strand break repair and genetic disease. *Nature Rev. Genet.* **9** (2008), 619–631.

(17) Chabes, A., Georgieva, B., Domkin, V., Zhao, X., Rothstein, R., and Thelander, L. Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell* **112**, 3 (2003), 391–401.

(18) Chan, C. T. Y., Pang, Y. L. J., Deng, W., Babu, I. R., Dyavaiah, M., Begley, T. J., and Dedon, P. C. Reprogramming of tRNA modifications

controls the oxidative stress response by codon-biased translation of proteins. *Nat. Comm.* **3**, 937 (2012), 1–9.

(19) Chen, H., Fan, M., Pfeffer, L. M., and Laribee, R. N. The histone H3 lysine 56 acetylation pathway is regulated by target of rapamycin (TOR) signaling and functions directly in ribosomal RNA biogenesis. *Nuc. Ac. Res.* **40**, 14 (2012), 6534–6546.

(20) Codón, A. C., Gasent-Ramírez, J. M., and Benítez, T. Factors which affect the frequency of sporulation and tetrad formation in *Sachharomyces cerevisiae* baker's yeast. *App. Environ. Microbiol.* **61**, 2 (1995), 630–638.

(21) D'Autréaux, B., and Toledano, M. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. *Nat. Rev.* **8** (2007), 813–824.

(22) de Nadal, E., Zapater, M., Alepuz1, P. M., Sumoy, L., Mas1, G., and Posas, F. The MAPK Hog1 recruits Rpd3 histone deacetylase to activate osmoresponsive genes. *Nature* **427**, 6972 (2004), 370–374.

(23) de Oliveira, F. B., Harris, M., Brazauskas, P., de Bruin, R., and Smolka, M. Linking DNA replication checkpoint to MBF cell-cycle transciption reveals a distinct class of G1/S genes. *EMBO J.* **31,** 7 (2012), 1798–1810.

(24) Deckert, J., and Struhl, K. Histone acetylation at promoters is differentially affected by specific activators and repressors. *Mol. Cell Biol.* **21**, 8 (2001), 2726–2735.

(25) Downs, J., Lowndes, N., and Jackson, S. A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature* **408** (2000), 1001–1004.

(26) Draghici, S., Khatri, P., Eklund, A. C., , and Szallasi, Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet*. **22**, 2 (2006), 101–109.

(27) Elford, H. L., Freese, M., Passamani, E., and Morris, H. P. Ribonucleotide reductase and cell proliferaion. *J. Biol. Chem.* **245** (1970), 5228–5233.

(28) Elledge, S. J., and Davis, R. W. Two genes differentially regulated in the cell cycle and by DNA-damaging agents encode alternative regulatory subunits of ribonucleotide reductase. *Genes Dev.* **4**, 5 (1990), 740–751.

(29) Elledge, S. J., Zhou, Z., Allen, J. B., and Navas, T. A. DNA damage and cell cycle regulation of ribonucleotide reductase. *BioEssays* **15**, 5 (1993), 333–339.

(30) Endo-Ichikawa, Y., Kohno, H., Tokunaga, R., and Taketani, S. Induction in the gene *RNR3* in *Saccharomyces cerevisiae* upon exposure to different agents related to carcinogenesis. *Biochem. Pharm.* **50**, 10 (1995), 1695–1699.

(31) Fisher, R. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edin*. **52**, 5 (1918), 399—433.

(32) Gasch, A. P., Huang, M., Metzner, S., Botstein, D., Elledge, S. J., and Brown, P. O. Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. *Mol. Bio. Cell* **12**, 10 (Jan 2001), 2987––3003.

(33) Hagen, D. C., Mccaffrey, G., and Sprague, G. F. Evidence the yeast *STE3* gene encodes a receptor for the peptide pheromone a factor: gene sequence and implications for the structure of the presumed receptor. *PNAS* **83**, 5 (1986), 1418––1422.

(34) Hendry, J. A high-throughput screen for novel genes involved in maintaining genome stability and the DNA damage response pathway in *Saccharomyces cerevisiae*. Master’s thesis, University of Toronto, Canada, March 2015.

(35) Hendry, J. A., Tan, G., Ou, J., Boone, C., and Brown, G. W. Leveraging dna damage response signaling to identify yeast genes controlling genome stability. *G3* **5**, 3 (2015), 997–1006.

(36) Herskowitz, I. Map kinase pathways in yeast: For mating and more. *Cell* **80**, 2 (1995), 187—197.

(37) Hickman, M. J., and Winston, F. Heme levels switch the function of Hap1 *Saccharomyces cerevisae* between transcriptional activator and transcriptional repressor. *Mol. Cell. Biol.* **27** (2007), 7417–7424.

(38) Hoejmakers, J. Genome maintenance mechanisms for preventing cancer. *Nat. Rev. Genet.* **411**, 6835 (2001), 366–374.

(39) Huang, M., and Elledge, S. J. Identification of *RNR4*, encoding a second essential subunit of ribonuceotide reductase in Saccharomyces cerevisiae. *Mol. Cell Biol.* **17**, 10 (1997), 6105–6113.

(40) Huang, M., Zhou, Z., and Elledge, S. J. The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. *Cell* **94**, 5 (1998), 595—605.

(41) Huisinga, K. L., and Pugh, B. F. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* **13**, 4 (2001), 573–585.

(42) Jaehnig, E., Kuo, D., Hombauer, H., Ideker, T. G., and Kolodner, R. D. Checkpoint kinases regulate a global network of transcription factors in response to DNA damage. *Cell* **4**, 1 (2013), 174–188.

(43) Kadosh, D., and Struhl, K. Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. *Mol. Cell. Biol.* **18**, 9 (1998), 5121—5127.

(44) Kapoor, P., Bao, Y., Xiao, J., Espejo, A., Yang, L., Bedford, M. T., Peng, G., and Shen, X. Phosphorylation-dependent enhancement of Rad53 kinase activity through the INO80 chromatin remodeling complex. *Mol. Cell* **58**, 5 (2015), 863–869.

(45) Klinkenberg, L. G., Mennella, T. A., Luetkenhaus, K., and Zitomer, R. S. Combinatorial repression of the hypoxic genes of *Saccharomyces cerevisiae* by DNA binding proteins rox1 and mot3. *Eukaryot. Cell* **4**, 4 (2005), 649–660.

(46) Klinkenberg, L. G., Webb, T., and Zitomer, R. S. Synergy among differntially regulated repressors of the ribonucleotide diphosphate reductase genes of *Saccharomyces cerevisiae*. *Eukaryot. Cell* **5**, 7 (2006), 1007–1017.

(47) Kondo, N., Takahashi, A., Ono, K., and Ohnishi, T. DNA damage induced by alkylating agents and repair pathways. *J. Nuc. Acids* **2010** (2010), 543531.

(48) Krogan, N. J., and Greenblatt, J. F. Characterization of a six-subunit holo-elongator complex required for the regulated expression of a group of genes in *Saccharomyces cerevisiae. Mol. Cell Biol.* **21**, 23 (2001), 8203–8212.

(49) Krokan, H. E., Standal, R., and Slupphaug, G. DNA glycosylases in the base excision repair of DNA. *Biochem. J.* **325**, 1 (1997), 1–16.

(50) Lee, J., Godon, C., Lagniel, G., Spector, D., Garin, J., Labarre, J., and Toledano, M. B. Yap1 and Skn7 control two specialized oxidative stress response regulons in yeast. *J. Biol. Chem.* **274**, 23 (1999), 16040–16046.

(51) Li, B., and Reese, J. C. Ssn6-Tup1 regulates *RNR3* by positioning nucleosomes and affecting the chromatin structure at the upstream repression sequnce. *J. Biol. Chem.* **276**, 36 (2001), 33788–33797.

(52) Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 6422 (1993), 709–715.

(53) Longhese, M. P., Foiani, M., Muzi-Falconi, M., Lucchini, G., and Plevani, P. DNA damage checkpoint in budding yeast. *EMBO J.* **17**, 19 (1998), 5525–5528.

(54) Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A. S. H., and Helleday, T. Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable *in vivo* DNA double-strand breaks. *Nuc. Acid. Res.* **33**, 12 (2005), 3799–3811.

(55) Ma, L.-H., Takanishi, C. L., and Wood, M. J. Molecular mechanism of oxidative stress perception by the Orp1 protein. *J. Biol. Chem.* **282**, 43 (2007), 31429–31436.

(56) Marchetti, M. A., Kumar, S., Hartsuiker, E., Maftahi, M., Carr, A. M., Freyer, G. A., Burhans, W. C., and Huberman, J. A. A single unbranched S-phase DNA damage and replication fork blockage checkpoint pathway. *PNAS* **99**, 11 (2002), 7472–7477.

(57) Mazumder, A., Pesudo, L. Q., McRee, S., Bathe, M., and Samson, L. D. Genome-wide single-cell-level screen for protein abundance and localization changes in response to DNA damage in *S. cerevisiae. Nuc Acid Res* **41**, 20 (2013), 9310–9324.

(58) Mazumder, A., Tummler, K., Bathe, M., and Samson, L. D. Single-cell analysis of ribonucleotide reductase transcriptional and translational response to DNA damage. *Annu. Rev. Biochem.* **33**, 3 (2012), 635–642.

(59) Meas, R., Smerdon, M. J., and Wyrick, J. J. The amino-terminal tails of histones H2A and H3 coordinate efficient base excision repair, DNA damage signaling and postreplication repair in *Saccharomyces cerevisiae. Nuc. Acids Res.* **43**, 10 (2015), 4990–5001.

(60) Mennella, T. A., Klinkenberg, L. G., and Zitomer, R. S. Recruitmnt of Tup1-Ssn6 by yeast hypoxic genes and chromatin-independent exclusion of TATA binding protein. *Eukaryot. Cell* **2**, 6 (2003), 1288–1303.

(61) Morrison, A. J., Highland, J., Krogan, N. J., Arbel-Eden, A., Greenblatt, J. F., Haber, J. E., and Shen, X. INO80 and γ-H2AX interaction links ATP-dependent chromatin remodeling to dna damage repair. *Cell* **119** (2004), 767–775.

(62) na Diaz, J. P., and Jiricny, J. Mammalian mismatch repair: error-free or error-prone? *Trends Biochem. Sci.* **37**, 5 (2012), 206–214.

(63) Neuman, R. J., and Rice, J. P. Two-locus models of disease. *Genet. Epidem.* **9**, 5 (1992), 347—365.

(64) Nicole Hustedt, S. M. G., and Shimada, K. Replication checkpoint: Tuning and coordination of replication forks in S phase. *Genes* **4**, 3 (2013), 388–434.

(65) Nordlund, P., and Reichard, P. Ribonucleotide reductase. *Annu. Rev. Biochem.* **75** (2006), 681–706.

(66) Peter, M., Gartner, A., Horecka, J., Ammerer, G., and Herskowitz, I. FAR1 links the signal transduction pathway to the cell cycle machinery in yeast. *Cell* **73**, 4 (1993), 747–760.

(67) Peter, M., and Herskowitz, I. Direct inhibition of the yeast cyclin-dependent kinase Cdc28-Cln by Far1. *Science* **265**, 5176 (1994), 1228–1231.

(68) Phenix, H., Morin, K., Batenchuk, C., Parker, J., Abedi, V., Yang, L., Tepliakova, L., Perkins, T. J., and Kaern, M. Quantitative epistasis analysis and pathway inference from genetic interaction data. *PLoS Comp. Biol.* **7**, 5 (2011), e1002048.

(69) Rana, A., Seinen, E., Siudeja, K., Muntendam, R., Srinivasan, B., van der Want, J. J., Hayflick, S., Reijngoud, D. J., Kayser, O., and Sibon, O. C. M. Pantethine rescues a Drosophila model for pantothenate kinase–associated neurodegeneration. *PNAS* **107**, 15 (2010), 6988—6993.

(70) Robertson, A. B., Klungland, A., Rognes, T., and Leiros, I. DNA repair in mammalian cells: base excision repair. *Cell. Mol. Life Sci.* **66** (2009), 981–993.

(71) Rowe, L. A., Degtyareva, N., and Doetsch, P. W. DNA damage-induced reactive oxygen species (ROS) stress response in *Saccharomyces cerevisiae. Free Radic. Biol. Med.* **45**, 8 (2008), 1167–1177.

(72) Rowley, N., Prip-Buus, C., Westermann, B., Brown, C., Schwarz, E., Barrell, B., and Neupert, W. Mdj1p, a novel chaperone of the DnaJ family, is involved in mitochondrial biogenesis and protein folding. *Cell* **77**, 2 (1994), 249–259.

(73) Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 5235 (1995), 467–470.

(74) Schenk, L., Meinel, D. M., Strässer, K., and Gerber, A. P. La-motif-dependent mRNA association with Slf1 promotes copper detoxification in yeast. *RNA* **18**, 3 (2012), 449–461.

(75) Sharma, V. M., Tomar, R. S., Dempsey, A. E., and Reese, J. C. Histone deacetylases RPD3 and HOS2 regulate the transcriptional activation of DNA damage-inducible genes. *Mol. Cell Biol.* **27**, 8 (2007), 3199–3210.

(76) Stolz, J., and Sauer, N. The fenpropimorph resistance gene *FEN2* from *Saccharomyces cerevisiae* encodes a plasma membrane $H^+$-pantothenate symporter. *J. Biol. Chem.* **274**, 26 (1999), 18747–18752.

(77) Stratton, M. Patterns of somatic mutation in human cancer genomes. *Eur. J Can. Supp.* **6**, 9 (2008), 6.

(78) Tamble, C. M., Onge, R. P. S., Giaever, G., Nislow, C., Williams, A. G., Stuart, J. M., and Lokey, R. S. The synthetic genetic interaction network reveals small molecules that target specific pathways in *Sacchromyces cerevisiae. Mol. BioSyst.* **7** (2011), 2019–2030.

(79) Tang, H.-M. V., Siu, K.-L., Wong, C.-M., and Jin, D.-Y. Loss of yeast peroxiredoxin Tsa1p induces genome instability through activation of the DNA damage checkpoint and elevation of dNTP levels. *PLoS Genet.* **5**, 10 (2009), e1000697.

(80) Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R.,

and Hopfner, K. P. Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell* **154**, 6 (2013), 1207–1219.

(81) Tsang, C. K., Liu, Y., Thomas, J., Zhang, Y., and Zheng, X. Superoxide dismutase 1 acts as a nuclear transcription factor to regulate oxidative stress resistance. *Nat. Comm.* **5**, 3446 (2014), 1–11.

(82) van Attikum, H., Fritsch, O., Hohn, B., and Gasser, S. M. Recruitment of the INO80 complex by H2A phosphorylation links ATP-dependent chromatin remodeling with DNA double-strand break repair. *Cell* **119** (2004), 777–788.

(83) Vidal, M., Strich, R., Esposito, R. E., and Gaberl, R. F. *RPD1* (*SIN3/UME4*) is required for maximal activation and repression of diverse yeast genes. *Mol. Cell. Biol.* **11**, 12 (1991), 6306—6316.

(84) Wagner, I., Arit, H., van Dyck, L., Langer, T., and Neupert, W. Molecular chaperones cooperate with PIM1 protease in the degradation of misfolded proteins in mitochondria. *EMBO J.* **13**, 21 (1994), 5135–5145.

(85) Whitehead, K., Kish, A., Pan, M., Kaur, A., Reiss, D. J., King, N., Hohmann, L., DiRuggiero, J., and Baliga, N. S. An integrated systems approach for understanding cellular responses to gamma radiation. *Mol. Syst. Biol.* **2** (2006), 47.

(86) Winkler, G. S., Kristjuhan, A., Erdjument-Bromage, H., Tempst, P., , and Svejstrup, J. Q. Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels in vivo. *PNAS* **99,** 6 (2007), 3517–3522.

(87) Winkler, G. S., Petrakis, T. G., Ethelberg, S., Tokunaga, M., Erdjument-Bromage, H., Tempst, P., and Svejstrup, J. Q. RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes. *J. Biol. Chem.* **276** (2001), 32743–32749.

(88) Workman, C. T., Mak, H. C., McCuine, S., Tagne, J.-B., Agarwal, M., Ozier, O., Begley, T. J., Samson, L. D., and Ideker, T. A systems approach to

mapping DNA damage response pathways. *Science* **312**, 5776 (2006), 1054–1059.

(89) Yao, R., Zhang, Z., An, X., Bucci, B., Perlstein, D. L., Stubbe, J., and Huang, M. Subcellular localization of yeast ribonucleotide reductase regulated by the DNA replication and DNA damage checkpoint pathways. *PNAS* **100**, 11 (2003), 6628–6633.

(90) Yoshimoto, H., Ohmae, M., and Yamashita, I. The *Saccharomyces cerevisiae GAM2/SIN3* protein plays a role in both activation and repression of transcription. *Mol. Gen. Genet.* **233**, 1–2 (1992), 327—330.

(91) Zegerman, P., and Diffley, J. F. X. DNA replication as a target of the DNA damage checkpoint. *DNA Rep.* **8**, 9 (2009), 1077–1088.

(92) Zhang, H., Kruk, J., and Reese, J. Dissection of coactivator requirement at *RNR3* reveals unexpected contributions from TFIID and SAGA. *J. Biol. Chem.* **283**, 41 (2008), 27360—-27368.

(93) Zhang, M., Zhang, C., Li, J., Hanna, M., Zhang, X., Dai, H., and Xiao, W. (2011) Inacivation of *YAP1* enhances sensitivity of the yeast *RNR3-lacZ* genotoxicity testing system to a broad range of DNA-damaging agents. *Toxicol. Sci.* **120**(2) : 310–21.

(94) Zhao, X., and Rothstein, R. The Dun1 checkpoint kinase phosphorylates and regulates the ribonucleotide reductase inhibitor Sml1. *PNAS* **99**, 6 (2002), 3746—3751.

(95) Curtin, N.J. (2012) DNA repair dysregulation from cancer driver to therapeutic target. *Nat. Rev.* **12**: 801–817.

(96) Daley, J.M., Laan, R.L.V., Suresh, A., and Wilson, T.E. DNA joint dependence of pol X family polymerase action in nonhomologous end joining. *J. Biol. Chem.* **280** (2005): 29030–29037.

(97) Kim, C. Paulus, B.F., Wold, M.S. Interactions of human replication protein A with oligonucleotides. *Biochem.* **33** (1994): 14197–14206.

(98) Kumaran, S., Kozlov, A.G., Lohma.n, T.M. Saccharomyces cerevisiae replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochem.* **45** (2006): 11958–11973.

(99) Gietz, R.D., and Schiestl, R.H. High-efficicency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protocol.* **2**, 1 (2007): 31–34.