2020

A Study of Inactive Enzyme-Homologues: The Biochemical and Biological Function of ECM14 in Saccharomyces Cerevisiae

Robert Christian McDonald
Andrews University, mcdonalr@andrews.edu

ABSTRACT

A STUDY OF INACTIVE ENZYME-HOMOLOGUES:
THE BIOCHEMICAL AND BIOLOGICAL FUNCTION OF ECM14
IN *SACCHAROMYCES CEREVISIAE*

By

R. Christian McDonald

Chair: Peter J. Lyons, PhD

ABSTRACT OF GRADUATE STUDENT RESEARCH

Thesis

Andrews University

College of Arts and Sciences

Title: A STUDY OF INACTIVE ENZYME-HOMOLOGUES: THE BIOCHEMICAL AND BIOLOGICAL FUNCTION OF ECM14 IN SACCHAROMYCES CEREVISIAE

Name of Researcher: R. Christian McDonald

Name and Degree of Faculty Chair: Peter J. Lyons, PhD

Date Completed: September 2020

Like most major enzyme families, the M14 family of metallocarboxypeptidases (MCPs) contains several pseudoenzymes predicted to lack enzyme activity and with unknown molecular function. The genome of the yeast *Saccharomyces cerevisiae* encodes only one member of the M14 MCP family, a pseudoenzyme named Ecm14 proposed to function in the extracellular matrix. Ecm14 is found throughout ascomycete fungi, with a group of related pseudoenzymes found in basidiomycetes. Although the prodomain of Ecm14 can be cleaved *in vivo* and *in vitro* by endopeptidases, suggesting an activation mechanism, no activity has been detected using standard carboxypeptidase substrates.

In order to examine the function of Ecm14 using an unbiased screen, a series of synthetic lethal assays were performed. Because biochemical pathways typically involve

multiple key gene components, many of these key gene players provide redundancy if other genes experience deleterious mutations. The synthetic lethal screen identifies novel mutants whose survival depends on a particular gene of interest. Approximately 27,000 yeast colonies were screened; however, only twenty-two putative ecm14 synthetic lethal mutants were identified. Further analysis showed many to be synthetic lethal with auxotrophic marker (plasmid reporting and selection) genes and synthetic lethality due to multiple mutations, suggesting that there are few, if any, single *S. cerevisiae* genes that present synthetic lethal interactions with ecm14. Recent research by others suggests a role for Ecm14 in processes such as vesicle-mediated transport and aggregate invasion, a fungal process that has been selected against in modern laboratory strains of *S. cerevisiae*. Thus, Ecm14 is likely an important, conserved, and "activatable" component of extracellular processes that are not critical to most modern yeast strains.

Andrews University

College of Arts and Sciences

A STUDY OF INACTIVE ENZYME-HOMOLOGUES: THE BIOCHEMICAL AND BIOLOGICAL FUNCTION OF ECM14 IN *SACCHAROMYCES CEREVISIAE*

A Thesis

Presented in Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

R. Christian McDonald

September 2020

A STUDY OF INACTIVE ENZYME-HOMOLOGUES: THE BIOCHEMICAL AND BIOLOGICAL FUNCTION OF ECM14 IN *SACCHAROMYCES CEREVISIAE*

A Thesis

Presented in Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

R. Christian McDonald

APPROVAL BY THE COMMITTEE:

Peter J. Lyons, PhD

Denise L. Smith, PhD

H. Thomas Goodwin, PhD

Date Approved

TABLE OF CONTENTS

PAGE:

LIST OF FIGURES

FIGURE: PAGE:

LIST OF TABLES

TABLE: PAGE:

LIST OF ABBREVIATIONS

5-FOA	5-Fluoroorotic Acid
CDS	Coding Sequence
CIAP	Calf Intestinal Alkaline Phosphatase
CPA	Carboxypeptidase A
CW/CFW	Calcofluor White
ECM	Extracellular Mutant
EMS	Ethyl Methanesulfonate
HR	Homologous Recombination
LB	Luria Broth
MCP	Metallocarboxypeptidase
ORF	Open Reading Frame
SC+5FOA	Synthetic Complete Plus 5-Fluoroorotic Acid
SC-Ura	Synthetic Complete Minus Uracil
SDM	Site-Directed Mutagenesis
SL	Synthetic Lethal
YPD/YEPD	Yeast Extract, Peptone, Dextrose
YCp	Yeast Centromere Plasmid
YPGal	Yeast Extract, Peptone, Galactose
YPGal+Ura	Yeast Extract, Peptone, Galactose, Plus Uracil
YPGal+CFW	Yeast Extract, Peptone, Galactose, Plus Calcofluor White
YPGal+Ura+CFW	Yeast Extract, Peptone, Galactose, Plus Uracil & Calcofluor White
YRp	Yeast Replicating Plasmid

ACKNOWLEDGEMENTS

Funding for this research was provided by Andrews University Faculty Research Grants to Peter J. Lyons and by Andrews University Grants in Aid of Research to R. Christian McDonald.

I wish to thank my advisor, Dr. Peter Lyons, for his expertise and support throughout my time working on this research project. Research is often tricky and filled with many unforeseen challenges and setbacks along the way, both in the laboratory and life in general. I truly appreciate both his patience and continued encouragement to savor the small successes and not to cling onto the many mistakes that will inevitably be made along the way.

I wish to thank my committee members, Dr. Tom Goodwin and Dr. Denise Smith, for their continued support and invaluable perspective throughout this project. Additionally, I wish to thank the entire faculty of the Department of Biology at Andrews University for the many opportunities afforded to me during both my undergraduate and graduate programs.

I wish to thank Dr. Melanie Dobson (Dalhousie University, Nova Scotia, Canada) for her expertise and generous gifts of starter yeast strains and various plasmids required to kickstart this project. I wish to thank Dr. Doug Koshland (University of California, Berkeley, California), and Dr. Steve Henikoff (Fred Hutchinson Cancer Research Center, Seattle, Washington) for their email correspondence and assistance in corroborating the sequence of the original pSLS1 map and its constituent gene fragments. I wish to thank Deidre and Ryan Casey from The Sequencing Center (Fort Collins, Colorado) for their expertise and willingness to accommodate our whole-genome sequencing needs on a

shoestring budget. Finally, I would also like to thank GSL Biotech, LLC for providing a short-term academic license for SnapGene which was immensely helpful for organizing and producing the high-quality DNA and plasmid maps found within this thesis.

I wish to thank my colleagues and friends who have helped me stay positive and motivated throughout this project. I am also very grateful to my family for their continued encouragement and many prayers along the way. I wish to thank my dearest wife, Kaitlin, for her unfaltering support and patience while I was avoiding completing this thesis. I simply could not have completed this without her love, support, and encouragement. Finally, I wish to thank God for the cognition and fortitude required to do research.

For Darwin McDonald

Father. Husband. Friend.

CHAPTER 1

INTRODUCTION

Genomes and the Evolution of Pseudoenzymes

Genomes contain a variety of components encoded in the ubiquitous four-base language of nucleic acids. Although genes have traditionally been defined by their ability to direct the translation of protein, many genes fundamentally lack all protein-coding ability. These non-coding pseudogenes were once thought to be "junk DNA." However, numerous investigations have demonstrated that these non-coding DNA sequences indeed play critical roles in directing protein synthesis and regulating gene expression through their RNA products. Notable examples of these ubiquitous non-coding RNAs include ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (Pink et al., 2011).

Other genes, though coding for proteins, may not actually encode for the functions that might be predicted based solely on their homology with other known and annotated sequences. Although key active site substitutions in what appear to be enzyme-coding genes might suggest a loss of catalytic activity, these substitutions might instead point towards either a presently unknown and uncharacterized substrate or some novel biochemical process.

It has been suggested that all major enzyme families across all kingdoms of life likely contain inactive enzyme-homologues (Adrain & Freeman, 2012; Eyers & Murphy, 2016; Kwon et al., 2019; Ribeiro et al., 2019). In general, these pseudoenzymes are characterized by highly conserved domains that appear structurally like those domains found in otherwise catalytically active enzymes. However, substitutions of key active site amino acids (or the absence of entire catalytic motifs altogether) in these conserved

domains suggest a loss of expected function in pseudoenzymes. From an evolutionary perspective, enzymes might evolve an altered substrate specificity, an altered reaction to be catalyzed, or, as in the extreme case of pseudoenzymes, a complete loss of catalytic activity. However, the inclusion of these pseudoenzymes in our present understanding of signal transduction networks and regulatory pathways is still relatively new and emerging within the current paradigm of systems biology. Thus, pseudoenzymes have only recently been considered as possible key players (Eyers & Murphy, 2016). In general, much work remains in understanding these ubiquitous pseudoenzymes and evaluating their usefulness as potential targets (or anti-targets) in the development of novel therapeutic agents.

The two most well documented pseudoenzymes are of the kinase and phosphatase families (Eyers & Murphy, 2013; Reiterer et al., 2014). [In general, kinases attach phosphate groups and phosphatases remove phosphate groups.] These inactive enzyme-homologues have been demonstrated to function both as allosteric modulators (either by activation or inhibition) and as non-catalytic competitors in signal transduction pathways (Eyers & Murphy, 2016). Therefore, the present evidence suggests that pseudoenzymes are likely functional, not as catalytically active enzymes, but rather as binding proteins that only transiently interact with their substrates.

The central hypothesis for understanding the evolution of pseudoenzymes first involves the genomic duplication of an "active" ancestral enzyme open reading frame followed by subsequent mutation events that, while off-selection, can repurpose a domain found in otherwise catalytically active enzyme (*see* **Figure *1***). These genetic copies, though being functionally redundant, lack the traditional evolutionary pressure required to maintain the otherwise catalytic motifs and phenotype(s) of their ancestral enzyme and,

thus, allow these genes to accumulate novel mutations (Bergthorsson et al., 2007; Eyers & Murphy, 2016; Ohno, 1970).

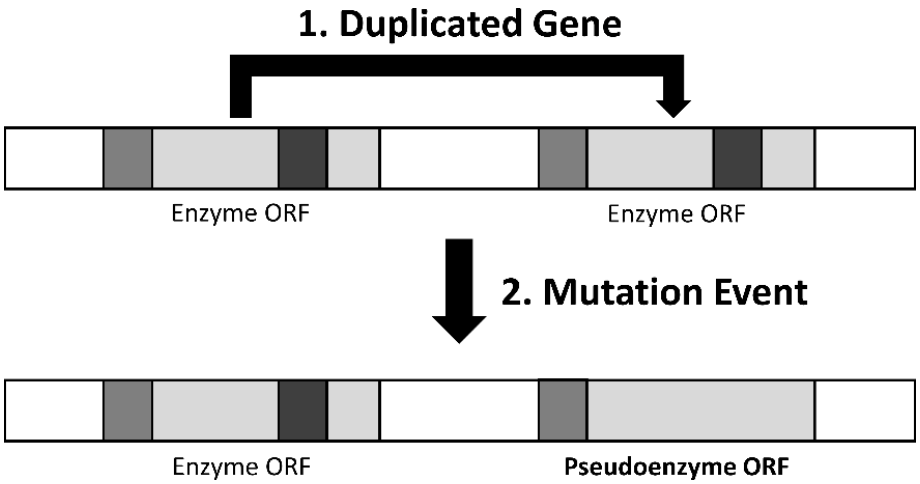

Figure 1. The evolution of pseudoenzymes from active enzyme-homologues.

The evolution of pseudoenzymes might first (**1**) involves the genomic duplication of an 'active' ancestral enzyme open reading frame followed by (**2**) subsequent mutation events that, in turn, repurpose a domain found in otherwise catalytically active enzyme in some potentially novel role.

The two primary models for understanding how genes (and, thus by proxy, their protein or enzyme products) evolve novel functions through this broad mechanism of gene duplication are 1) *neofunctionalization* and 2) *subfunctionalization* (Force et al., 1999; Noda-Garcia & Barona-Gomez, 2013; Ohno, 1970). The accumulation of mutation events in a duplicated gene over time typically results in the ubiquitous non-coding pseudogenes. However, a divergence of function can occur when mutations either give rise to a novel function or cause an otherwise promiscuous ancestral enzyme (one with several distinct sub-functions) to further refine and specialize on a sub-function.

Catalytic motifs are defined in part by the presence of specific active site residues that catalyze some specific biochemical reaction(s). However, these motifs are also defined by their ability to interact with very specific ligands in potentially non-catalytic ways.

Reiterer et al. (2014) suggested that at least some active enzymes might have paradoxically evolved from pseudoenzymes. However, in the absence of selective pressures to maintain the catalytic motif in these duplicated genes, evolution is allowed to further tweak the non-catalytic function of these pseudoenzymes and, in turn, cause their sequences to further deviate over time from their original ancestral sequences.

Although previous investigations involving pseudoenzymes have suggested possible non-catalytic functions, it has been suggested that even catalytically active enzymes too might be able to perform these novel, unexpected secondary functions (so-called promiscuous or "moonlighting" enzymes) (Jeffery, 2003). Therefore, both active enzymes and their inactive homologues might share these previously uncharacterized and noncatalytic roles. Because the structural motifs of active enzymes are inherently predisposed to facilitating very specific substrate-binding interactions, these structures serve as ideal candidates to being evolutionarily repurposed and reused in some non-catalytic capacity (Eyers & Murphy, 2016). Therefore, studies involving pseudoenzymes can potentially inform our understanding of even catalytically active enzymes in general by suggesting possible functions beyond just the prototypical native catalytic function.

One possible model for understanding the function of pseudoenzymes that have been implicated in substrate binding and localization is the rhomboid protease family, which includes several active members as well as inactive enzyme-homologues. [Proteases and proteolysis are discussed in more detail in a subsequent introductory section.] Active rhomboid proteases function in the proteolysis of transmembrane proteins and the subsequent release of signaling molecules from a variety of cellular organelles (Lemberg, 2011). Active rhomboid proteases have been demonstrated to function in the epidermal

growth factor receptor signaling pathway in *C. elegans*, *Drosophila*, and mammalian cells (Adrain & Freeman, 2012; Blobel et al., 2009; Dutt et al., 2004).These active proteases cut membrane-bound growth factor precursors and liberate active growth factors which in turn bind to receptors on neighboring cells and initiate cascades of growth signals. In contrast, several inactive rhomboid pseudoproteases (e.g. Derlin-1 and iRhoms) have adopted novel, non-traditional function that seem to complement the function of active rhomboid proteases: they bind to and help destabilize substrate proteins targeted by active rhomboid proteases in the endoplasmic reticulum-associated degradation of misfolded proteins (Adrain & Freeman, 2012; Greenblatt et al., 2011; Zettl et al., 2011).

How these pseudoenzymes adapt and evolve to fill novel biological and biochemical niches in general remains to be fully understood. The broad classification of putative pseudoenzymes is largely based on batch *in silico* sequence comparisons against sequences of known catalytic domains that have a wealth of experimental validation in the literature. In other words, such comparisons only work *well* when using catalytic source sequences with very well-defined mutations that demonstrably nullify the expected catalytic activity. However, this type of comparison does suggest probable knowledge gaps and the inescapable possibility that many enzymes are actually erroneously classified as catalytically active in online enzyme databases (or, conversely, many pseudoenzymes incorrectly classified as catalytically inactive) due to a fundamental lack of targeted experimental evidence (Eyers & Murphy, 2016; Ribeiro et al., 2019).

Because pseudoenzymes are thought to be present in all major enzyme families across all domains of life, the majority of these pseudoenzymes largely remain either unidentified, misidentified, or entirely uncharacterized. For instance, adequately validating

a *complete* lack of catalytic function fundamentally requires that *all* possible substrates be known—a lofty and seemingly insurmountable requirement (Ribeiro et al., 2019). Furthermore, the obscure functional nature of these pseudoenzymes makes them inherently difficult to associate with a phenotype. However, a quantifiable phenotype is fundamentally needed in order to systematically study the biochemical and biological implications of the key active site substitutions found in these pseudoenzymes. Once a quantifiable phenotype is either identified or engineered, a series of targeted experiments can be performed that selectively interrogate each of the key active site substitutions and individually assess their necessity and/or usefulness. One goal of this study is to design such a targeted approach to the investigation of a suspected pseudoenzyme found in the baker's yeast *Saccharomyces cerevisiae*.

Although targeted experiments focused on the biochemical and biological nature of pseudoenzymes remain to be particularly challenging, there are several key strategies and techniques that can be used to build a general case for function (or the lack thereof). These strategies might include sequence comparison and protein homology modeling, classifying and grouping novel genes as being involved in broad functional classes, phylogenic tree analysis and visualization, enzymatic prodomain cleavage followed by Edman N-terminal degradation sequencing, enzyme activity assays, and *in vivo* protein localization and gene expression assays.

Although prior investigations have demonstrated that several pseudoenzymes have also adopted novel, non-catalytic functions, much work remains in order to accurately identify, clarify and characterize the function of these inactive enzyme-homologous (Greenblatt et al., 2011; Pils & Schultz, 2004; Zettl et al., 2011).

Saccharomyces cerevisiae as a Model Organism

One flexible model organism for studying the biochemical and biological nature of pseudoenzymes is the ascomycete fungi *Saccharomyces cerevisiae* (*see* ***Figure 2***). Because the yeast *S. cerevisiae* has an extensive history both as a model organism and in the baking and brewing industries, many well-demonstrated techniques and tools already exist that make its genetic manipulation both inexpensive and robust.

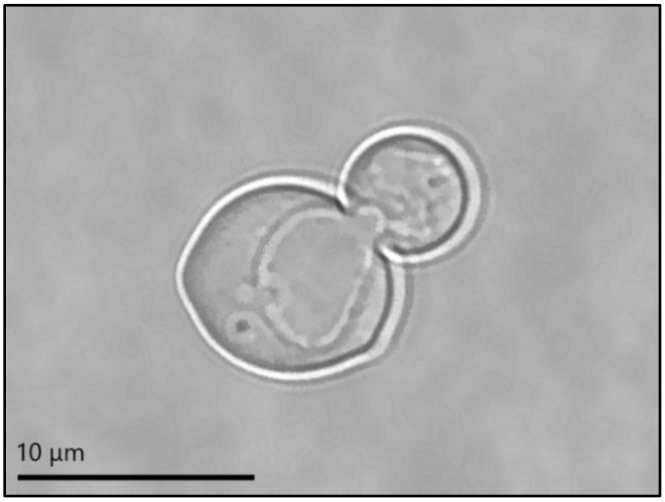

Figure 2. A budding haploid S. cerevisiae yeast cell.

A micrograph of a budding haploid *S. cerevisiae* yeast cell. The large central vacuole in the parent yeast cell can be observed poking into the daughter bud.

As a eukaryotic organism, yeast are capable of reproducing sexually and exhibit an alternation of generations between diploid, spore-producing sporophytes and haploid, gamete-producing gametophytes. Additionally, humans and yeast share over one thousand orthologous genes (~23% homologous genes to humans) (Laurent et al., 2016; Liu et al., 2017). Therefore, these features together make the yeast *S. cerevisiae* a prime candidate for high-throughput gene function studies that require both precise genetic manipulation and sexual reproduction capabilities (qualities that are of critical importance to this investigation in particular).

Yeast have considerably simpler genomes than those of higher order eukaryotes. Like bacteria, yeast can readily be transformed with plasmid DNA (Hinnen et al., 1978). These plasmids can either replicate independently (sometimes even behaving like miniature chromosomes) or integrate directly into the yeast genome, sometimes tagging or even deleting endogenous genes via homologous recombination (Duina et al., 2014). The capability to selectively target, tag, and disrupt endogenous genes with a high degree of precision has made yeast a preeminent model organism. Furthermore, the sophisticated portfolio of genomic and molecular tools available to yeast biologists further qualify *S. cerevisiae* as a premier experimental subject.

Chromosome III (315-kb) of the yeast *S. cerevisiae* genome was the first chromosome from any organism to be completely sequenced (Oliver et al., 1992). In addition to the monumental achievements related to chromosome III, the prevalence of *S. cerevisiae* in both commercial industry and human health likely motivated the worldwide collaboration to sequence the remaining 15 chromosomes (Goffeau et al., 1996). As such, the yeast genome was the first eukaryotic genome to be entirely sequenced. To date, the yeast genome has been thoroughly sequenced many times over, using every major generation of whole-genome sequencing technology including Sanger, Illumina®, PacBio®, and Oxford Nanopore® (Cherry et al., 2012; Engel et al., 2014; Goodwin et al., 2015; Matheson et al., 2017).The *Saccharomyces Genome Database* (SGD) is the prominent, online database and curated search tool for exploring the wealth of yeast genome data (Cherry et al., 2012). Whole-genome sequencing of the yeast *S. cerevisiae* genome is significant to this investigation.

Characteristics of the Fungal Kingdom

The Kingdom Fungi is a highly diverse group of heterotrophic eukaryotes that can obtain nutrients either as saprophytes or as symbionts (e.g. with plant roots as a mycorrhiza). Members of the Kingdom Fungi can range from the unicellular budding yeasts (e.g. *S. cerevisiae*) and aquatic chytrids to larger mushrooms and even the colossal Armillarians. For example, a single individual of *A. bulbosa*, with an estimated biomass of 10,000 kg, has been identified in a northern Michigan hardwood forest (Smith et al., 1992). This single clonal individual spans nearly 40 acres. Even larger *Armillarians* have been identified in the Pacific Northwest (Ferguson et al., 2003). In general, only about 100,000 fungal species have been identified to date. However, conservative estimates suggest that the fungal kingdom likely contains between 2.2 and 3.8 million species (Hawksworth & Lücking, 2017). Thus, only about 3-8% of fungal species have been named so far.

Unlike members of the other three eukaryotic kingdoms (Animalia, Plantae, and Protista), fungi characteristically contain cell walls that utilize chitin polymers. Additionally, chitin is the primary structural component in the exoskeleton of arthropods. Additionally, members of the Kingdom Plantae also contain cell walls. However, the major carbohydrate in plant cell walls is cellulose. Cellulose consists of linear chains of repeating β-1,4-linked *D*-glucose monomers. In contrast, glucose is replaced by *N*-acetylglucosamine monomers in chitin nanofibrils. Chitin plays an important role in the structure of fungal cell walls. In contrast to the more elastic β-glucan fibers, chitin confers the bulk of the wall's rigidity and its ability to withstand both chemical and physical abuse (Fesel & Zuccaro, 2016). This resiliency stems in part from the insolubility of chitin in water. [Yeast

cell wall biology is of interest in this study and is discussed in more detail in a subsequent introductory sub-section.]

Members of the Kingdom Fungi exhibit an alternation of generations between diploid, spore-producing sporophytes and haploid, gamete-producing gametophytes. As a member of the Kingdom Fungi, *S. cerevisiae*, too, exhibits this phenomenon. However, the gametophyte generation typically predominates. During both stages, the yeast *S. cerevisiae* grows via mitotic budding, a type of asexual reproduction. However, meiosis produces haploid spores from mature diploid sporophytes in response to stress during the early stages of sporulation (Neiman, 2011). Haploids are generally more efficient at reproducing and surviving. The metabolic trade-offs between these two ploidy stages suggest an adaptive function of meiosis and, thus, an underlying survival mechanism in response to stress conditions (Zörgö et al., 2013).

Although the yeast *S. cerevisiae* can reproduce sexually, outcrossing is generally quite rare due to both mating types being produced in such close proximity within a single ascus (Ruderfer et al., 2006). Therefore, the rarity of outcrossing suggests that meiosis is conserved, not as a means of introducing genetic variation, but rather as a means of enhancing phenotypic stability and repairing DNA following periods of prolonged metabolic stress (Birdsell & Wills, 2003). This inherent stability of the fungal genome further qualifies *S. cerevisiae* as a prime model organism.

Diversity of the Ascomycete Fungi

The ascomycetes represent the largest phyla of the Kingdom Fungi and consists of over 64,000 species (Hawksworth & Lücking, 2017). The unifying characteristic of Ascomycota is the presence of an ascus, a reproductive sac-like structure that contains the haploid ascospores (Neiman, 2005). The meiotic divisions during ascus formation do not culminate in cytokinesis. Instead, the meiotic nuclear products are individually captured and encased in new cell membrane and wall within the now anucleate ascus 'sac.' As a member of Ascomycota, *S. cerevisiae* too produce asci during sexual reproduction. However, the haploid stage predominates in domesticated yeast species.

The phylum Ascomycota is subdivided into the three distinct subphyla Pezizomycotina, Saccharomycotina, and Taphrinomycotina (Eriksson & Winka, 1997). The subphyla Pezizomycotina is the largest subdivision of Ascomycota and contains over 30,000 named species (Kirk et al., 2001). Pezizomycotinas are characterized by a predominately filamentous morphology and represent most of the fruiting ascomycetes. Pezizomycotinas can be found in both aquatic and terrestrial environments and as symbionts (e.g. mycorrhizae and lichens) and as pathogens (Spatafora et al., 2006; Sugiyama et al., 2006). Pezizomycotinas play significant roles in both human health and disease. For example, *Penicillium rubens* (or *P. chrysogenum*) is infamously known for its production of penicillin, one of the most widely used antibiotic agents, and *Tolypocladium inflatum* for its production of cyclosporin A, a potent immunosuppressant drug (Houbraken et al., 2011; Wenger, 1984). The World Health Organization considers both penicillin and cyclosporin A to be essential drugs within the context of even the most basic health-care systems. Finally, the pezizomycotina *Neurospora crassa*, named for the distinct axon-like

morphology of the spores, is a popular model organism in biochemistry, genetics, and molecular biology. The draft sequence of *N. crassa* was reported in 2003 and has since been thoroughly characterized and refined (Galagan et al., 2003).

The subphylum Saccharomycotina consists of the classical ascomycete 'yeasts' including *Candida albicans*, *Ashbya gossypii*, and *Saccharomyces cerevisiae* (Eriksson & Winka, 1997). Although many members of this subphylum have tremendous commercial applications, some are indeed pathogenic and play significant roles in human health and disease. For instance, infection by *C. albicans* is the leading cause of mycosis-related death in the United States (Pfaller & Diekema, 2007). Although *C. albicans* is usually a commensal resident of the human gastrointestinal tract, immunocompromised and/or genetically predisposed individuals are especially susceptible to developing deadly candidiasis infections (Martins et al., 2014). *A. gossypii* has been the primary industrial commercial producer of riboflavin (vitamin B2) for many decades (Aguiar et al., 2015).

Finally, the subphylum Taphrinomycotina consists of a diverse variety of taxa including Neolectomycetes, Pneumocystidomycetes, and Schizosaccharomycetes (Eriksson & Winka, 1997). The Neolecta species are characterized by fleshly yellowish-orange unbranched (or branched) fruiting bodies (Landvik et al., 2003). A notable member of the Pneumocystidomycetes is *Pneumocystis jirovecii*, a commensal resident in the human lungs. Similar to candidiasis infections, *P. jirovecii* can cause deadly Pneumocystis pneumonia in immunocompromised individuals (Aliouat-Denis et al., 2008). Finally, although similar to *S. cerevisiae*, the Schizosaccharomycetes reproduce exclusively via medial fission, hence the nickname 'fission yeast' (Hoffman et al., 2015).

The Biology and Ecology of *Saccharomyces cerevisiae*

The ability to maintain and propagate monoclonal strains of *Saccharomyces cerevisiae* have allowed scientists to study the genetics of what is effectively a few single individuals and, thus, rigorously characterize and document a few well-studied strains (Greig & Leu, 2009). Although select yeast strains are commonly employed in a variety of commercial processes such as baking/brewing and as a model organism in the laboratory, it remains unclear whether *S. cerevisiae* truly exists as a naturally occurring 'wild' species (Duina et al., 2014; Greig & Leu, 2009). Perhaps the distribution of *S. cerevisiae* in nature is, instead, largely artificial and simply the result of domestication.

In these tightly controlled and unnatural domesticated contexts, yeasts have very little reason, if any, to adapt and, as such, very little is known about how *S. cerevisiae* responds to both intra- and interspecific competition within these "natural" microbial communities (Boynton & Greig, 2014; Jouhten et al., 2016). This significant knowledge gap severely handicaps the work towards a contextually accurate understanding of the relationship between the ecology, evolution, and genetics of *S. cerevisiae*. Therefore, the ecology of *S. cerevisiae* is vastly different from other ascomycetes in that this ecology is likely the result of domestication and, thus, is largely man-made.

Genomic analyses have suggested that *S. cerevisiae* originated 'out-of-China' (Peter et al., 2018). This Chinese origin coincides with some of the earliest evidence of alcoholic drinks in recorded human history. *Saccharomyces paradoxus*, the closest known relative of *S. cerevisiae*, can be found as a commensal resident on the bark and leaves of deciduous trees throughout the northern hemisphere (Boynton & Greig, 2014; Greig & Leu, 2009). However, the bark and leaves of these deciduous trees are notably low-sugar

substrates. In contrast, *S. cerevisiae* most commonly grows on distinctly high-sugar substrates (e.g. the skin of grapes) and is well known for its efficiency in fermenting sugars to ethanol (Duina et al., 2014). Fortunately, *S. paradoxus* remains to be marred by the pressures of domestication and, thus, serves as a potential surrogate for studying and understanding the ecology, evolution, and genetics of these hypothetical 'wild' yeast populations (Dunham & Louis, 2011; Greig & Leu, 2009).

In general, *Saccharomyces* species share similar biochemical phenotypes and morphologies (Boynton & Greig, 2014). Although *S. cerevisiae* and *S. paradoxus* are both mesophilic, the ideal temperature for *S. paradoxus* is slightly lower than that of *S. cerevisiae*. Furthermore, the biogeographical distribution of four distinct *S. paradoxus* strains throughout the northern hemisphere suggest that these yeasts can adapt to distinct climatic conditions and, thus, likely contain temperature-dependent genetic components (Leducq et al., 2014).

Biology of the Yeast Cell Wall

The cell wall of the yeast *Saccharomyces cerevisiae* is a dynamic extracellular structural matrix that consists of several polysaccharide-based polymers including β-glucans (polymers of uridine diphosphate-*N*-glucose, UDP-Glc), chitin (a polymer of uridine diphosphate-*N*-acetylglucosamine, UDP-GlcNAc), and mannoproteins (proteins glycosylated with mannose) (Lesage & Bussey, 2006). The ability to regulate the biosynthesis and remodeling of the cell wall plays a critical role in maintaining cellular integrity during a variety of cellular events (i.e. division, growth, and mating) and in response to both environmental and metabolic stress conditions (Popolo et al., 2001). Thus, the biosynthesis and the remodeling of the cell wall are tightly controlled processes and critical to the survival of the yeast.

In general, the cell wall in fungi and plants creates a strong, polysaccharide-protein-based framework by which individual cells can act as structural *bricks* within some larger framework or body plan (Somerville et al., 2004). However, this role is irrelevant in unicellular fungi such as the yeast *S. cerevisiae*. In addition to these roles in maintaining cellular shape and protecting against stressful conditions, the cell wall is also involved in regulating the influx and efflux of water by osmosis by allowing cells to elastically expand or contract in response to changes in extracellular osmolarity (Klis et al., 2006). Although there is much inter-kingdom diversity in both cell wall composition and structure, the underlying secretory mechanisms and pathways involved in eukaryotic cell wall biosynthesis are, for the most part, highly conserved (Lesage & Bussey, 2006). In the yeast *S. cerevisiae*, the cell wall consists of approximately 85% polysaccharides and 15% proteins by mass (Nguyen et al., 1998). Furthermore, the cell wall can account for

approximately one-third of the dry mass of the yeast S. *cerevisiae* (Aguilar-Uscanga & Francois, 2003).

The yeast cell wall primarily consists of an inner chitin/β-glucan matrix layer and an outer layer of mannoproteins. The space between the yeast cell wall and plasma membrane is traditionally referred to as the *periplasmic space*. Two distinct varieties of β-glucans are found in yeast. These include β-1,3-glucans, which form the primary polysaccharide backbone, and shorter β-1,6-glucans, which form elongated, inter-backbone branches (Lesage & Bussey, 2006). This branching network of 1,3 and 1,6 β-glucans serves to both cross-link chitin polymers via β-1,4 glycosidic linkages and provide an overall spring-like resistance to the cell wall (Klis et al., 2006). Furthermore, the relative amounts of β-1,3 and β-1,6-glucans vary depending on the morphological state of the yeast (Magnelli et al., 2002). β-1,3-glucans are synthesized and translocated to the periplasmic space by β-1,3-glucan synthase, a transmembrane enzyme. This enzyme synthesizes β-1,3-glucan from UDP-Glc, a cytoplasmic precursor, and extrudes newly synthesized linear β-1,3-glucan strands into the periplasmic space where they are incorporated into the cell wall. β-1,3-glucan synthase consists of two catalytic subunits (encoded by FKS1 and GSC2) and one regulatory subunit (encoded by RHO1). Rho1 integrates a variety of signals from sensors in both the cell wall and the intracellular actin cytoskeleton (Lesage & Bussey, 2006). However, much work remains in fully characterizing and understanding the central role of Rho1 in modulating β-1,3-glucan synthase activity.

Chitin plays an important role in the maintenance of the yeast cell wall. In contrast to the more elastic β-glucan matrix, chitin confers the bulk of the walls rigidity and its ability to withstand both chemical and physical abuse (Fesel & Zuccaro, 2016). This

resiliency stems in part from the insolubility of chitin in water. Chitin is predominantly found close to the plasma membrane. Like β-1,3-glucan synthase, two transmembrane chitin synthases (Chs2 and Chs3) synthesize chitin by modifying and adding cytoplasmic UDP-GlcNAc precursors to the growing chain within the periplasmic space (Lesage & Bussey, 2006). These chitin synthases are activated at different times. Chs2 predominates during cytokinesis and Chs3 predominates during initial bud formation.

The fungal cell surface and cell wall play critical roles in budding, mating, and cell fusion (Lussier et al., 1997). As such, the cell surface and cell wall vary immensely in composition, shape, and structure throughout the fungal life cycle.

Calcofluor White: An Inhibitor of the Cell Wall

Calcofluor white (fluorescent brightener 28) is a fluorescent stain that associates strongly with chitin and cellulose (both structural polymers comprised of polysaccharides) and is commonly used in the detection of fungal, algal, and plant cell walls (Harrington & Raper, 1968; Herth & Schnepf, 1980; Nagata & Takebe, 1970). Calcofluor white has also been used to identify chitin in chitin-containing organisms such as arthropods and certain sponges (Ehrlich et al., 2013).

The active fluorochrome is the disodium salt of a highly conjugated stilbene derivative (Harrington & Raper, 1968) (*see* ***Figure 3***). Using ultraviolet and electron microscopy, Herth and Schnepf (1980) demonstrated that calcofluor white tightly associates via hydrogen bonding parallel to the longitudinal axis of chitin and cellulose. However, this strong association between calcofluor white and the longitudinal axis of chitin and cellulose can impair the nucleation and elongation of these polymers.

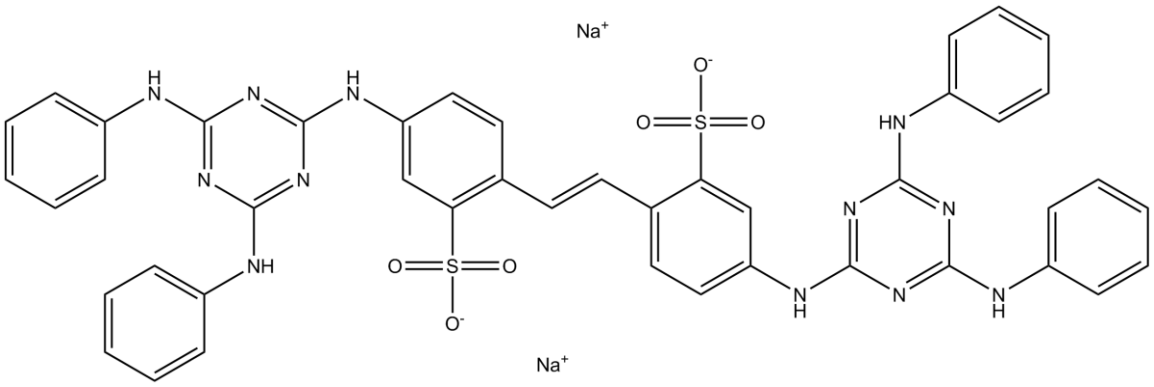

Figure 3. The molecular structure of calcofluor white.

The active fluorochrome of calcofluor white fluoresces when it associates with β-1,4 and β-1,3 glycosidic linkages found in cellulose and chitin, the primary structural component of plant and fungal cell walls, respectively. MW=960.95

Calcofluor white has been shown to negatively affect endogenous chitin polymerization and remodeling of the fungal cell wall. Roncero and Duran (1985) noted that a 2% calcofluor white solution dramatically inhibits the growth of wild-type *S. cerevisiae*. In addition to complete growth inhibition at high concentrations, Roncero and Duran (1985) observed varying responses in *S. cerevisiae* to differing calcofluor white concentrations. For instance, certain yeast strains were not appreciably affected by calcofluor white concentrations below 15 μg/ml. However, at intermediate concentrations with other strains, Roncero and Duran (1985) observed budding yeast failing to separate by producing distinctly abnormal multicellular aggregates and, thus, resulting in a dramatic decreased cell growth rate.

The Diversity of Peptidases and Proteolysis

Proteases (peptidases) are enzymes that catalyze the post-translational (and usually irreversible) hydrolysis of peptide bonds covalently joining adjacent amino acids in substrate polypeptide chains. In general, these enzymes are classified according to the underlying mechanism of peptide bond hydrolysis. The MEROPS database represents the de facto standard in the field for classifying these diverse enzymes, their substrates and their inhibitors (Rawlings et al., 2018). For example, aspartic (A-), glutamic (G-), and metallo (M-) type peptidases typically employ a one-step catalysis; and cysteine (C-), serine (S-), and threonine (T-) peptidases typically employ a two-step catalysis.

In one-step catalysis, an acidic amino acid (i.e. aspartate or glutamate) or a divalent cation (i.e. zinc) first polarizes a water molecule which then, in turn, directly catalyzes the hydrolysis of some targeted peptide bond via a classical nucleophilic attack and, thus, liberates two cleavage products. However, a two-step protease is first transiently bound to the targeted peptide covalently which, in turn, destabilizes the original peptide bond and displaces one cleavage product. However, the second cleavage product remains covalently linked to the protease until a water molecule can eventually perform a classical nucleophilic attack on this transient protease-substrate bond and, thus, liberate the second cleavage product.

Carboxypeptidases are proteases that more specifically catalyze the hydrolysis of carboxy-terminal (c-terminal) amino acids from substrate polypeptide chains (Gomis-Ruth, 2008). As a peptidase, carboxypeptidases too are classified according to the conventional MEROPS enzyme types (A-, C-, G-, M-, S-, T-, or U-) based on their structure and underlying catalytic mechanism (Rawlings et al., 2018).

The M14 Metallocarboxypeptidase Family

Metallocarboxypeptidases (MCPs) or M-type carboxypeptidases are of interest in this study. Like other proteases, the catalytic activity of MCPs has been implicated in a variety of post-translational processing pathways, such as the processing of pro-hormones and neuropeptides (Arolas et al., 2007). One notable family of MCPs is the human M14 family that contains 21 catalytically active members and four members predicted to be catalytically inactive (Pascual et al., 1989; Petrera et al., 2014; Reznik & Fricker, 2001; Tan et al., 1993) (*see* ***Figure 4***).

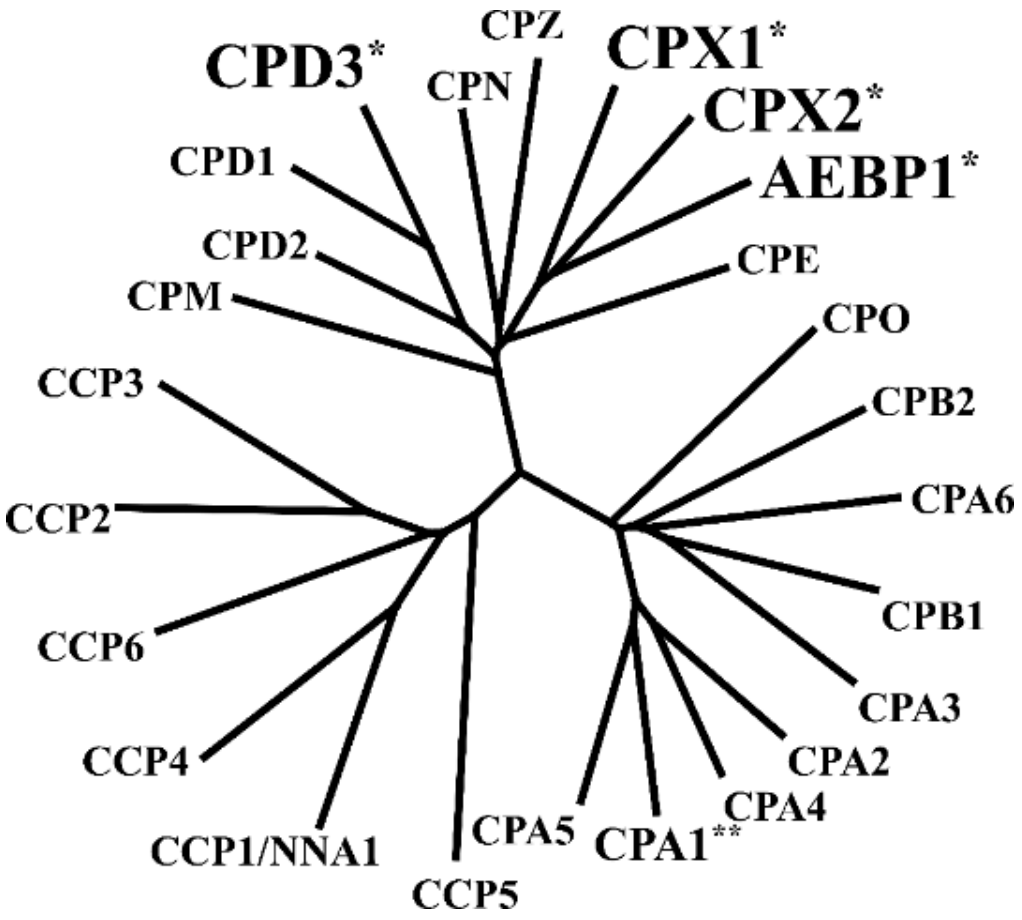

Figure 4. The human M14 metallocarboxypeptidase family.

The human M14 family contains 21 catalytically active members and four members predicted to be catalytically inactive (*****). The protein coded by ECM14 in *Saccharomyces cerevisiae* is homologous to human CPA1 (******).

Members of the human M14 family have been implicated in a variety of physiological functions and disease pathways. Three notable active members of the M14 family are carboxypeptidase B2 (CBP2/TAFI/), carboxypeptidase E (CPE), and cytosolic carboxypeptidase 1 (CCP1/Nna1/AGTPBP1). These three members briefly illustrate the functional diversity of this enzyme family.

1. **Carboxypeptidase B2 (CBP2/TAFI/CPU)** functions in blood plasma as an inhibitor in the fibrinolytic pathway (Eaton et al., 1991). Active plasmin alone is a very powerful hydrolyzer of fibrin, the fibrous structural protein involved in blood clot formation (Ezihe-Ejiofor & Hutchinson, 2013). CPB2 is synthesized by the liver as the zymogen proCPB2, a catalytically inactive precursor. Active CPB2 inhibits the fibrinolysis pathway by preferentially cleaving c-terminal lysine residues from fibrin, making fibrin less favorable for degradation by activated plasmin. Because the effect of active CPB2 is complemented with fibrin cross-linking and stabilization by coagulation factor XIII, a CPB2-deficiency alone does not lead to grossly impaired hemostasis (Nagashima et al., 2002). However, deficiencies in both CPB2 and coagulation factor XIII result in premature clot lysis and overall hemostasis dysregulation.
2. **Carboxypeptidase E (CPE/CPH)** plays a ubiquitous role in the biosynthesis of nearly all neuropeptides and peptide hormones within the human body (Fricker, 2018). These peptides include insulin, vasopressin, oxytocin, and the enkephalins. Mice homozygous for the fat spontaneous mutation (Cpe^{fat}/Cpe^{fat}) exhibit a variety of endocrine disorders including hyperglycemia, infertility, insulin resistance, and obesity (Naggert et al., 1995). Although CPE mutations

within the human population are not common, a truncating mutation has been documented in a morbidly obese female with type II diabetes mellitus (Alsters et al., 2015).

3. **Cytosolic carboxypeptidase 1 (CCP1/Nna1/AGTPBP1)** has been implicated in the deglutamylation (a post-translational modification) of tubulin during a variety of cellular events including axonal regeneration, axogenesis, dendritogenesis, and the synthesis and maintenance of the ciliary/flagellar axoneme (Kalinina et al., 2007; Otazo et al., 2013). Defective tubulin deglutamylation is responsible for neurodegeneration in the classical Purkinje cell degeneration (pcd) mouse, which lacks functional CCP1. These mice exhibit adult-onset degeneration of cerebellar Purkinje neurons, olfactory bulb mitral neurons, retinal photoreceptors, and select thalamic neurons.

Catalytically active MCPs of the M14 family characteristically contain six conserved active site residues: H/His69 (histidine), E/Glu72 (glutamate) R/Arg127 (arginine), R/Arg145, H/His196, and E/Glu270 (glutamate) (Auld, 2004; Lyons & Fricker, 2011). [These numbers represent the positions of each residue in the mature polypeptide sequence of the prototypical carboxypeptidase, CPA1, from *Bos taurus*.]

H69, E72, and H196 play roles in coordinating with a required zinc ion; R127 and R145 are involved in the binding of c-terminal carboxylate groups; and E270 functions as a catalytic general base (*see* ***Figure 5***). Substitutions at any of these critical positions likely results in a loss of traditional carboxypeptidase-like activity.

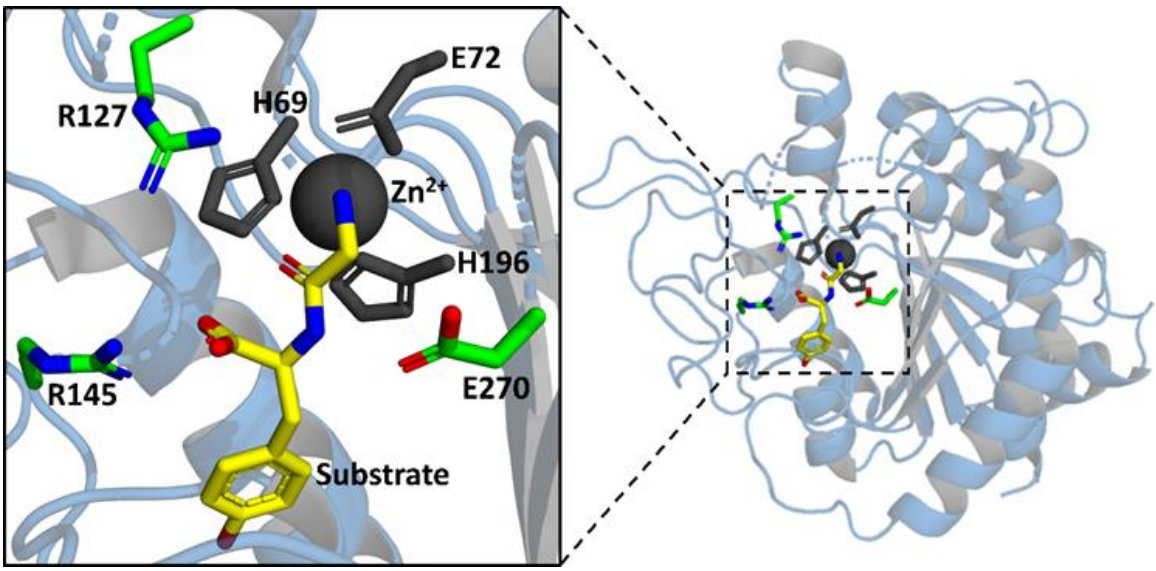

Figure 5. The active site of the prototypical carboxypeptidase from *Bos taurus*.

The active site of an metallocarboxypeptidase must contain two arginines (**R127** and **R145**) to bind the substrate c-terminus (shown in yellow), a glutamate (**E270**) to perform acid-base hydrolysis, and three other residues (**H69**, **E72**, and **H196**) for coordination with a zinc ion (gray sphere). The dipeptide glycyl-L-tyrosine substrate is shown in the active site (shown in yellow). Only the carboxypeptidase domain is shown. Model generated using PyMOL (PDB ID 3CPA).

ECM14: A Fungal Pseudopeptidase

The yeast *Saccharomyces cerevisiae* genome contains ≈6,000 genes (≈5,000 non-essential and ≈1,000 essential) (Winzeler et al., 1999). [Non-essential genes are defined as being those that can be singularly disrupted without causing significant harm to an organism.] Of these genes, only one gene, *ECM14*, codes for a member of the M14 family of MCPs. The protein coded by *ECM14* is homologous to human CPA1 (one of the 21 catalytically active members of the M14 MCP family) (*see above* ***Figure 5***). Although many amino acid residues are conserved between fungal Ecm14 and human CPA1, multiple substitutions in the presumed active site of Ecm14 seem to indicate a loss of traditional carboxypeptidase-like activity (Idowu et al., 2016) (*see* ***Figure 6***, ***Figure 7***).

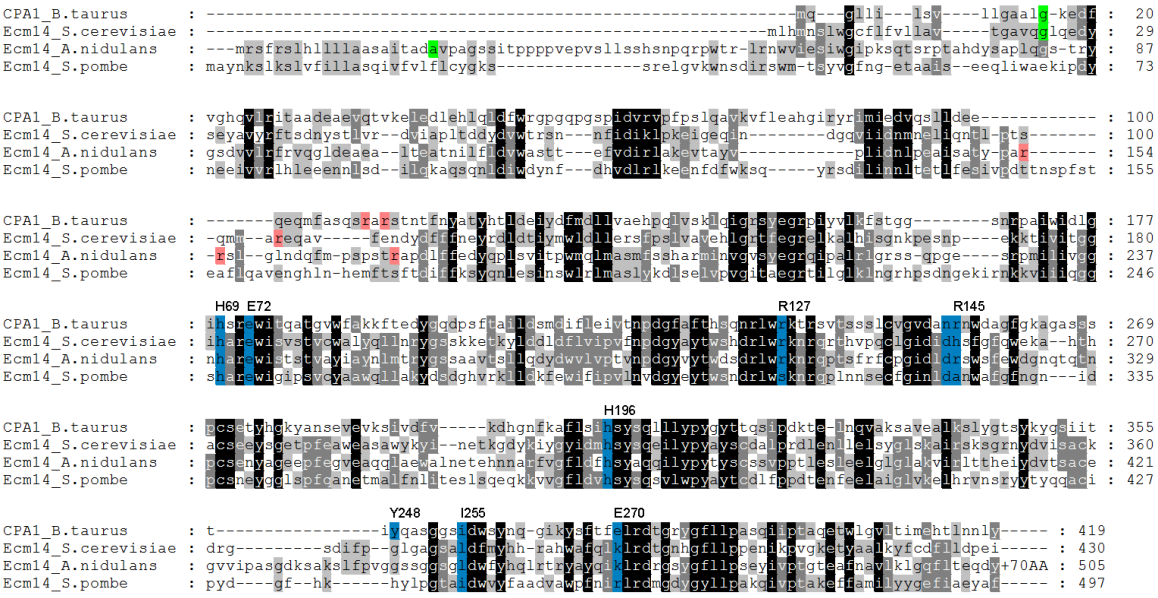

Figure 6. Comparison of Ecm14 with bovine CPA1 indicates that Ecm14 is likely an inactive member of the M14 carboxypeptidase family.

Ecm14 sequences from each of the three subphyla of Ascomycota are shown compared to *B. taurus* CPA1. The predicted cleavage sites for signal peptidase are shown in green. Possible sites at which a trypsin-like endopeptidase might cut to remove the prodomain are shown in pink. Critical active site amino acids are shown in blue, with the corresponding numbering from bovine CPA1 above.

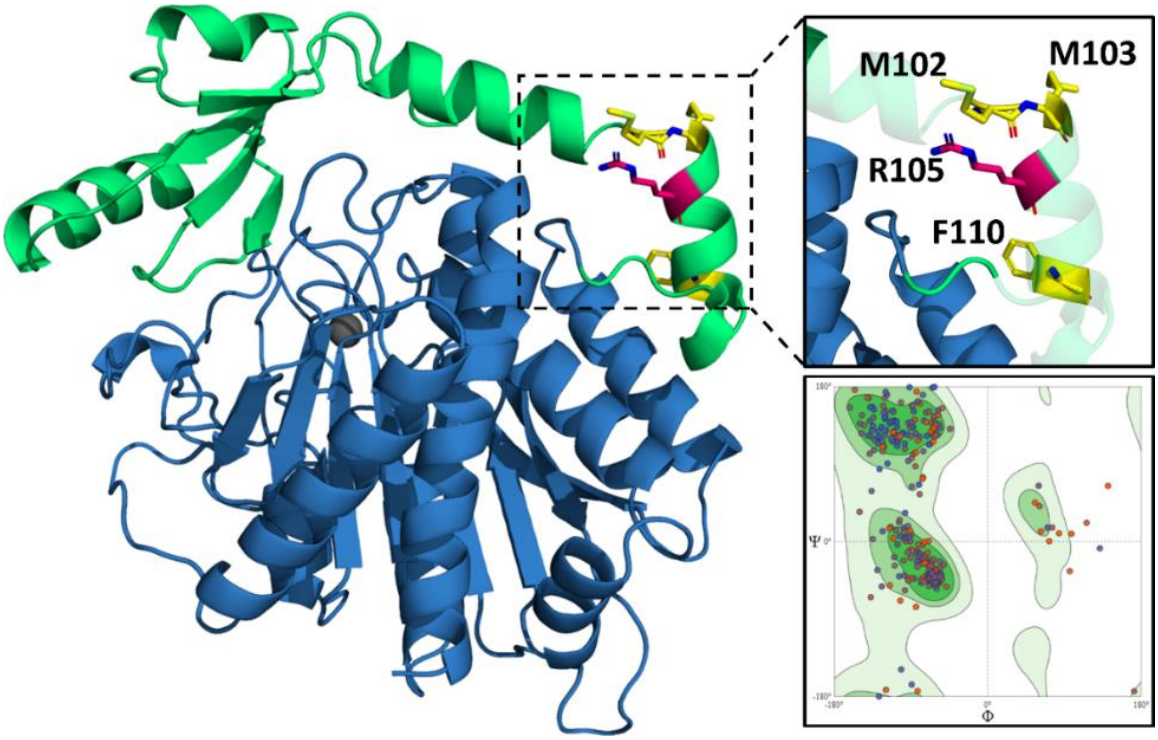

Figure 7. Ecm14 is similar to other metallocarboxypeptidases.

The tertiary structure of Ecm14 was modeled using SWISS-MODEL and an X-ray crystal structure for human CPA1 (PDB ID 4UF4). The catalytic domain is shown in blue and the prodomain in green. Endopeptidase cleavage sites for trypsin are shown in pink and for chymotrypsin in yellow. Models were generated using PyMOL. The resulting Ramachandran plot is also shown (bottom right).

At least within the context of routine laboratory growth conditions, *ECM14* is a non-essential gene in the yeast *S. cerevisiae* genome (Winzeler et al., 1999). However, the conservation of *ECM14* in the genome of ascomycete fungi, but not in other fungal phyla, suggests that Ecm14 does have some, albeit presently unknown, function (*see* **Figure 8**).

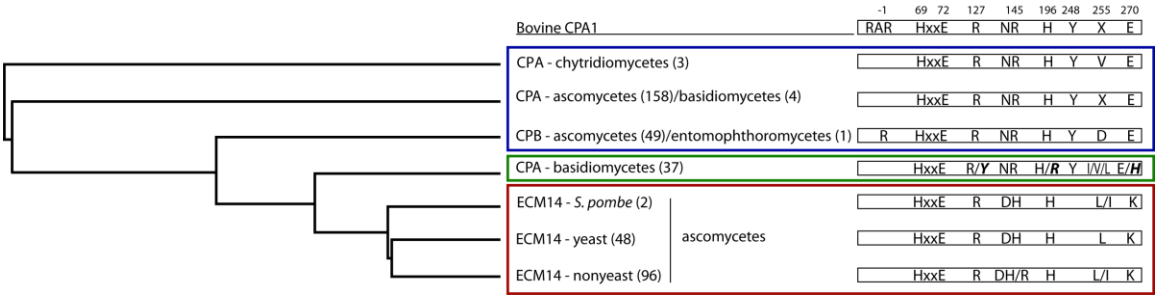

Figure 8. Ecm14 is conserved in Ascomycota, but not in other fungal phyla.

CPA1-like fungal proteins fall into three broad categories based on active site residues: standard CPA/CPB proteins containing all expected active site residues, with representatives from all subphyla of fungi (blue box), basidiomycete CPA proteins containing expected active site residues with the exception of 5 closely aligned proteins from 5 different species containing R127Y, H196R, and E270H substitutions (green box), and Ecm14-like proteins found only in ascomycetes and defined by the presence of N144D and E270K substitutions.

Prodomain cleavage is an essential feature of many catalytically active peptidases. For example, the model zinc protease, CPA1, is synthesized within the rough endoplasmic reticulum of pancreatic centroacinar cells as the zymogen proCPA1, a catalytically inactive precursor. This zymogen is secreted in response to stimulation of the pancreas by cholecystokinin (CCK), a digestive hormone. Once in the duodenum, proteolytic cleavage of the N-terminal prodomain (a 95-aa regulatory domain) liberates the catalytically active carboxypeptidase domain (*see above* ***Figure 7***). Because cleavage of a prodomain is typically associated with activation in peptidases, retention of prodomain cleavage in Ecm14 likely suggests some underlying function to be regulated.

Previous work by Idowu et al. (2016) has demonstrated that both trypsin and chymotrypsin are able to cleave the prodomain of Ecm14 *in vitro*. Edman N-terminal degradation sequencing was performed by The Synthesis & Sequencing Facility at Johns Hopkins University (Baltimore, M.D.) in order to identify the location of these endopeptidase cleavage sites (*see above* ***Figure 7***). Furthermore, previous work by Schott (2015) suggests that Ecm14 might be able to compete with active carboxypeptidases *in vitro* by binding and protecting c-terminal amino acid residues of potential substrates.

Although this previous work has preliminarily suggested some activity of Ecm14 *in vitro*, very little is known about the *in vivo* implications of the key amino acid substitutions in the presumed active site of Ecm14 (and pseudoenzymes in general). However, the present evidence suggests that Ecm14 is likely functional; not as a catalytically active enzyme, but rather as a binding protein that transiently interacts with other protein substrates and, thus, influences the intracellular localization and/or fate of these protein substrates.

The obscure functional nature of pseudoenzymes in general makes them inherently difficult to associate with a phenotype. Although previous studies have suggested several phenotypes for *ECM14*, the potential of these phenotypes for both reproducibility and quantification remains to be fully understood. However, a reproducible and readily quantifiable phenotype is fundamentally needed in order to systematically study the biochemical and biological implications of the key amino acid substitutions of pseudoenzymes in general.

Potential Phenotypes of *ECM14*

In a genome-wide screen focused on characterizing yeast genes involved in cell wall architecture and remodeling, Lussier et al. (1997) identified Ecm14 as a possible player in this broad functional classification. Understanding the impact of calcofluor white on chitin polymerization and cell wall biosynthesis, Lussier et al. (1997) began by initially screening for any altered sensitivity to calcofluor white in transposon-mutagenized yeast cells. Ram et al. (1994) first postulated that exposure to calcofluor white amplifies the effect of mutations in genes specifically involved in cell wall architecture and remodeling. Additional supplemental screens were used to further characterize the impact of each mutated gene on a variety of phenotypes known to be related to cell wall biology. In addition to ECM14, at least 81 other genes in the yeast *S. cerevisiae* genome were identified, that when mutated, confer at least some modified sensitivity to calcofluor white (Lussier et al., 1997). An altered cell wall hexose composition (an increased mannose: glucose ratio) and a hypersensitivity to the antifungal compounds Hygromycin B and Papulacandin B were also observed in ecm14 mutants.

In a genome-wide screen, Hillenmeyer et al. (2008) systematically subjected every non-essential gene knockout yeast strain to a variety of different chemicals and identified several potential fitness defects associated with ecm14 deletion. These fitness defects included a reduced sensitivity to paraquat, mercury chloride, and lithium chloride. However, Idowu et al. (2016) were unable to clearly replicate these results. It is theorized that a disturbed or weakened cell wall is simply not able to withstand drug or chemical concentrations as effectively as the more robust cell wall of wild-type yeast cells.

Therefore, a disturbed or weakened cell wall might also explain the fitness defects observed by Hillenmeyer et al. (2008) and Lussier et al. (1997).

Finally, Schoner et al. (2008) identified arp1 ecm14 double mutants as potentially being inviable—a case of synthetic lethality. [Actin-related protein 1 (Arp1) serves as a filamentous core backbone in the dynactin complex, which functions in nuclear migration during mitosis (Igarashi et al., 2005). Additionally, Arp1 plays an essential role in ensuring the completion of cell wall remodeling prior to mitotic initiation.] However, like the fitness defects observed by Hillenmeyer et al. (2008) and Lussier et al. (1997), attempts to replicate these results have remained either incomplete or unsuccessful. [The concept of synthetic lethality is discussed in more detail in a subsequent introductory sub-section below.]

Intracellular Localization and Potential Biological Role of Ecm14

In a cellular analysis of protein localization in *S. cerevisiae*, Huh et al. (2003) identified Ecm14 to be localized within the yeast vacuole, an important organelle involved in yeast cell wall architecture and remodeling. By inserting the coding sequence for green fluorescent protein (GFP) from *Aequorea victoria* (a bioluminescent jellyfish species) in-frame immediately preceding the endogenous stop codon for each yeast ORF (open reading frame), Huh et al. (2003) were able to provide a comprehensive map of *in vivo* protein localization in the yeast *S. cerevisiae*.

In yeast, the vacuole localizes the key enzymatic players of the macroautophagy pathway. [The yeast vacuole is functionally analogous to the lysosomes found in higher eukaryotes.] This ubiquitous homeostatic process facilitates the degradation of damaged organelles and misfolded proteins in response to a variety of metabolic signals and stressors (Levine et al., 2011). Furthermore, the vacuole is responsible for the degradation of several enzymes specifically implicated in cell wall biology (Wiederhold et al., 2009). For example, one recognized pathway for regulating chitin synthase II (Chs2) is via vacuolar degradation by the vacuolar aspartyl protease A (Pep4) (Chuang & Schekman, 1996). In response to some yet to be characterized signal, Chs2 is trafficked via endocytosis into the yeast vacuole by Sla2 for eventual degradation (Lesage & Bussey, 2006).

Like Ecm14, Proteinase C (Prc1, a serine carboxypeptidase Y) and Autophagy-42 (Atg42, a functional homologue of Prc1) are both localized in the yeast vacuole (Huh et al., 2003). Parzych et al. (2018) demonstrated that deficiencies in both Atg42 and Prc1 result in several defects in the yeast macroautophagy pathway. Following hydrolysis, the resultant components are released into the cytosol via newly synthesized permease-like

proteins (Yang et al., 2006). However, the present evidence suggests that yeast with mutations in both Atg42 and Prc1 have significant deficiencies in the breakdown of autophagic-bodies, the activation of vacuolar zymogens, and the synthesis of various autophagy-dependent proteins (Parzych et al., 2018).

Another vacuolar carboxypeptidase in *S. cerevisiae* is Carboxypeptidase S (Cps1). Bordallo and Suarez-Rendueles (1993) demonstrated that yeast pre-cultured in a nitrogen and glucose-rich growth medium show a dramatic increase in CPS1 expression when subsequently deprived of nitrogen. Like Atg42 and Prc1, the upregulation of Cps1 activity during periods of nutrient deprivation likely suggests a crucial mechanism of homeostasis. In the case of Cps1, this mechanism seems to invoke the recycling of certain peptides (as catalyzed by Cps1) for use as an alternative source of metabolic nitrogen during periods of nitrogen depravation (Bordallo & Suarez-Rendueles, 1993).

Although Ecm14 seems to lack traditional carboxypeptidase-like activity, its homology with active peptidases, its implication in resistance to calcofluor white, and its localization in the yeast vacuole seem to suggest a possible uncharacterized role in cell wall architecture and remodeling.

Synthetic Lethality and the Synthetic Lethal Screen

The phenomenon of synthetic lethality occurs when two or more genetic perturbations together result in cellular death (Dobzhansky, 1946; Nijman, 2011). Biochemical pathways typically involve multiple key genetic components and it is thought that many of these key gene players can provide some redundancy in the event that other genes experience deleterious mutations (*see* ***Figure 9***). Although enzymes are generally characterized based on a primary native reaction, enzymes (and potentially their inactive homologues) can also participate in numerous secondary side reactions and processes. This theory of "genetic robustness" suggests that organisms are quite resistant to major phenotypic changes despite inherent genetic variability, environmental changes and random mutation events.

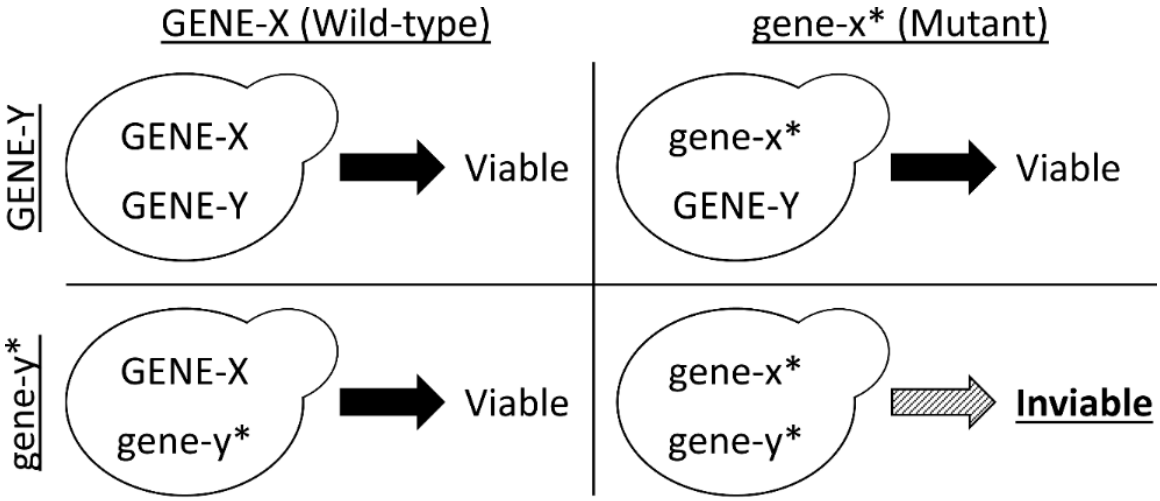

Figure 9. The phenomenon of synthetic lethality.

Biochemical pathways typically involve multiple key gene components and many of these key gene players provide redundancy if other genes experience deleterious mutations (*). The synthetic lethal screen identifies novel mutants (gene-y) whose survival depends on a gene of interest (GENE-X).

One method for exploring the function of non-essential genes is the synthetic lethal screen. This screen incorporates a convenient color-colony reporting assay to quickly visualize dependence on a plasmid carrying a gene of interest required for cell survival (Barbour et al., 2000; Koshland et al., 1985). The color-colony reporting assay exploits the *de novo* adenine biosynthesis pathway involving at least seven major genetic players (named *ADE1-7*). In short, cells carrying an *ade2* mutation accumulate the intermediate P-ribosylaminoimidazole, which is oxidized to a red pigment. However, the *ade3* mutation is epistatic to *ade2* and, thus, blocks the pathway at a point prior to pigment formation. Thus, both wild-type and *ade3*/*ade2* double mutant yeasts are white, and ade2 single mutant yeast are red (*see* ***Figure 10***). However, double *ade3*/*ade2* mutant yeasts transformed with a plasmid carrying a functional *ADE3* gene are red (*see* ***Figure 10***). Transformants that only weakly retain the plasmid will produce red and white sectored colonies to varying degrees.

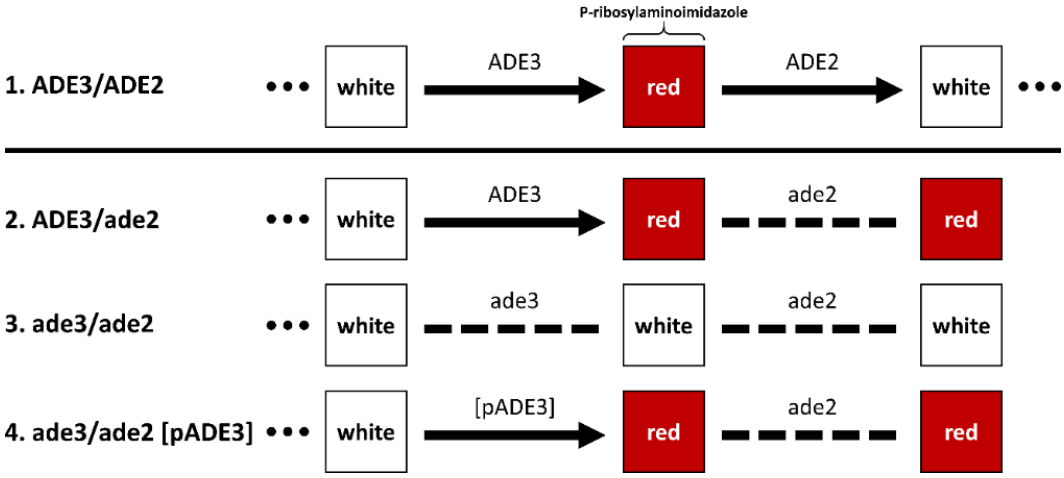

Figure 10. The genetic basis for the red-white color-colony assay in yeast.

(**1**) Wild-type yeast are white. (**2**) Yeast that contain a single *ade2* mutation accumulate a red pigment (oxidized P-ribosylaminoimidazole). (**3**) Yeast that contain both *ade3* and *ade2* mutations result in white colonies. (**4**) These double mutants that are transformed with a plasmid containing a functional *ADE3* gene [pADE3] are red.

A synthetic lethal screen begins by first deleting a gene of interest in a yeast strain that is already deficient in the adenine biosynthesis pathway (*ade3*/*ade2* mutants that produce white colonies). [The yeast gene deletion strategy is discussed in more detail in a subsequent methods sub-section.] Additionally, this host yeast strain must also contain an auxotrophic marker for selecting transformants (e.g. *ura3*Δ is commonly used) and complementary auxotrophic or drug resistance markers for selecting diploids from mating reactions (Sadowski et al., 2008). In *URA3*-based selection, the plasmid DNA has a wild-type copy of the yeast *URA3* gene, as well as the *URA3* promoter; however, the host yeast strain lacks the *URA3* allele. Thus, yeast cells transformed with the plasmid have a functional copy of the *URA3* gene, which allows them to grow on media lacking uracil. Conversely, *URA3*-based selection also enables negative selection against plasmids containing *URA3* when grown on media containing 5-fluoroorotic acid (5-FOA). When 5-FOA is added to the growth media, the active *URA3* gene product (orotidine 5'-phosphate decarboxylase) converts 5-FOA into the toxic compound 5-fluorouracil (5-FU). These URA3-containing plasmids are effectively destabilized and rapidly lost when grown on media containing 5-FOA. The *S. Cerevisiae* W303 strain (MATa/MATα {*leu2*-3,112 *trp1*-1 *can1*-100 *ura3*-1 *ade2*-1 *his3*-11,15}) is commonly used as the parent strain when performing synthetic lethal assays (Cherry et al., 2012).

After creating a suitable knockout W303-derived yeast strain, a wild-type copy of the gene of interest is cloned into a plasmid specifically designed for yeast synthetic lethal screens (Barbour et al., 2000). As a shuttle vector, this pSLS1 plasmid contains sequences for proliferation and selection in both bacteria and yeast. As a vector for synthetic lethal

screens, this pSLS1 plasmid also contains a functional *ADE3* gene and a yeast centromere (*CEN4*) under the influence of a galactose promotor (pGAL1) (*see* ***Table 2*** *and* ***Figure 11***).

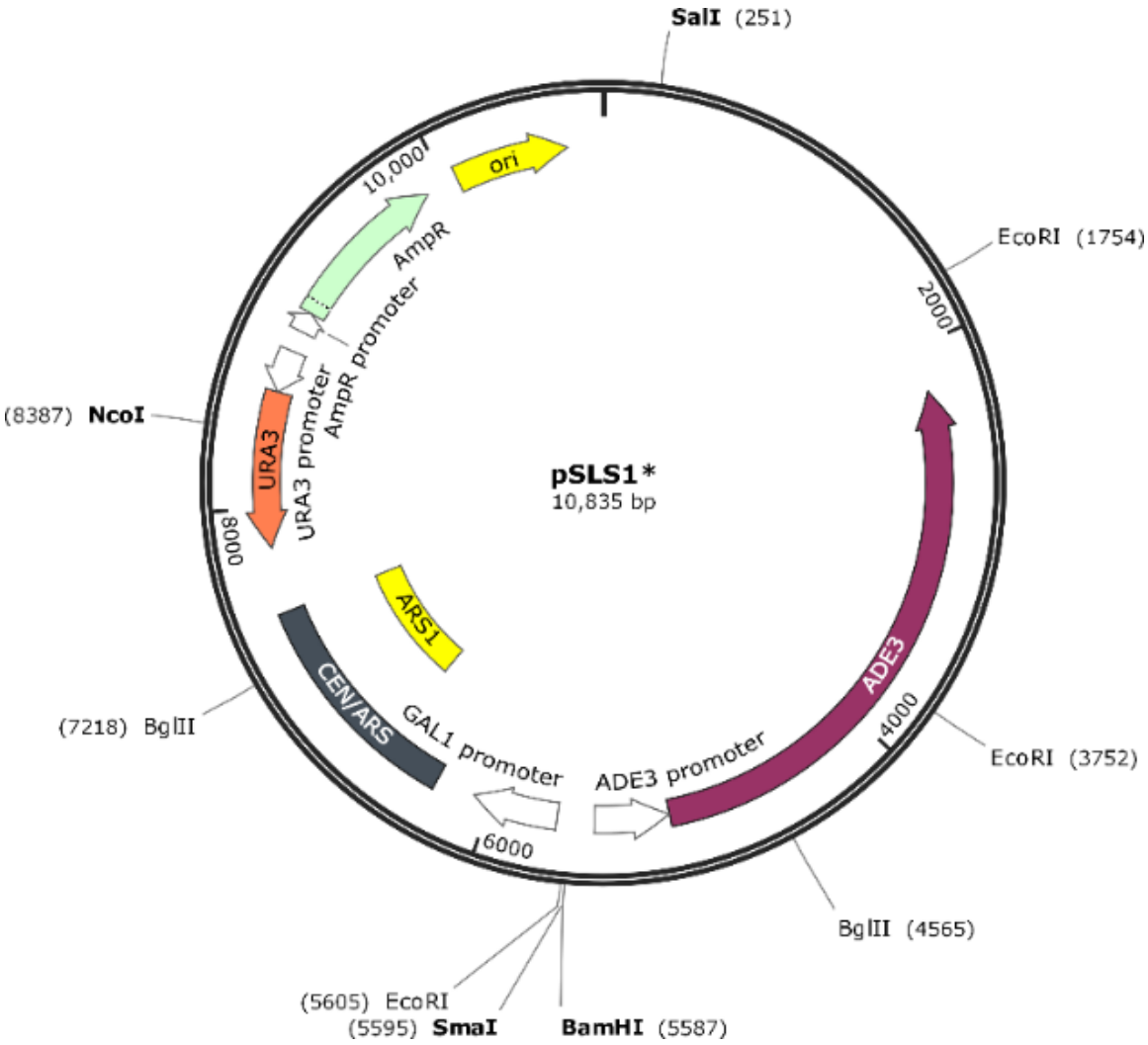

Figure 11. The pSLS1 plasmid for yeast synthetic lethal screens.

The pSLS1 plasmid contains the bacterial origin of replication sequence (ORI), the ampicillin resistance gene (AmpR) for selection in bacteria, the CEN/ARS sequence for replication in yeast, the *URA3* gene for selection in yeast, and the *ADE3* gene for color-colony reporting.

When grown on media containing galactose, the galactose promotor erroneously induces transcription through the traditionally *un-transcribed CEN4* sequence (*see above* ***Figure 11***). The *CEN4* sequence normally causes these yeast centromere plasmids (YCp) to behave like low-copy, "mini-chromosomes" during cell division (Clarke & Carbon, 1980). However, erroneous transcription through the centromere sequence causes these YCp plasmids to instead behave like highly unstable yeast replicating plasmids (YRp). Therefore, providing galactose as a carbon source functions as a *molecular switch* causing the pSLS1 centromere plasmid to instead behave more like a replicating plasmid (Barbour et al., 2000). This ability to regulate plasmid stability *in vivo* is an essential requirement for effective synthetic lethal screens and has been shown to significantly reduce the number of false positives that must be subsequently screened and interrogated (Barbour et al., 2000).

The synthetic lethal screen functions on the basis that dependence of the gene of interest for survival will cause yeast to retain the plasmid and, thus, produce non-sectored red colonies when grown on galactose-containing media (Barbour et al., 2000). However, yeast that are not dependent on the plasmid for survival will rapidly lose the plasmid when grown on galactose-containing media and produce red and white sectored colonies, or even solid white colonies, depending on the rate of plasmid retention. It is thought these sectored colonies are caused by the natural differential retention of the plasmid under various non-essential contexts and should ideally be avoided when selecting candidate yeast colonies for further investigation.

Spontaneous mutations occur at a rate far too low to be practically useful for most genetic screens. However, carefully prescribed treatment with either ethyl methansulfonate (EMS) or ultraviolet (UV) light can significantly increase the rate of mutations per gene

without the excessive killing of cells (Winston, 2008). As a potent base-altering agent, EMS is known to add an ethyl group to guanine, forming O^6-ethylguanine (*see* ***Figure 12***). This alkylated guanine can no longer form the typical three hydrogen bonds with cytosine and, thus, can erroneously be mispaired with thymine by DNA polymerase during DNA replication, causing G/C→A/T transitions (Brookes & Lawley, 1961). Since stop codons are rich in ATs (i.e. TAA, TGA, and TAG), EMS-based mutagenesis is effective for generating nonsense mutations. Because truncated gene products are generally not useful or even partially functional, these types of nonsense mutations are preferred when conducting synthetic lethal screens.

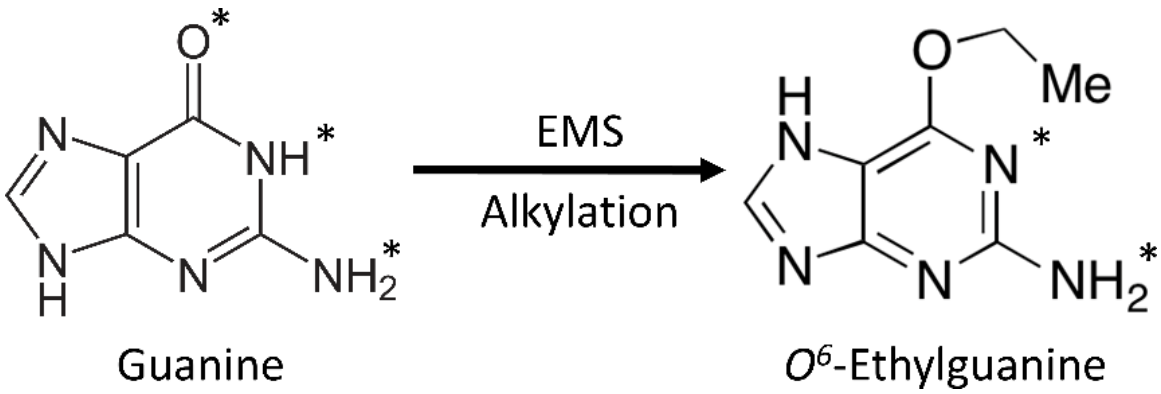

Figure 12. EMS-induced alkylation of guanine to form O6-ethylguanine.

Guanine can form three hydrogen bonds (*****). However, O^6-ethylguanine can only form two hydrogen bonds (*****) and, thus, can erroneously mispair with thymine during DNA replication.

The synthetic lethal screen can be described by three primary steps. First, a W303-derived knockout yeast strain is transformed with a pSLS1-based plasmid containing a wild-type copy of the gene of interest (GENE-X). Next, random mutagenesis introduces random point mutations throughout the yeast genome. Finally, these mutagenized yeasts are cultured and screened on galactose-rich media to induce plasmid destabilization and the accumulation of the red pigment (*see* ***Figure 13***).

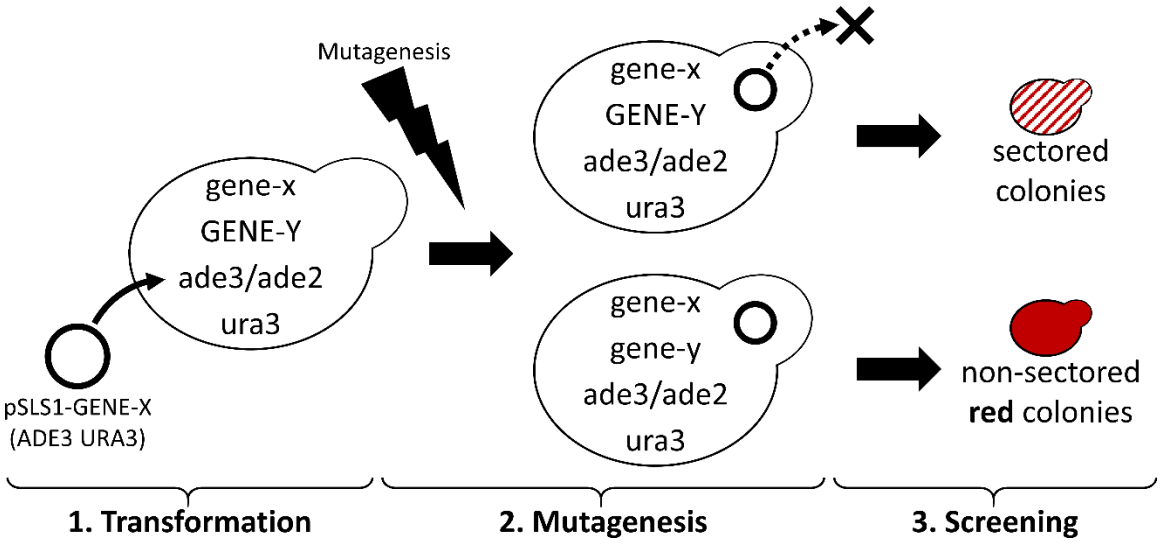

Figure 13. The synthetic lethal assay is used to identify novel mutant yeast strains whose survival is dependent on a gene of interest.

(**1**) A wild-type gene of interest (GENE-X) is first moved from the yeast genome into the pSLS1 vector (pSLS1-GENE-X). (**2**) Random genomic point mutations are introduced throughout the yeast genome using ethyl methansulfonate (EMS) or ultraviolet (UV) light. (**3**) Colonies that depend on the gene of interest for survival can be visualized by screening for non-sectored red colonies.

CHAPTER 2

GENERAL MATERIALS AND METHODS

Yeast Cell Culture and Media

YEPD media (Yeast Extract 10 g/l, Peptone 20 g/l, Dextrose 20 g/l) was used for the routine maintenance and propagation of yeast. When necessary, this routine YEPD growth media was also supplemented with 8 ml/l of 0.5% (w/v) adenine (YEPD+Ade) for yeast strains deficient in adenine biosynthesis. Filter-sterilized Geneticin® (G418 sulfate, 500 μg/ml) was used to select for yeast containing the KanMX allele. Filter-sterilized nourseothricin sulfate (NTC, 100 μg/ml) was used to select for yeast containing the NatMX allele. Both G418 and NTC were added to warm media (≈55°C). When preparing Petri dishes, agar was added at 20 g/l.

SC-Ura media (Synthetic Complete minus Uracil) was used to positively select for yeast transformed with a plasmid containing the URA3 allele. This SC-Ura media contained 1% (w/v) succinic acid, 0.6% (w/v) sodium hydroxide, 0.67% (w/v) yeast nitrogenous base, and 1% (w/v) casamino acids. A 0.5% (w/v) autoclaved adenine stock was added at 8 ml/l. A 1% (w/v) filter-sterilized tryptophan stock was added at 2 ml/l (small drips of 3 M sodium hydroxide were added, as necessary, to increase tryptophan solubility in water). A 20% (w/v) dextrose stock was autoclaved separately in order to prevent caramelization and undesirable Maillard side reactions with amino acids in solution. This 20% (w/v) dextrose stock was added at 100 ml/l. Adenine, tryptophan and dextrose were added to warm media (≈55°C). When preparing Petri dishes, agar was added at 20 g/l.

SC+5FOA media (Synthetic Complete plus 5-Fluoroorotic Acid) was used to negatively select for yeast transformed with a plasmid containing the URA3 allele. This

SC+5FOA media was made as above, with the addition of a 0.2% (w/v) autoclaved uracil stock that was added at 5 ml/l. Finally, 100X 5-FOA from Zymo Research (100 mg/ml in DMSO) was added to warm media at 10 ml/l.

YPGal media (Yeast Extract 10 g/l, Peptone 20 g/l, Galactose 20 g/l) was used for the destabilization of yeast centromere plasmids (YCp) under the influence of the galactose promoter (GAL1). A 20% (w/v) galactose stock was either autoclaved separately or filter-sterilized in order to prevent caramelization and undesirable Maillard side reactions with amino acids in solution. Galactose was added to warm media (≈55°C). When preparing Petri dishes, agar was added at 20 g/l.

Bacterial Cell Culture and Media

Luria broth (sodium chloride 10 g/l, tryptone 10 g/l, and yeast extract 5 g/l) was used for the routine maintenance and propagation of DH5α *Escherichia coli*. Filter-sterilized 1000X ampicillin stock (100 mg/ml) was added to warm media (≈55°C) at 1 ml/l (or µl/ml). When preparing plates, agar was added at 15 g/l. For blue-white screening plates, filter-sterilized 100X IPTG (100 mM) and filter-sterilized 100X X-Gal (20 mg/ml) were added to warm media (≈55°C) at 10 ml/l (or µl/ml).

Gel Extraction and Cleanup of Large DNA Fragments

The Qiagen QIAquick™ Gel Extraction Kit was used to cleanup DNA fragments from agarose gels following various enzymatic reactions. Although the QIAquick Gel Extraction Kit is only officially rated for up to 10 µg DNA (70-bp to 10-kb), several modifications to the included quick-start protocol allowed for the dependable extraction of larger DNA fragments (>10-kb). Bands of interest were sparingly visualized using UV transillumination, excised using clean scalpel blades (either fresh or cleaned using RNase AWAY™ as a DNA decontamination reagent), and transferred into tared 2-ml microcentrifuge tubes. Gel slices were excised and trimmed to <400 mg per extraction tube. Gel slices were incubated at 50°C for 10 minutes in three gel volumes of Buffer QG. Each incubating gel slice was vortexed every 2 minutes in order to thoroughly dissolve each gel slice. One gel volume of isopropanol was added to each tube and thoroughly vortexed. Each sample was applied to the center of a fresh QIAquick spin column, centrifuged for 90 seconds at 21,130g, and the flow-through discarded (this was repeated as needed for sample volumes >750 µl). An additional 500 µl of fresh Buffer QG was washed through each column in order to remove all traces of agarose. 750 µl of Buffer PE was carefully added to the center of each spin column and allowed to incubate at room temperature for >5 minutes. The Buffer PE wash was centrifuged for 90 seconds at 21,130g and the flow-through discarded. Each spin column was subsequently rotated 180 degrees and centrifuged for an additional 90 seconds at 21,130g to remove any residual Buffer PE. Residual Buffer PE (containing ethanol) can interfere with downstream applications. DNA was eluted by carefully applying 50 µl of warm Buffer EB (55-65°C) to the center of each spin column and allowing each to incubate at room temperature for >5 minutes. Each spin

column was placed into a clean 1.5-ml microcentrifuge tube and centrifuged for 90 seconds at 21,130g to elute the DNA. Multiple spin columns were used for samples exceeding 400 mg. Serial elutions were used to successively collect and concentrate DNA from multiple columns. Final preparations were stored at -20°C.

Isolation of Yeast Genomic DNA for PCR-based Applications

Yeast genomic DNA (gDNA) was isolated using an inexpensive and quick approach based on a technique by Looke et al. (2011) from the Institute of Molecular and Cell Biology, University of Tartu, Estonia. Either a single colony or 200 µl of mid-log phase liquid culture (1-5x10^7 cells/ml, OD_{600} = 0.4-1.7) was resuspended in 100 µl of fresh 200 mM lithium acetate, 1% SDS solution and incubated for 10 minutes at 70°C. 300 µl of 100% ethanol was added and samples were thoroughly vortexed prior to 3 minutes of centrifugation at 21,130g. Samples were carefully decanted and the supernatant discarded. Pellets were subsequently re-washed with 500 µl of 70% ethanol and allowed to air-dry for >10 minutes prior to being resuspended in 100 µl of Qiagen® Buffer EB (10 mM Tris-Cl, pH 8.5). Samples were briefly centrifuged for 15 seconds to collect cell debris and 1 µl of supernatant was used for PCR template DNA.

Isolation of Plasmid DNA From Bacterial Cells Using the NID Method

Plasmid DNA was routinely isolated using the "one-step, one-tube" non-ionic detergent (NID) method demonstrated by Lezin et al. (2011). The NID extraction buffer (5% sucrose, 20-50 mM EDTA, 50 mM Tris pH 8, 0.75 M NH_4Cl, and 0.5% Triton X-100) was stored without enzyme at 4°C. A separate 100X enzyme stock (10 mg/ml lysozyme, 2.5 mg/ml RNase A, 50 mM Tris pH 8, and 50% glycerol) was stored at -20°C and added as necessary. For small 50-µl minipreps, LB/Amp cultures (2-4 ml) were inoculated with fresh monoclonal bacterial colonies-of-interest and cultured overnight for 16-18 hours at 37°C with shaking at 250 RPM. These overnight cultures were split into 2-ml microcentrifuge tubes and pelleted at 5,000g for 1 minute. Each decanted pellet was thoroughly resuspended in 150 µl of extraction buffer (with enzymes) and incubated at 65°C for 5-7 minutes. For overnight cultures >2 ml, identical pellets were serially resuspended and collected in the same 150 µl of extraction buffer (with enzymes) prior to incubation for increased plasmid concentrations. Suspensions were centrifuged at 21,130g for 10 minutes and each cellular pellet was carefully removed using a sterile toothpick. Plasmid DNA was precipitated and pelleted by adding 120 µl of isopropanol, followed by thorough mixing and centrifugation at 5,000g for 10 minutes. Finally, 500 µl of 70% ethanol was added, followed by centrifugation at 21,130g for 10 minutes. After removing the ethanol supernatant, each pellet was allowed to thoroughly dry for 10-15 minutes prior to being dissolved in 50 µl of Qiagen Buffer EB.

Large 200-µl maxipreps were prepared by first inoculating 1-ml LB/Amp seed cultures with fresh monoclonal bacterial colonies-of-interest and incubating for 8 hours at 37°C with shaking at 250 RPM. These 8-hr seed cultures were then used to inoculate larger

100-ml LB/Amp cultures that were subsequently incubated overnight at 37°C with shaking at 250 RPM. Each 100-ml overnight culture was divided between three sterile 50-ml centrifuge tubes and pelleted at 5,000g for 5 minutes in a refrigerated centrifuge (4°C). Identical pellets were serially resuspended using 10 ml of prepared NID extraction buffer and consolidated into one 50-ml centrifuge tube. Suspensions were incubated in a water bath at 65°C for 8 minutes and centrifuged at 13,000g for 30 minutes in a refrigerated centrifuge (4°C). Plasmid DNA was precipitated and pelleted by adding 1 ml of isopropanol, followed by thorough mixing and centrifugation at 13,000g for 15 minutes in a refrigerated centrifuge (4°C). Each plasmid pellet was thoroughly resuspended in 1.5 ml of 70% ethanol, transferred to a clean 1.5-ml microcentrifuge tube, and centrifuged at 21,130g for 5 minutes at room temperature. After removing the ethanol supernatant, each pellet was allowed to thoroughly dry for 10-15 minutes prior to being dissolved in 500 µl of Qiagen Buffer EB.

DNA Quantification using ImageJ

DNA agarose gels were imaged using the AlphaImager HP™ gel imaging system and analyzed using NIH ImageJ (Rueden et al., 2017). These images were intentionally captured using a very low auto exposure profile to ensure more accurate final band quantification. Each gel image was opened in ImageJ and a bounding box was drawn around the entire DNA ladder using the Rectangle Tool. This first lane was marked by either pressing Control + 1 or via the Analyze > Gels > "Select First Lane" menu option. The box was then duplicated onto each lane of interest by dragging and marking these subsequent lanes by either pressing Control + 2 or via the Analyze > Gels > "Select Next Lane" menu option. The "Invert peaks" option was unselected (Analyze > Gels > "Gel Analyzer Options…"). Finally, lanes were plotted by either pressing Control + 3 or via the Analyze > Gels > "Plot Lanes" menu option (*see* ***Figure 14***).

Each peak of interest was area-bound by drawing a line using the straight-line tool underneath each peak between the predominate points of inflections above baseline. The magic wand tool was then used to select each enclosed, area-bound peak and each peak was labeled using the Analyze > Gels > "Label Peaks" menu option. Usually the bright 3.0-kb TriDye band was used as a reference standard for all DNA quantification experiments. When using a 5 µl load of TriDye, the 3.0-kb band has as reference mass (R_{mass}) of 62.5 ng. The masses of each band of interest (B_{mass}) were calculated by simply relating the known peak area (R_{area}) and mass (R_{mass}) of the 3.0-kb TriDye band to the peak area (B_{area}) of bands of interest in subsequent lanes (*see* ***Figure 14***):

$$B_{mass} = \frac{B_{area} R_{mass}}{R_{area}}$$

The mass of each band of interest was then converted to concentration (ng/µl) by dividing by the volume of DNA that was loaded into the agarose gel. This volume of DNA was usually the amount used in a preceding diagnostic restriction digest (e.g. 5 µl).

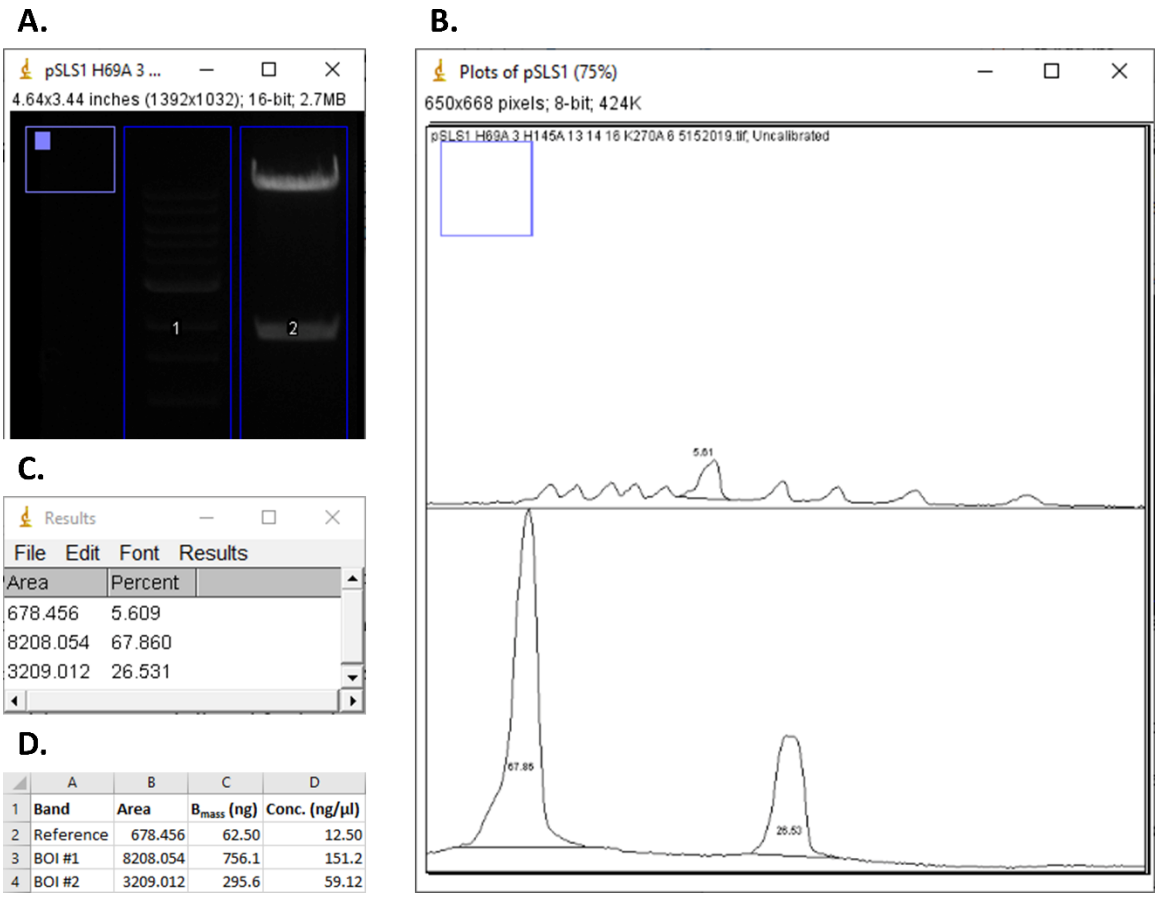

Figure 14. DNA quantification using ImageJ and Microsoft Excel.

(A) The DNA ladder lane and any additional lane-of-interest is selected in ImageJ using the rectangle tool. (**B**) The peak intensities of each lane are then automatically plotted, and each peak is area-bound using the straight-line tool. (**C**) Each area-bound peak is then selected using the magic-wand tool and labeled. (**D**) Finally. these calculated peak intensities are converted to DNA concentrations for each band-of-interest (BOI).

CHAPTER 3

PLASMID CONSTRUCTION METHODS AND RESULTS

High-Fidelity PCR Amplification of Wild-type ECM14

Primer sequences for amplifying the *Saccharomyces cerevisiae* ECM14 (YHR132C) open reading frame (ORF) were obtained from the Saccharomyces Genome Deletion Project (Baudin et al., 1993; Wach et al., 1994). These forward (_A) and reverse (_D) ORF primers are designed to also include several hundred bases of flanking sequence around the targeted gene to ensure adequate coverage of both endogenous promotor and terminator sequences. As such, amplicons produced using these forward and reverse primers (e.g. ECM14_A and ECM14_D) are referred to as "expression cassettes." In addition to the standard ECM14_A and ECM14_D primers, slightly modified primers were designed to include a 5' *Bam*HI restriction site (ECM14_A-*Bam*HI and ECM14_D-*Bam*HI) for future ligation into various plasmid vectors. These *Bam*HI primers also included a 4-bp TATA sitting sequence preceding the GGATCC *Bam*HI recognition site on the 5' end to facilitate optimal enzyme binding and interaction. All primers were ordered from Thermo Fisher Scientific® (*see* ***Table 1***) and reconstituted with Buffer EB (10 mM Tris-Cl, pH 8.5) to 100 µM. These reconstituted primers were further diluted to a final working concentration of 10 µM before being used in PCR or Sanger sequencing reactions.

Genomic DNA was extracted from ECM14 wild-type Y1239 yeast cells cultured on standard YEPD media. The high-fidelity Platinum SuperFi® DNA Polymerase from Invitrogen® was used to amplify the ECM14 ORF. A PCR master mix was prepared for n+1 reactions (n=4) by combining 127.5 µl of sterile Milli-Q® water, 50 µl of 5X SuperFi buffer, 20 µl of 2mM dNTP mix, and 2.5 µl of SuperFi DNA Polymerase (2 U/µl). Either

5 µl or 1 µl of Y1239 gDNA was used as template DNA with both the standard ECM14 and modified ECM14-*Bam*HI primers to a final volume of 50 µl per reaction. Forward and reverse primers were added to a final concentration of 0.5 µM. Each reaction tube was capped and carefully mixed before placing into a Bio-Rad® T100 Thermal Cycler. The following thermal profile was performed: initial denaturation at 98°C for 30 seconds, followed by 35 cycles of denaturation at 98°C for 10 seconds, annealing at 63.6°C for 10 seconds, and extension at 72°C for 60 seconds. The profile was completed with a final extension at 72°C for 5 minutes. Each PCR reaction was combined with 10 µl of New England Biolabs'® Gel Loading Dye, Purple (6X) and ran in a 0.8% agarose gel (0.5 µg/ml EtBr) for visualization and gel extraction/cleanup (*see* ***Figure 15***).

The pSLS1 vector contains only three unique restriction sites (*Bam*HI, *Sal*I, and *Sma*I) for subcloning applications (*see above* ***Figure 11***). Thus, it was required to utilize custom PCR primers in order to introduce novel restriction sites (preferably those with sticky ends) flanking the traditional ECM14 amplicon. The original ECM14_A and ECM14_D primers from the *Saccharomyces* Genome Deletion Project were used as a control to demonstrate amplification of wild-type ECM14 (1.9 kb) from Y1239 gDNA.

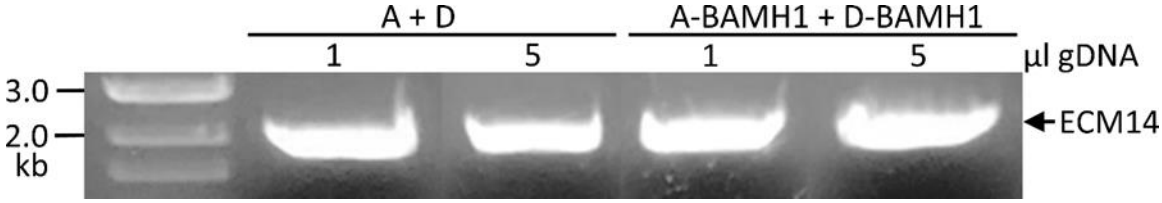

Figure 15. High-fidelity PCR amplification of wild-type ECM14.

The original ECM14_A and ECM14_D primers were used to demonstrate amplification of wild-type ECM14 (1.9 kb) from Y1239 gDNA. Custom ECM14_A and ECM14_D primers were designed to harbor terminal BamH1 restriction sites to facilitate future ligation into various plasmid vectors. Either 1 µl or 5 µl of template Y1239 gDNA was used in each PCR reaction. All four reactions were successful in amplifying ECM14.

Cloning of Wild-type ECM14 into pBluescript

The ECM14 PCR product harboring terminal *Bam*HI restriction sites and the target pBluescript II SK(+) plasmid vector were both digested with the New England Biolabs' *Bam*HI-HF restriction enzyme. [Although there are four commercially available variants of the pBluescript II vector, these differences were not of interest in this study. Any one of these pBluescript variants can be used as a routine cloning plasmid.] For the ECM14-*Bam*HI PCR product, 50 µl of gel-purified PCR product was combined with 6 µl of CutSmart Buffer, 2 µl of sterile Milli-Q water, and 2 µl of *Bam*HI-HF enzyme. For pBluescript II SK(+), 2 µl of pBluescript plasmid (1 µg/µl) was combined with 15 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of BamHI-HF enzyme. Both digestion reactions were thoroughly incubated at 37°C for >1 hour and were cleaned up immediately using the QIAquick gel extraction kit without actually running each digest in an agarose gel (3 volumes of Buffer QG and 1 volume of isopropanol were simply added prior to applying each digest to a fresh spin column). The pBluescript-*Bam*HI digest was additionally treated with 1 µl of New England Biolabs' Calf Intestinal Alkaline Phosphatase (CIAP) to remove 5' terminal phosphate groups and, thus, prevent vector re-circularization during subsequent ligation reactions. This CIAP-treated pBluescript-*Bam*HI digest was also cleaned up immediately using the QIAquick kit. Each cleaned-up digest was quantified in a 0.8% agarose gel (0.5 µg/ml EtBr) with 5 µl TriDye Ladder as a reference standard.

The ECM14-*Bam*HI insert was ligated into the pBluescript-*Bam*HI-CIAP vector using T4 DNA ligase with a standard vector-to-insert ratio of 1:3. The ligation reaction contained 6 µl of pBluescript-*Bam*HI-CIAP (0.042 pmol), 4 µl of ECM14-*Bam*HI (0.126

pmol), 7 µl of sterile Milli-Q water, 2 µl of T4 DNA ligase buffer, and 1 µl of T4 DNA ligase. The ligation reaction was incubated at 24°C for 1 hour. Three 50-µl aliquots of Invitrogen Subcloning Efficiency™ DH5α Competent Cells were slowly thawed on wet ice and mixed with either 5 µl of sterile Milli-Q water (negative control), 5 µl pBluescript-ECM14 ligation mixture, or 1 µl pBluescript (positive control). Each transformation mixture was incubated on wet ice for 30 minutes, heat shocked at 42°C for 45 seconds, and returned to wet ice for 5 minutes. Each reaction was incubated for 60 minutes at 37°C in 200 µl of fresh LB without ampicillin to allow for adequate expression of the ampicillin resistance gene prior to plating. These outgrowth cultures were subsequently plated on dry IPTG/X-Gal LB agar plates with ampicillin (starting with the negative control first), incubated at 37°C overnight for >20 hours, and screened for white colonies.

Four putative pBluescript-ECM14 colonies (white) were cultured overnight, miniprepped using the NID method, and digested with *Bam*HI-HF enzyme at 37°C for 60 minutes. Each of these diagnostic restriction digests included 5 µl of plasmid DNA, 12 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of *Bam*HI-HF enzyme. Each digest was electrophoresed in a 0.8% agarose gel (0.5 µg/ml EtBr) at 100V until sufficient band separation (*see* ***Figure 16***). The combined peak area of both the pBluescript band and the ECM14 band were used to calculate the concentration of each plasmid preparation.

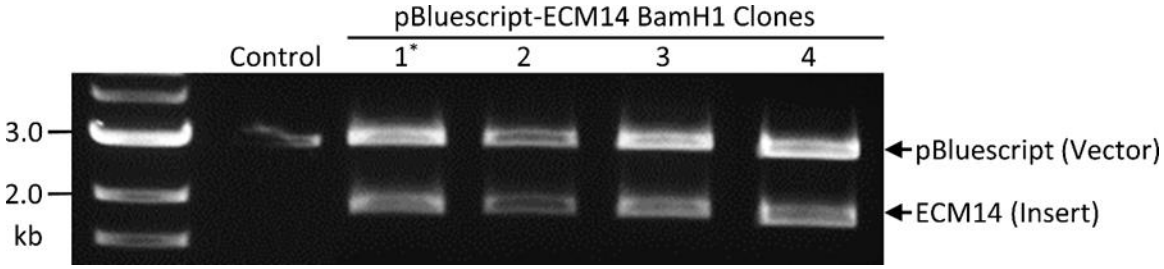

Figure 16. Cloning of Wild-type ECM14 into pBluescript.

The wild-type ECM14 PCR product harboring terminal *Bam*HI sites and the pBluescript II SK(+) vector were first digested with BamH1. The pBluescript vector was then treated with an alkaline phosphatase (CIAP) prior to ligation by T4 DNA ligase. Four putative pBluescript-ECM14 preparations and empty pBluescript II (3.0 kb) were digested with *Bam*HI to confirm the presence or absence of the ECM14 gene insert (1.9 kb).

Confirmation of ECM14 Sequence Accuracy

Sanger sequencing was performed by GenScript (Piscataway, N.J.) The four putative pBluescript-ECM14 plasmids were sequenced using the ECM14_A, ECM14_D, and ECM14_Internal primers (*see* ***Table 1***) to ensure adequate sequencing coverage of the ECM14 gene insert. Three sequencing samples for each plasmid were prepared by premixing 10 µl of plasmid DNA (>50 ng/µl), 2.5 µl (50 pmol) of primer (either ECM14_A, ECM14_D, or ECM14_Internal), and 2.5 µl of sterile Milli-Q water. Samples were shipped overnight in 1.5-ml microcentrifuge tubes sealed with Parafilm.

Sequencing results were first processed by visually inspecting the chromatograms using either Chromas or UGENE for base call quality and subsequently truncating any poor terminal sequence as necessary (Okonechnikov et al., 2012). Each sequence was then aligned with the wild-type ECM14 reference sequence using NCBI Nucleotide BLAST (BLASTn) and inspected for any sequence gaps and erroneous base identities. Any discrepancies were further considered by manually re-inspecting the chromatograms and, when possible, comparing regions of overlapping sequence coverage for any additional clues for determining consensus. Finally, the three contiguous sequences were assembled using the CAP3 sequence assembly program and saved in FASTA format.

The ECM14 insert sequence produced by this study was merged with the pBluescript II SK(+) sequence in SnapGene to produce new recombinant pBluescript-ECM14 maps for both orientations (*see* ***Figure 17***).

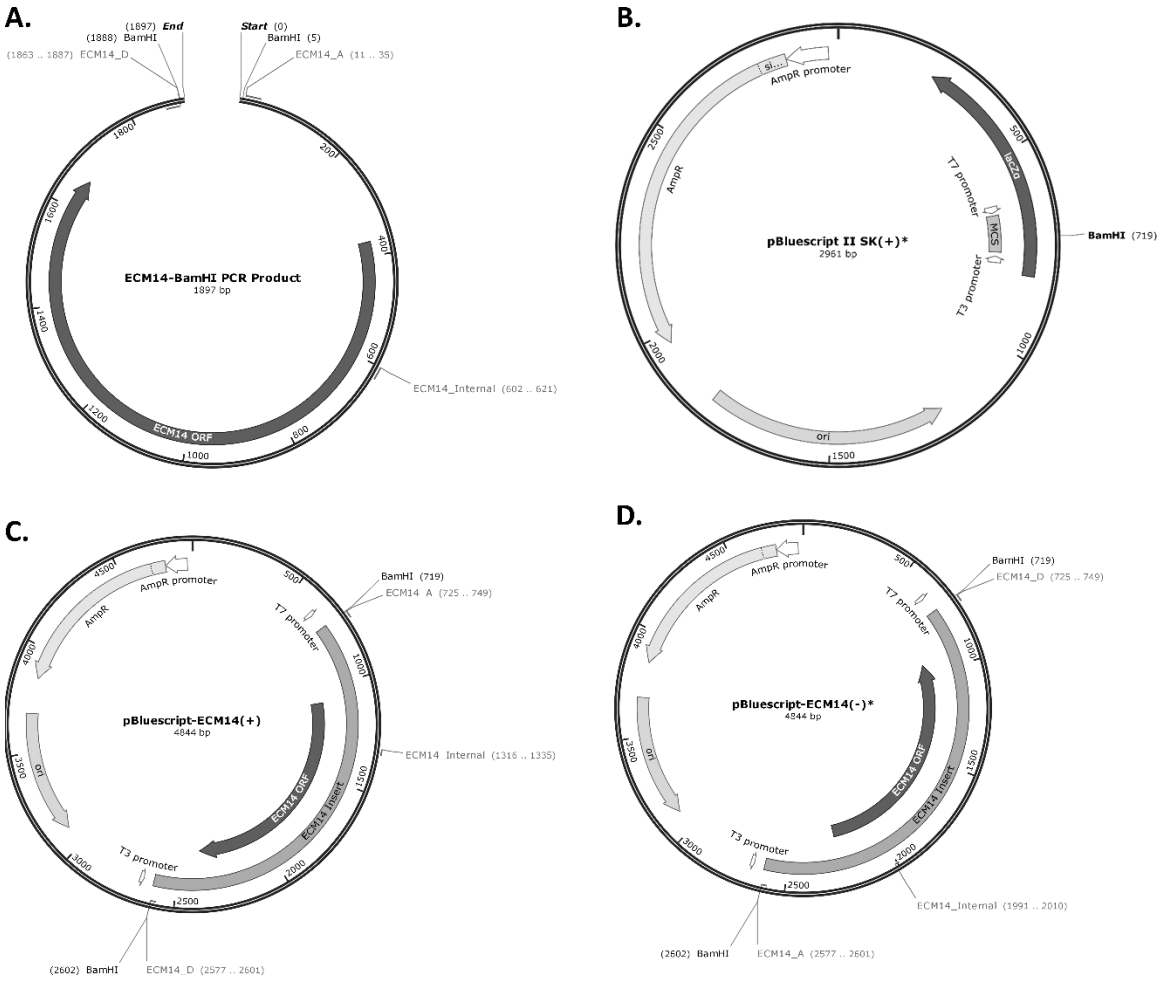

Figure 17. Diagrams of the ECM14-BamHI PCR product, the pBluescript II SK(+) vector, and both orientations of the pBluescript-ECM14 recombinant plasmid.

(**A**) The ECM14 PCR product harboring terminal *Bam*HI restriction sites. Relevant primer sites for amplification and sequencing are annotated. (**B**) The pBluescript II SK(+) plasmid contains a multiple cloning site (MCS) and blue-white screening capability (LacZα). (**C**, **D**) Either ECM14 insert orientation is possible in the recombinant pBluescript-ECM14 plasmid due to only one restriction enzyme being used for cloning. Only plasmid features relevant to this study are annotated (*).

Subcloning of ECM14 from pBluescript-ECM14 into pSLS1

The wild-type ECM14 gene insert was subcloned from the pBluescript-ECM14 (4,838-bp) parent plasmid into the pSLS1 (10.8-kb) destination plasmid. At least 2 µg of both pBluescript-ECM14 and pSLS1 were thoroughly digested with *Bam*HI to ensure adequate recovery yields for future downstream applications. For pBluescript-ECM14, 12 µl of pBluescript-ECM14 plasmid (170 ng/µl) was combined with 32 µl of sterile Milli-Q water, 5 µl of CutSmart Buffer, and 1 µl of *Bam*HI-HF enzyme. For pSLS1, 7 µl of pSLS1 plasmid (300 ng/µl) was combined with 37 µl of sterile Milli-Q water, 5 µl of CutSmart Buffer, and 1 µl of *Bam*HI-HF enzyme. Both digestion reactions were allowed to incubate for >5 hours at 37°C. The pSLS1-*Bam*HI digest was additionally treated with 2 µl of Calf Intestinal Alkaline Phosphatase (CIAP) to remove 5' terminal phosphate groups and, thus, prevent vector re-circularization during subsequent ligation reactions. Both digests were electrophoresed in a 0.8% agarose gel (0.5 µg/ml EtBr) for 60 minutes in order allow for sufficient separation of the pBluescript (2,961-bp) and ECM14 (1,883-bp) bands. The pSLS1 and ECM14 bands were both gel extracted using the modified protocol for large DNA fragments, ran in an additional 0.8% agarose gel (0.5 µg/ml EtBr), and quantified using ImageJ.

The pSLS1-ECM14 ligation reaction was performed overnight at 16°C with a 1:3 vector-to-insert ratio. This ligation reaction contained 5 µl of ECM14-*Bam*HI (20 ng/µl, 0.081 pmol), 2 µl of pSLS1-*Bam*HI-CIAP (94 ng/µl, 0.026 pmol), 10 µl of sterile Milli-Q water, 2 µl of T4 DNA Ligase Buffer, and 1 µl of T4 DNA Ligase. An additional ligation reaction lacking ECM14-*Bam*HI was also performed as a test for CIAP treatment efficacy.

Additionally, the pSLS1-ECM14 ligation reaction was further diluted 5X as recommended by the Invitrogen Max Efficiency™ DH5α transformation protocol.

A fresh 200-µl aliquot of Invitrogen Max Efficiency DH5α was slowly thawed on wet ice, split into four 50-µl aliquots, and gently mixed with either 5 µl of the pSLS1-ECM14 ligation reaction, 5 µl of the 5X dilution, 5 µl of the CIAP ligation control, or 1 µl of pUC19 control plasmid. Each transformation mixture was incubated on wet ice for 30 minutes, heat shocked at 42°C for 45 seconds, and returned to wet ice for 2 minutes. Each reaction was incubated for 60 minutes at 37°C in 450 µl of fresh SOC without ampicillin to allow for adequate expression of the ampicillin resistance gene prior to plating. These outgrowth cultures were subsequently plated on dry LB agar plates (250 µl/plate) with ampicillin and incubated at 37°C overnight for >20 hours. The pUC19 control transformation was plated on dry IPTG/X-Gal LB agar plates with ampicillin.

Twelve putative pSLS1-ECM14 colonies were cultured overnight and miniprepped using the NID method. Nine putative pSLS1-ECM14 plasmids (1, 3, 4, 5, 6, 7, 8, 11 and 12) were randomly selected and digested with *Bam*HI-HF at 37°C for 60 minutes. Each of these diagnostic digests included 5 µl of plasmid DNA, 12 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of *Bam*HI-HF enzyme. Each digest was electrophoresed in a 0.8% agarose gel (0.5 µg/ml EtBr) and imaged using the AlphaImager HP gel imaging system (*see* ***Figure 18***). Only those digests that clearly liberated both a pSLS1 band (10.8-kb) and an ECM14 band (1,883-bp) were considered for further interrogation (1, 3, 5, and 7). The remaining digests (4, 6, 8, 11, and 12) contained a pBluescript band (2,961-bp) instead of a pSLS1 band and were discarded.

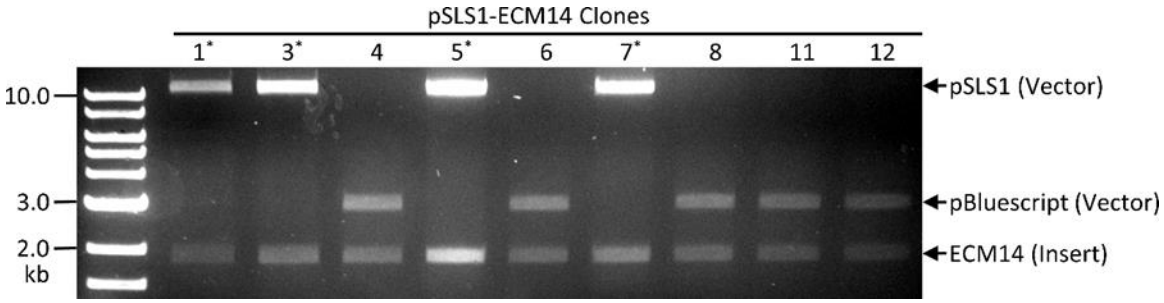

Figure 18. Cloning of wild-type ECM14 from pBluescript-ECM14 into pSLS1.

The sequenced-verified ECM14 gene insert was subcloned from the pBluescript-ECM14 (4.8 kb) parent plasmid into the pSLS1 (10.8 kb) destination vector. Nine putative pSLS1-ECM14 preparations were screened using a *Bam*HI diagnostic digest. Only those digests that clearly liberated both a pSLS1 band (10.8 kb) and an ECM14 band (1.9 kb) were considered for further interrogation (*).

Determination of ECM14 Insert Orientation in pSLS1-ECM14

A sequential diagnostic restriction digest with *Sma*I and *Eco*RI was performed in order to determine the orientation of the ECM14 gene insert in pSLS1-ECM14. Because only a single restriction enzyme was used during the original subcloning procedure, either insert orientation was equally probable. Although the ECM14 gene insert includes both endogenous promotor and terminator sequences (and, thus, should be transcribable in either orientation), it was still desirable to document the orientation for future reference.

*Sma*I and *Eco*RI were intentionally chosen for this experiment based on two underlying assumptions: 1) S*ma*I cuts only within the parent vector and 2) *Eco*RI cuts asymmetrically within the gene insert. [These two basic assumptions can be applied to similar experiments.] The restriction sites within the ECM14 insert sequence produced by this study were mapped using EMBOSS Restrict (Hancock & Bishop, 2014) and SnapGene (from GSL Biotech; available at snapgene.com). It was noted that *Eco*RI cuts the ECM14 insert sequence 378 bases from one end.

Due to incompatible incubation temperatures, the *Sma*I digest was performed first at the lower temperature of 25°C (5 µl of plasmid DNA, 12 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of *Sma*I enzyme). The *Eco*RI digest was performed next at 37°C (5 µl of plasmid DNA, 12 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of EcoRI-HF enzyme). Each diagnostic digest was thoroughly incubated for >1 hour. Each diagnostic digest was combined with 4 µl of New England Biolabs' Gel Loading Dye, Purple (6X), ran in a 0.8% agarose gel (0.5 µg/ml EtBr) at 100V until sufficient band separation, and imaged using the AlphaImager HP gel imaging system (*see* ***Figure 19***).

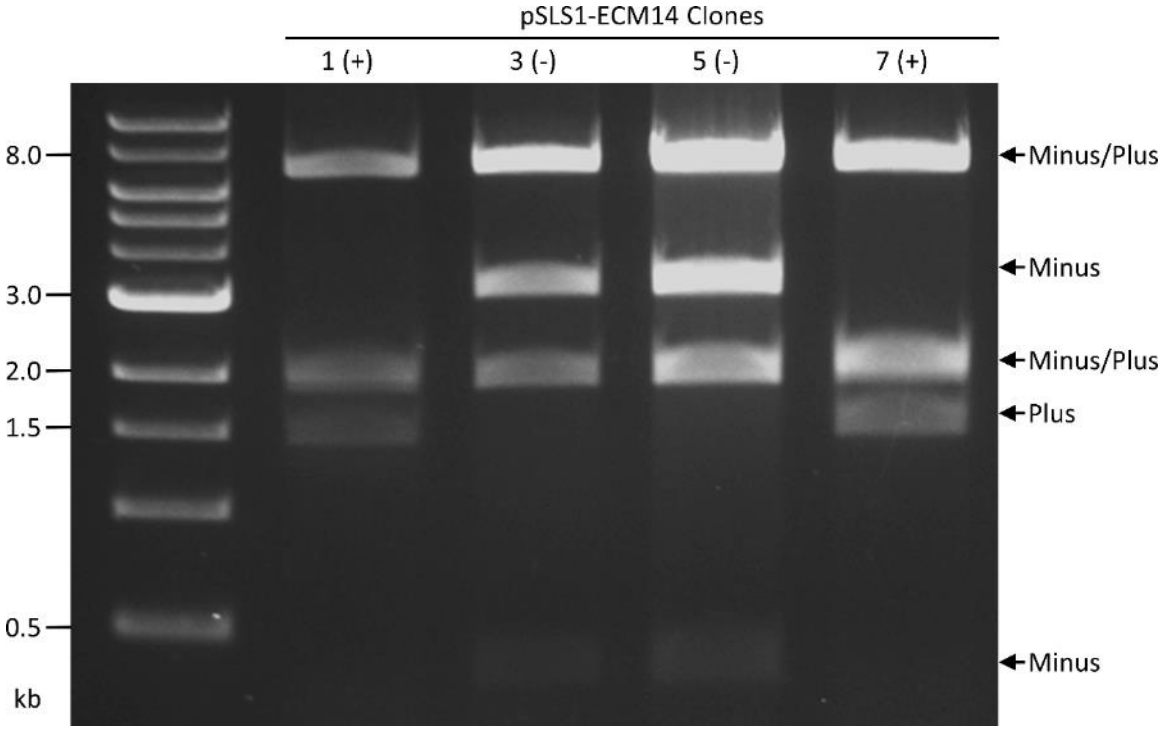

Figure 19. Determination of ECM14 insert orientation in pSLS1-ECM14.

A sequential diagnostic restriction digest with *Sma*I and *Eco*RI was performed in order to determine the orientation of the ECM14 gene insert in pSLS1-ECM14.

The complete restriction map for the pSLS1 plasmid was recreated *in silico* using the original plasmid construction workflow reported by Barbour et al. (2000). It was noted in a personal communication with Steve Brill (Rutgers University, Piscataway, N.J.) that the pJM555 plasmid that was used as the *ADE3* fragment source by Barbour et al. (2000) is now internally called pJM2555. Additional personal communications with Melanie Dobson (Dalhousie University, Nova Scotia, Canada), Doug Koshland (University of California, Berkeley, California), and Steve Henikoff (Fred Hutchinson Cancer Research Center, Seattle, Washington) further corroborated the sequence of the original pSLS1 map and its constituent gene fragments.

The sequence-complete pSLS1 map revealed an *Eco*RI site at position 5,605 (downstream of the *Sma*I cloning site at position 5,595) that was previously undocumented

by Barbour et al. (2000). Therefore, the *EcoRI* double digest with *SmaI* that was performed in this study was unnecessary. Besides an additional 10-bp *SmaI-EcoRI* fragment, a single *EcoRI* digest would have produced nearly identical restriction fragments (*see* ***Figure 20***).

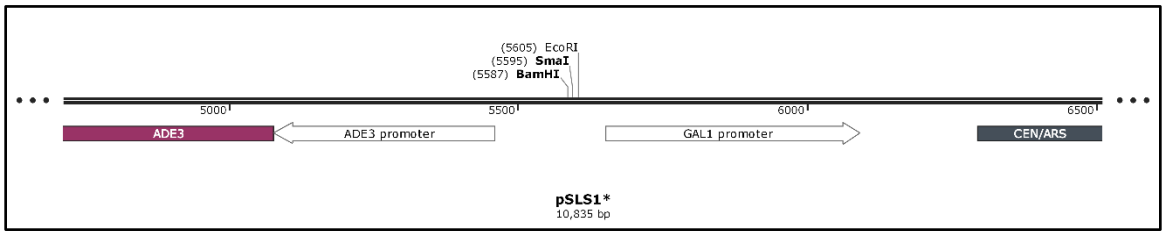

Figure 20. A truncated segment of the sequence-complete pSLS1 map showing the previously undocumented EcoRI site at position 5,605.

The ECM14 insert sequence produced by this study and the orientation data were merged with the newly constructed pSLS1 map using SnapGene to produce the novel recombinant pSLS1-ECM14 maps for both insert orientations (*see* ***Figure 21***). The orientation nomenclature of pSLS1-ECM14 was defined according to the direction of ECM14 transcription relative to the GAL1 promotor (*plus* in the same direction; *minus* in the opposite direction).

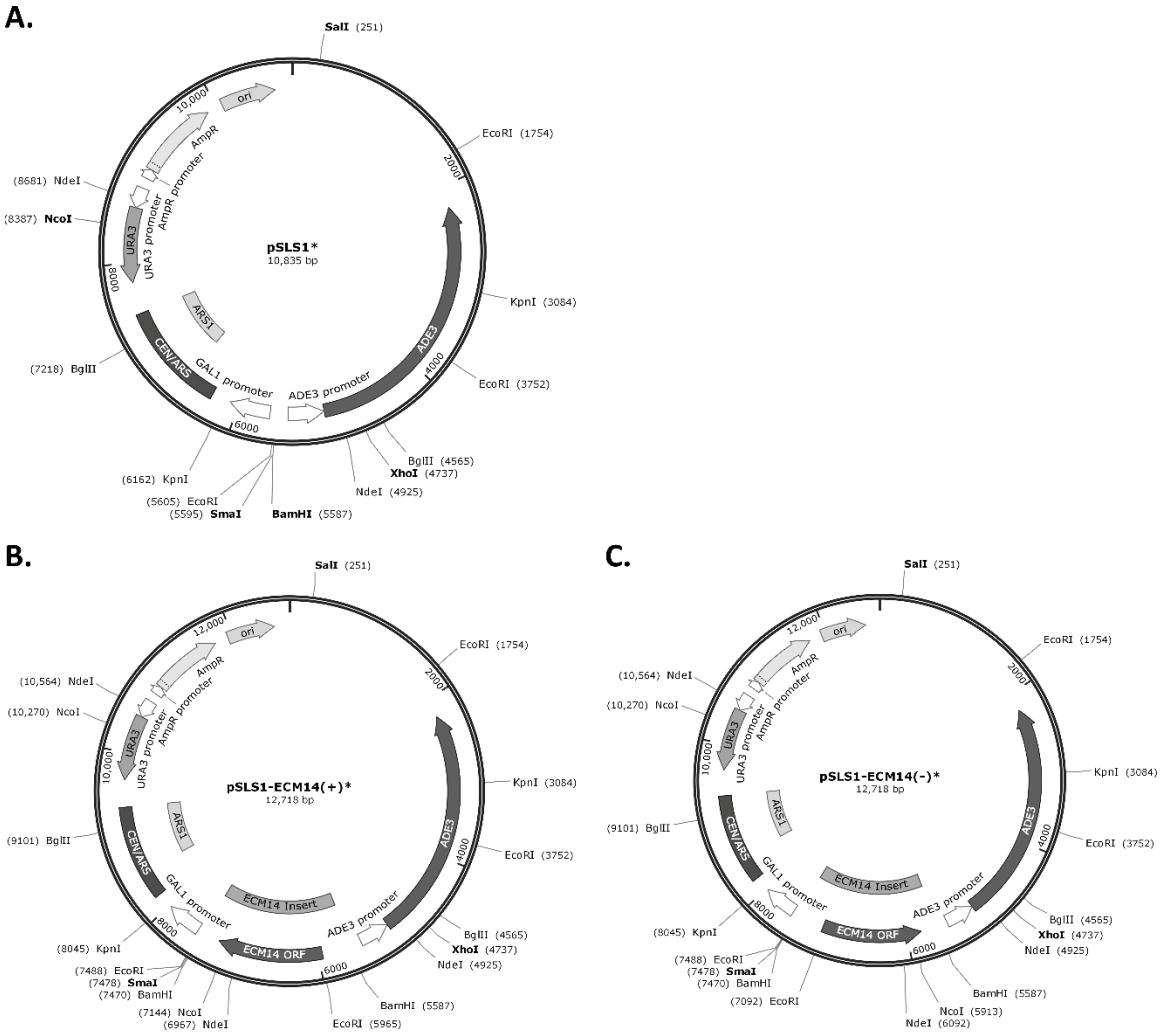

Figure 21. Diagrams of the original pSLS1 vector and both orientations of the novel pSLS1-ECM14 recombinant plasmid.

(**A**) The original pSLS1 plasmid contains the bacterial origin of replication sequence (ORI), the ampicillin resistance gene (AmpR) for selection in bacteria, the CEN/ARS sequence for replication in yeast, the URA3 gene for selection in yeast, and the *ADE3* gene for color-colony reporting (Barbour et al., 2000; Xiao, 2006). The CEN sequence is under the influence of the GAL1 galactose promotor. (**B**, **C**) The recombinant pSLS1-ECM14 plasmids containing either ECM14 insert orientation. Only the minus insert orientation was used for this study as this ECM14 ORF is transcribed away from the CEN/ARS region. Only plasmid features relevant to this study are annotated (*).

QuikChange Site-Directed Mutagenesis of ECM14

QuikChange site-directed mutagenesis (SDM) was used to introduce targeted changes to the wild-type ECM14 ORF. Because QuikChange SDM is more efficient with smaller plasmid constructs, the Sanger-sequenced pBluescript-ECM14 (4.8-kb) plasmid was selected instead of the larger pSLS1-ECM14 (12.7-kb) plasmid as the template harboring the wild-type ECM14 sequence to be mutagenized (*see* ***Table 2***). QuikChange SDM operates on the functional principles of PCR using complementary mutagenic primers harboring 13-20 bp of homology, the mutation of interest (1-3 bp), and another 13-20 bp of homology (*see* ***Table 1***). Because the sense and antisense mutagenic primers are typically complementary (3' overhangs are generally recommended), QuikChange SDM is inherently quite inefficient because these primers can erroneously self-anneal and produce primer dimers. Finally, the original parent plasmid is degraded prior to transformation and amplification in *E. coli* using the restriction enzyme *Dpn*I, which selectively digests methylated dsDNA (PCR synthesized dsDNA is not methylated) (*see* ***Figure 22***).

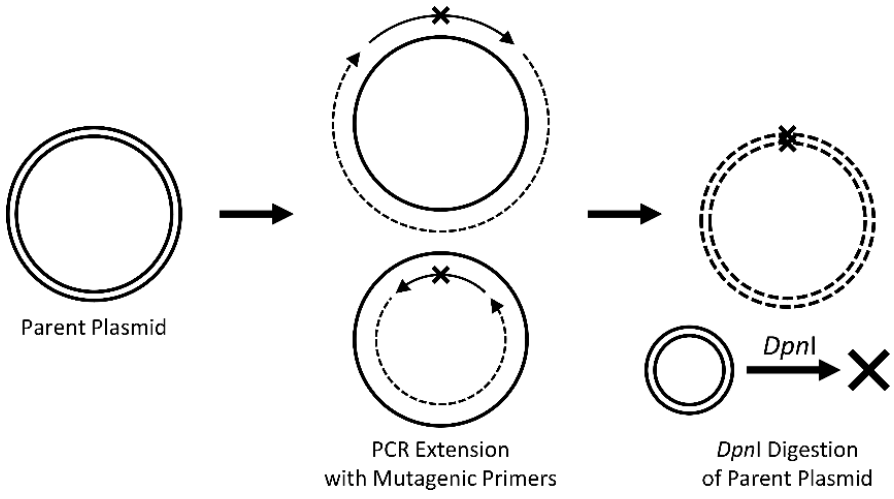

Figure 22. The basic QuikChange site-directed mutagenesis workflow.

A parent plasmid harboring the wild-type sequence to be mutated is extended using PCR with complementary mutagenic primers that introduce a 1-3 bp substitution.

Three ECM14 residues were selected for QuikChange SDM: H69, H145, and K270 (*see **Table 3***). These three residues/positions were chosen based on their conservation and demonstratable relevance in catalytically active MCP orthologs: one selected as being critical for zinc binding (H69), one selected for substrate binding (R145), and one selected for catalysis (E270). The positively charged side chain of histidine (H69 and H196) is fundamentally required for proper zinc coordination in catalytically active MCP orthologs. Therefore, the **ECM14-H69A** mutant specifically disrupts this predicted zinc coordination by substituting this histidine (H) for an alanine (A), a small and uncharged amino acid. Because histidine and arginine (R) are both positively charged, the **ECM14-H145A** mutant specifically disrupts predicted substrate binding by replacing this conserved, positively charged residue with a small and uncharged alanine residue. Finally, because the glutamate (E) at position 270 is required for catalysis in active MCP orthologs, the **ECM14-K270E** and **ECM14-K270A** mutants specifically interrogate the relevance of the lysine (K) at this position in wild-type Ecm14.

All mutagenic primers were ordered from Thermo Fisher Scientific (*see **Table 1***) and reconstituted with Buffer EB (10 mM Tris-Cl, pH 8.5) to 100 μM. These reconstituted primers were further diluted to a final working concentration of 10 μM before being used in QuikChange SDM PCR reactions. The high-fidelity Platinum SuperFi DNA Polymerase from Invitrogen was used to perform the QuikChange SDM. A PCR master mix was prepared for n+1 reactions (n=4) by combining 142.5 μl of sterile Milli-Q water, 50 μl of 5X SuperFi buffer, 25 μl of 2 mM dNTP mix, and 2.5 μl of SuperFi DNA Polymerase (2 U/μl). The Sanger-sequenced pBluescript-ECM14 parent plasmid stock was diluted to a working concentration of 0.2 ng/μl (1 μl being used per SDM reaction). The amount of

parent pBluescript-ECM14 plasmid was minimized to ensure a high mutagenesis efficiency and thorough *Dpn*I digestion prior to transformation. Forward and reverse mutagenic primers were added to a final concentration of 0.5 μM. Because of the inherent inefficiencies of QuikChange SDM, three separate attempts with varying reaction conditions were performed in order to successfully produce the four desired ECM14 mutants. The location of each site to be mutagenized was mapped and visualized on the pBluescript-ECM14 map using SnapGene (*see* ***Figure 23***).

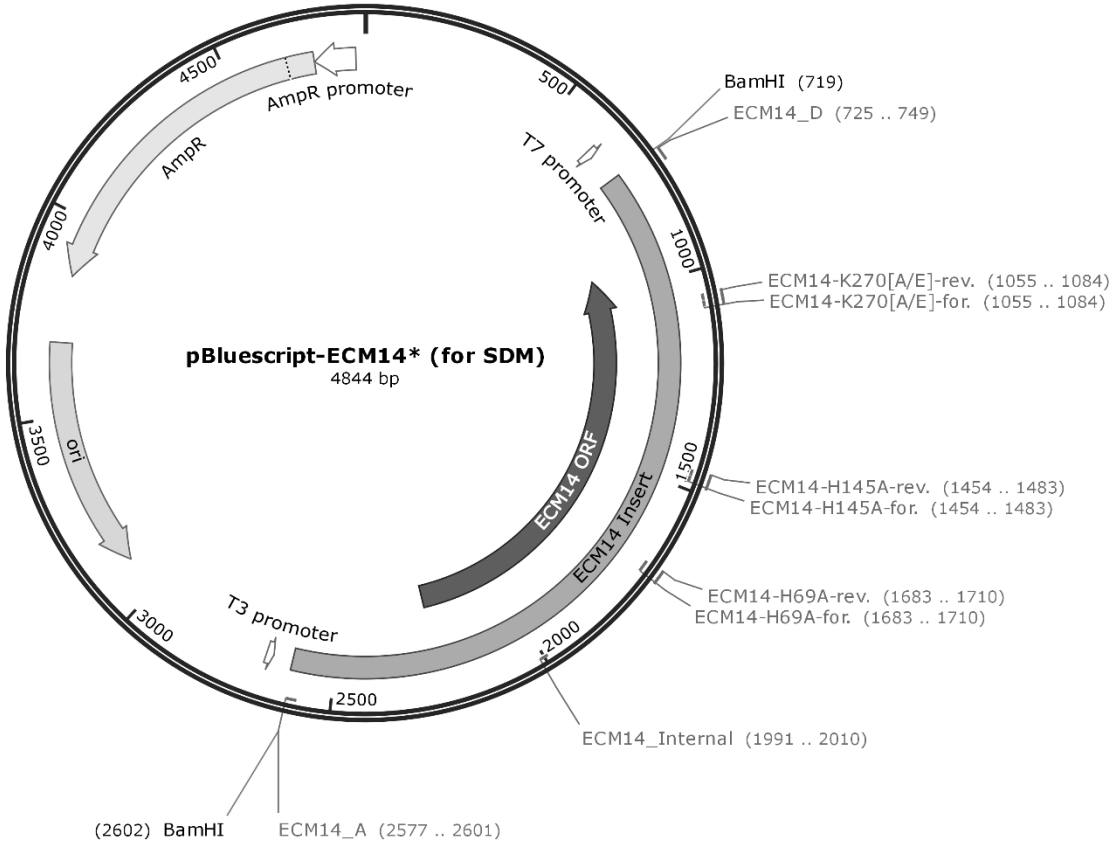

Figure 23. Locations for site-directed mutagenesis of pBluescript-ECM14.

The pBluescript-ECM14 plasmid was used instead of the pSLS1-ECM14 plasmid as the template for site-directed mutagenesis due to its smaller size. Relevant primer binding sites are annotated (H69A, H145A, K270A, and K270E). Only plasmid features relevant to this study are shown (*).

The first SDM attempt was performed using a non-traditional two-step thermal cycling protocol as recommended in the SDM application note for Platinum SuperFi DNA Polymerase. The following two-step thermal cycling profile was performed: initial denaturation at 98°C for 30 seconds, followed by 30 cycles of denaturation at 98°C for 10 seconds and a combined annealing-extension at 72°C for 2 minutes and 40 seconds (≈ 30 sec/kb). The profile was completed with a final extension at 72°C for 5 minutes. The parent plasmid was digested by adding 1 µl of New England Biolabs' *Dpn*I restriction enzyme to each reaction and incubating in the thermal cycler heat block for 1 hour at 37°C.

The second and third SDM attempts were performed using the more traditional three-step thermal cycling protocol. For the second SDM attempt, the following three-step thermal cycling profile was performed: initial denaturation at 98°C for 30 seconds, followed by 30 cycles of denaturation at 98°C for 10 seconds, annealing at 63°C for 10 seconds, and extension at 72°C for 2 minutes and 40 seconds (≈ 30 sec/kb). The profile was completed with a final extension at 72°C for 5 minutes. Additionally, the third SDM attempt also included 3% dimethyl sulfoxide (DMSO) in the PCR master mix to help discourage secondary structure formation in the template DNA. Consequently, because DMSO is known to lower primer melting temperatures, an annealing temperature of 59°C was utilized for the third SDM attempt. Like the first SDM attempt, the parent plasmids were digested by adding 1 µl of New England Biolabs' *Dpn*I restriction enzyme to each reaction and incubating in the thermal cycler heat block for 1 hour at 37°C.

A portion of each *Dpn*I-treated reaction (15 µl) was combined with 3 µl of New England Biolabs' Gel Loading Dye, Purple (6X) and ran in a 0.8% agarose gel (0.5 µg/ml EtBr) for verification prior to transformation and amplification in *E. coli*. (*See* ***Figure 24***).

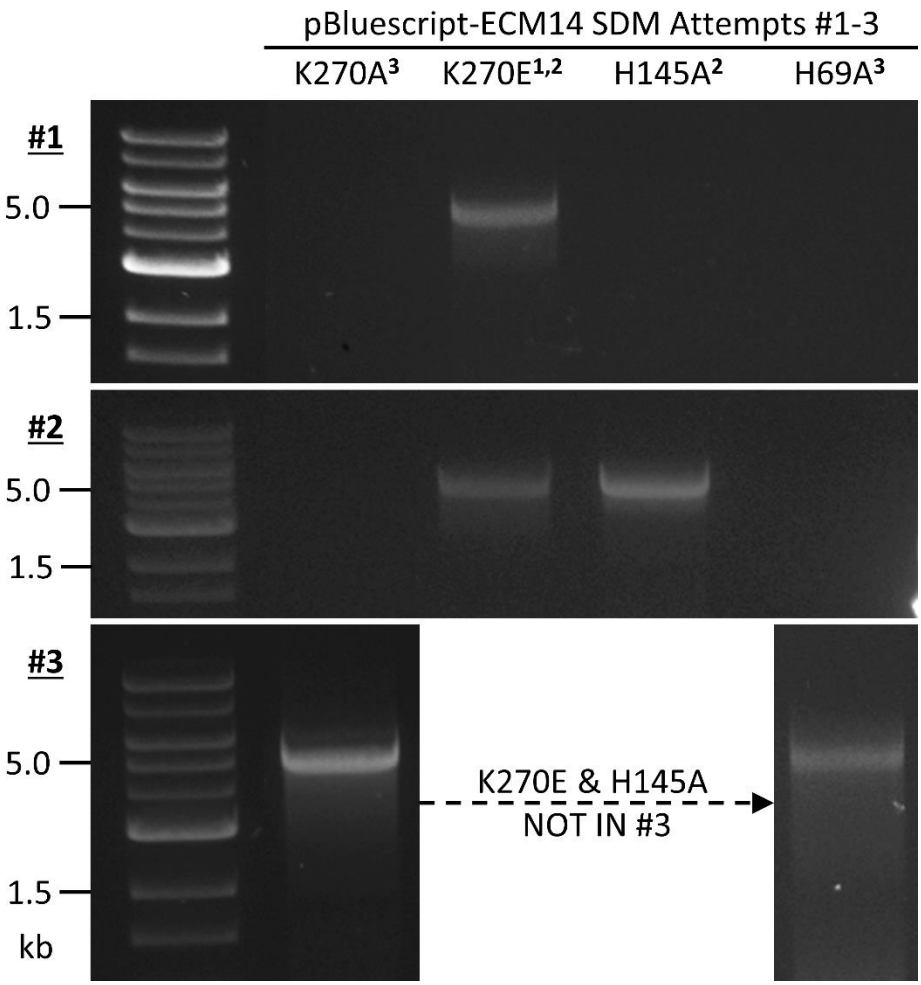

Figure 24. QuikChange site-directed mutagenesis of pBluescript-ECM14.

Three QuikChange SDM attempts were required to successfully produce the four desired ECM14 mutants. The K270E mutant was successfully synthesized in the first attempt; K270E and H145A in the second attempt; and K270A and H69A in the third attempt (K270E and H145A were not attempted in the third attempt). PCR reactions were treated with *Dpn*I prior to visualization and transformation.

Four 50-µl aliquots of Invitrogen Subcloning Efficiency DH5α Competent Cells were slowly thawed on wet ice and each mixed with 5 µl of a *Dpn*I-treated SDM reaction. Each transformation mixture was incubated on wet ice for 30 minutes, heat shocked at 42°C for 45 seconds, and returned to wet ice for 5 minutes. Each reaction was incubated at 37°C for 60 minutes in 200 µl of fresh LB without ampicillin to allow for adequate expression of the ampicillin resistance gene prior to plating. These outgrowth cultures were subsequently plated on dry LB agar plates with ampicillin, incubated at 37°C overnight for >20 hours, and screened for colonies. All four transformations yielded several colonies.

In order to potentially improve the effectiveness and resolution of downstream Sanger sequencing, the Qiagen QIAprep® Spin Miniprep Kit was used instead of the cheaper NID method for extracting and purifying each putative mutant pBluescript-ECM14 plasmid. Two monoclonal colonies from each transformation plate were used to inoculate 4 ml of LB/Amp and each culture was incubated overnight for 16-18 hours at 37°C with shaking at 250 RPM. Each overnight culture was pelleted at 6,800g for 3 minutes and resuspended in 250 µl of Buffer P1. Each Buffer P1 suspension was mixed with 250 µl of Buffer P2 and incubated at 25°C for 4 minutes. Immediately following incubation, these cell lysis reactions were thoroughly mixed with 350 µl of Buffer N3. Each of these neutralized reactions were centrifuged at 17,900g for 10 minutes to pellet cellular debris. Supernatants were transferred to a QIAprep spin column and centrifuged at 17,900g for 60 seconds. The flow-through was discarded. Each spin column was then washed with 500 µl of Buffer PB and centrifuged at 17,900g for 60 seconds. The flow-through was discarded. Finally, each spin column was washed with 750 µl of Buffer PE and centrifuged for 17,900g for 60 seconds. The flow-through was discarded. Any residual Buffer PE was

thoroughly removed with an additional round of centrifugation at 17,900g for 60 seconds. Plasmid DNA was eluted by carefully applying 50 µl of warm Buffer EB (55-65°C) to the center of each spin column and allowing each to incubate at room temperature for >5 minutes. Each spin column was placed into a clean 1.5-ml microcentrifuge tube and centrifuged for 90 seconds at 21,130g to elute the plasmid DNA.

Sanger sequencing was performed by GenScript (Piscataway, N.J.). The four putative mutant pBluescript-ECM14 plasmids were sequenced using the T7, T3, and ECM14_Internal primers (*see* ***Table 1***) to ensure adequate sequencing coverage of the ECM14 gene insert. Three sequencing samples for each plasmid were prepared by premixing 10 µl of plasmid DNA (>50 ng/µl), 2.5 µl (50 pmol) of primer (either T7, T3, or ECM14_Internal), and 2.5 µl of sterile Milli-Q water. Samples were shipped overnight in 1.5-ml microcentrifuge tubes sealed with Parafilm. [An initial attempt at sequencing the putative mutant pBluescript-ECM14 plasmids was performed using the ECM14_A and ECM14_D primers. However, these sequencing runs were inconclusive.]

Sequencing results were first processed by visually inspecting the chromatograms using either Chromas or UGENE for base call quality and subsequently truncating any poor terminal sequence as necessary. Each sequence was then aligned with the wild-type ECM14 reference sequence produced by this study. Complete, end-to-end sequencing coverage of each mutant ECM14 gene was not achieved. However, the mutagenized site for each was successfully sequenced and, thus, provided adequate confirmation that each of the four desired mutations were present (*see* ***Figure 25, Figure 26,*** *and* ***Figure 27***). Finally, each mutant ECM14 gene was subcloned into pSLS1 and the resulting insert orientations determined using the methodology demonstrated in this study.

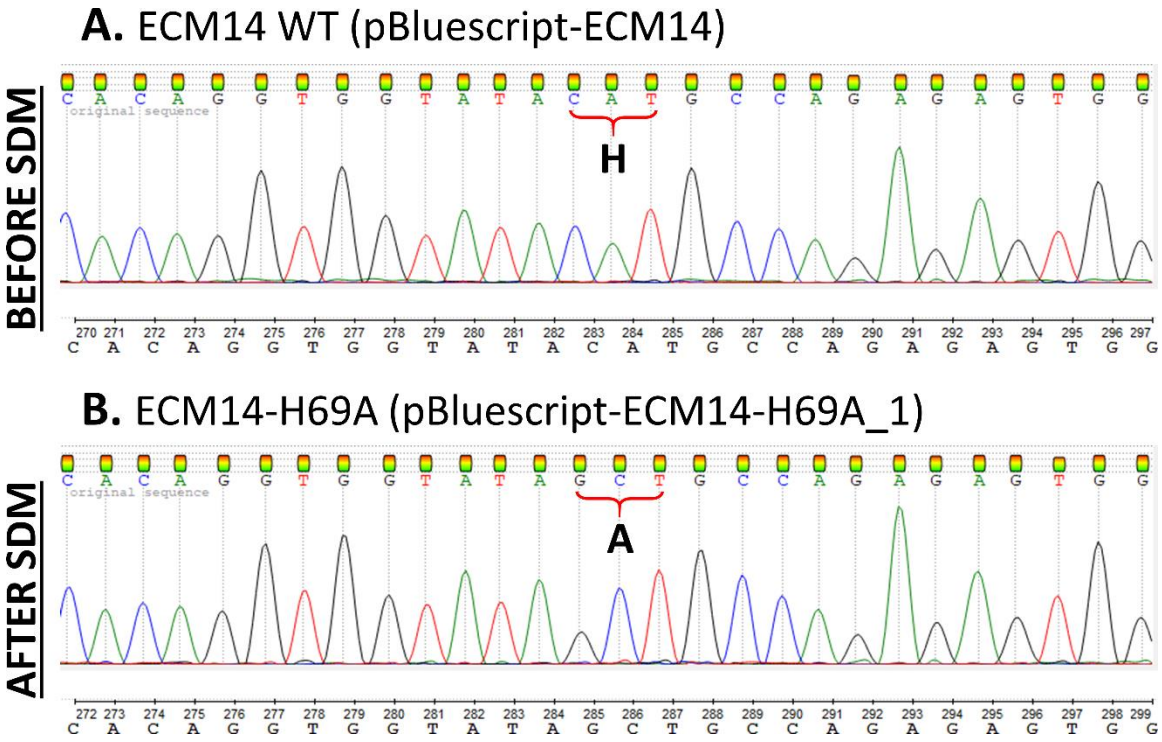

Figure 25. Confirmation of successful ECM14-H69A SDM

Three putative pBluescript-ECM14-H69A plasmids were sequenced using the T7, T3, and ECM14_Internal primers (*see* **Table 1**) to ensure adequate sequencing coverage of the ECM14 gene insert. Only the sequence and chromatograms of the 28-30 bases around mutagenic site are shown. (**A**) Wild-type pBluescript-ECM14 was used as a "**before SDM**" control. (**B**) The resultant "**after SDM**" mutagenized ECM14 sequence. The codon of interest is annotated in both sequences and the resultant amino acid identified. Sanger sequencing was performed by GenScript.

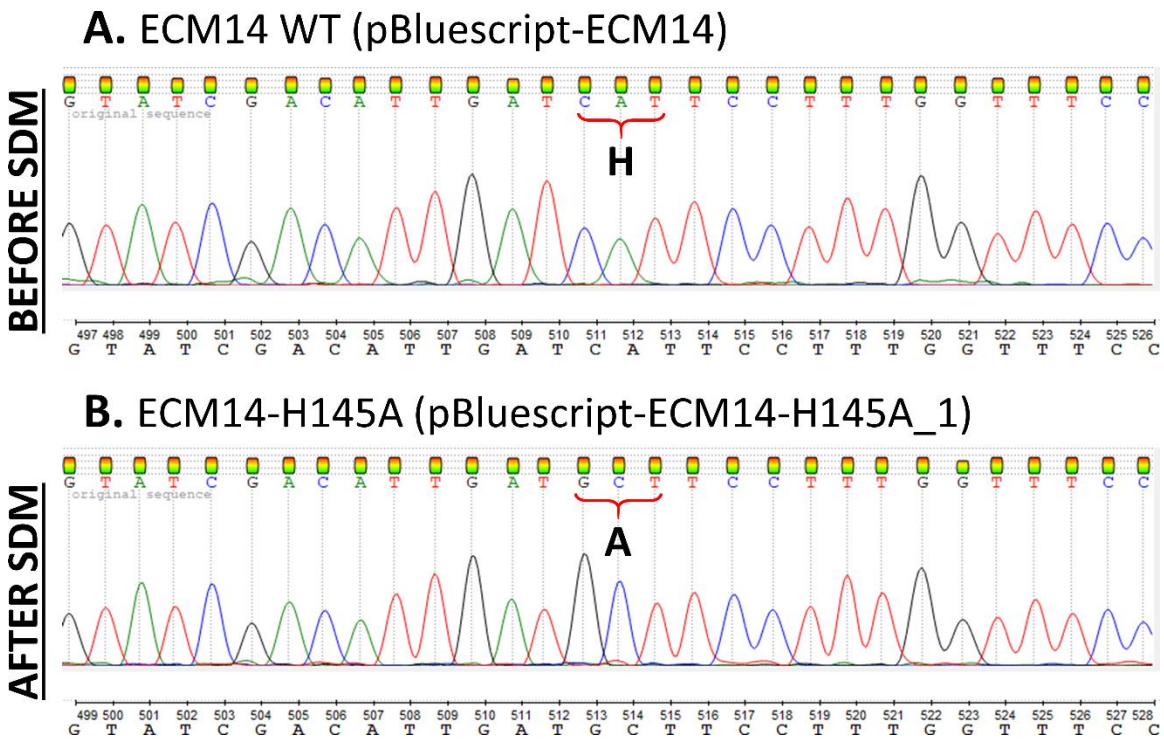

Figure 26. Confirmation of successful ECM14-H145A SDM.

Three putative pBluescript-ECM14-H145A plasmids were sequenced using the T7, T3, and ECM14_Internal primers (*see* **Table 1**) to ensure adequate sequencing coverage of the ECM14 gene insert. Only the sequence and chromatograms of the 28-30 bases around mutagenic site are shown. (**A**) Wild-type pBluescript-ECM14 was used as a "**before SDM**" control. (**B**) The resultant "**after SDM**" mutagenized ECM14 sequence. The codon of interest is annotated in both sequences and the resultant amino acid identified. Sanger sequencing was performed by GenScript.

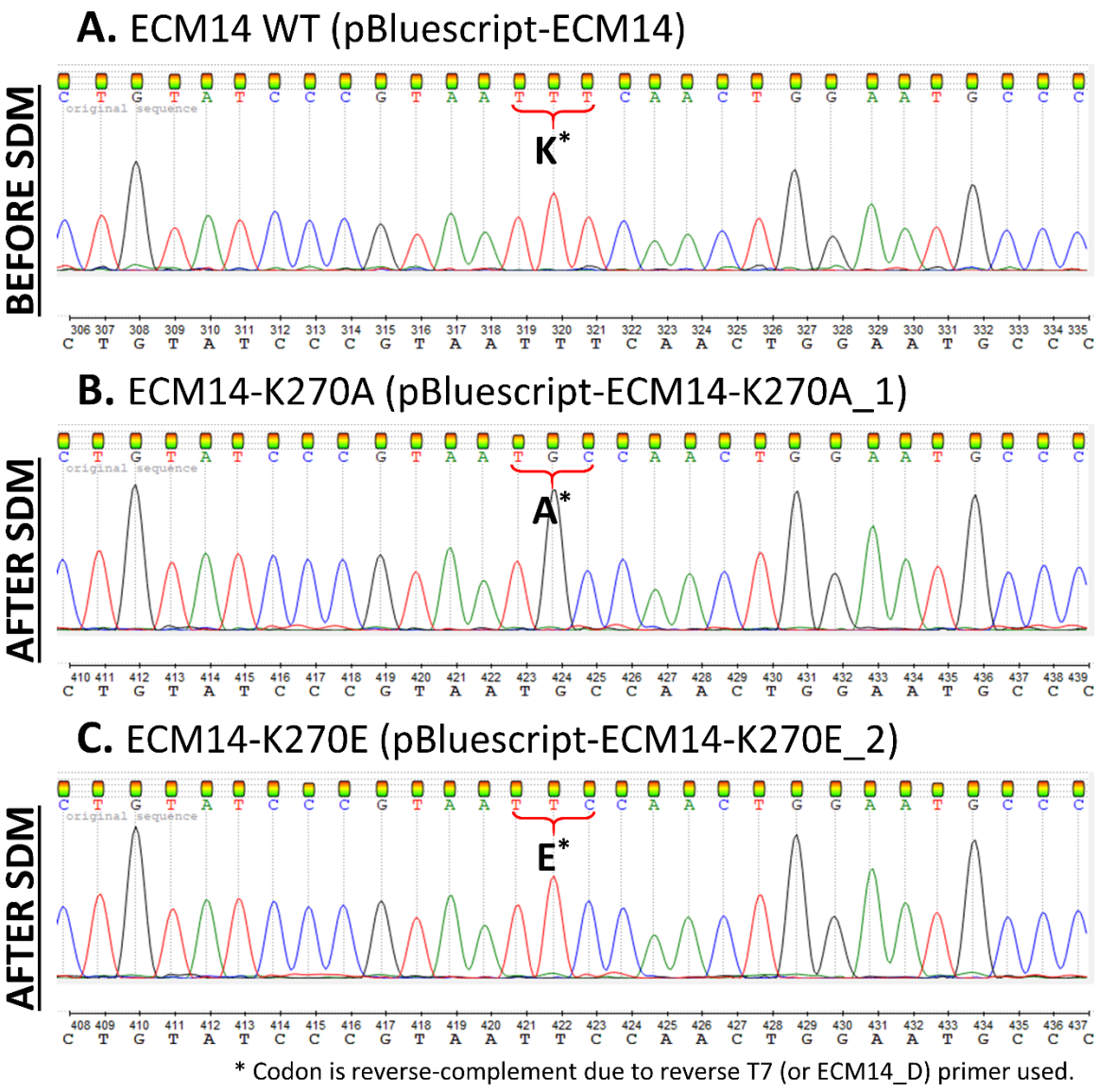

Figure 27. Confirmation of successful ECM14-K270[A/E] SDM.

Three putative pBluescript-ECM14-K270[A/E] plasmids were sequenced using the T7, T3, and ECM14_Internal primers (*see* **Table 1**) to ensure adequate sequencing coverage of the ECM14 gene insert. Only the sequence and chromatograms of the 28-30 bases around mutagenic site are shown. (**A**) Wild-type pBluescript-ECM14 was used as a "**before SDM**" control. (**B**, **C**) The resultant "**after SDM**" mutagenized ECM14 sequence. The codon of interest is annotated in both sequences and the resultant amino acid identified. Codons are reverse-complement due to the reverse T7 primer used (*****). Sanger sequencing was performed by GenScript.

Three minus orientations of each of the four mutant pSLS1-ECM14 plasmids were miniprepped using the Qiagen QIAprep Spin Miniprep Kit (H69A_1, H69A_3, H69A_6; H145A_13, H145A_14, H145A_16; K270A_1, K270A_5, K270A_6; and K270E_6, K270E_7, K270E_11). The mutated site in each plasmid was Sanger sequenced by Genscript using only a single primer. The ECM14_Internal primer was used to sequence the H69A and H145A sites; and the ECM14_D primer was used to sequence the K270A and K270E sites (*see* ***Table 1***). Because the high-fidelity Platinum SuperFi DNA Polymerase was used to perform the original site-directed mutagenesis in pBluescript-ECM14, it was not necessary to re-sequence the entire mutant ECM14 gene. However, the purpose of this final round of Sanger sequencing was simply to validate the subcloning of each mutant ECM14 gene from pBluescript into pSLS1 (*see* ***Table 2***).

The mutated sites in H145A_13, H145A_14, H145A_16, K270A_1, K207A_5, and K270E_11 were successfully sequenced and confirmed. However, the remaining samples were inconclusive due to failed sequencing runs that Genscript was unable to resolve by simply tweaking the sequencing conditions and re-running these failed runs. However, those successful mutant pSLS1-ECM14 sequences, along with the successful mutant pBluescript-ECM14 sequences, provided confidence that the subcloning of each mutant ECM14 sequence from pBluescript to pSLS1 was successful.

CHAPTER 4

YEAST STRAIN CONSTRUCTION METHODS AND RESULTS

The Yeast Gene Deletion Strategy

The primary method for deleting genes in the yeast genome is the mechanism of chromosomal integration by homologous recombination (HR) (*see* ***Figure 28***). Because point and non-sense mutations can spontaneously revert (producing so-called *revertants*), removing the entire gene by HR is a more effective method for permanently deleting genes.

The first step in yeast gene deletion involves a polymerase chain reaction (PCR) in order to amplify and extend a plasmid-borne MX sequence with 45-bp upstream and downstream extensions flanking your favorite gene (YFG). These unique upstream and downstream extensions are critical for ensuring that HR occurs over the intended genomic interval. These MX "expression" cassettes also include the powerful, *always-on* promotor (P_{TEF}) and terminator (T_{TEF}) sequence derived from the heterologous *Ashbya gossyppi* TEF gene. Finally, subsequent diagnostic PCR reactions are performed in order to interrogate both the location and the size the MX allele in the yeast genome. In general, this approach can be used to replace any non-essential yeast gene with a variety of drug resistance (e.g. Geneticin/Kanamycin, Nourseothricin, etc.) or autotrophic (e.g. LEU2, TRP1, URA3, etc.) selection markers (Baudin et al., 1993; Wach et al., 1994).

The commercially available Yeast Knock-out (YKO) Collection was produced by selectively replacing, via this process of chromosomal integration by homologous recombination, each of the ≈5,000 non-essential genes with the Geneticin® resistance MX cassette (KanMX). The original ecm14Δ clones used in this investigation were derived from the YKO collection.

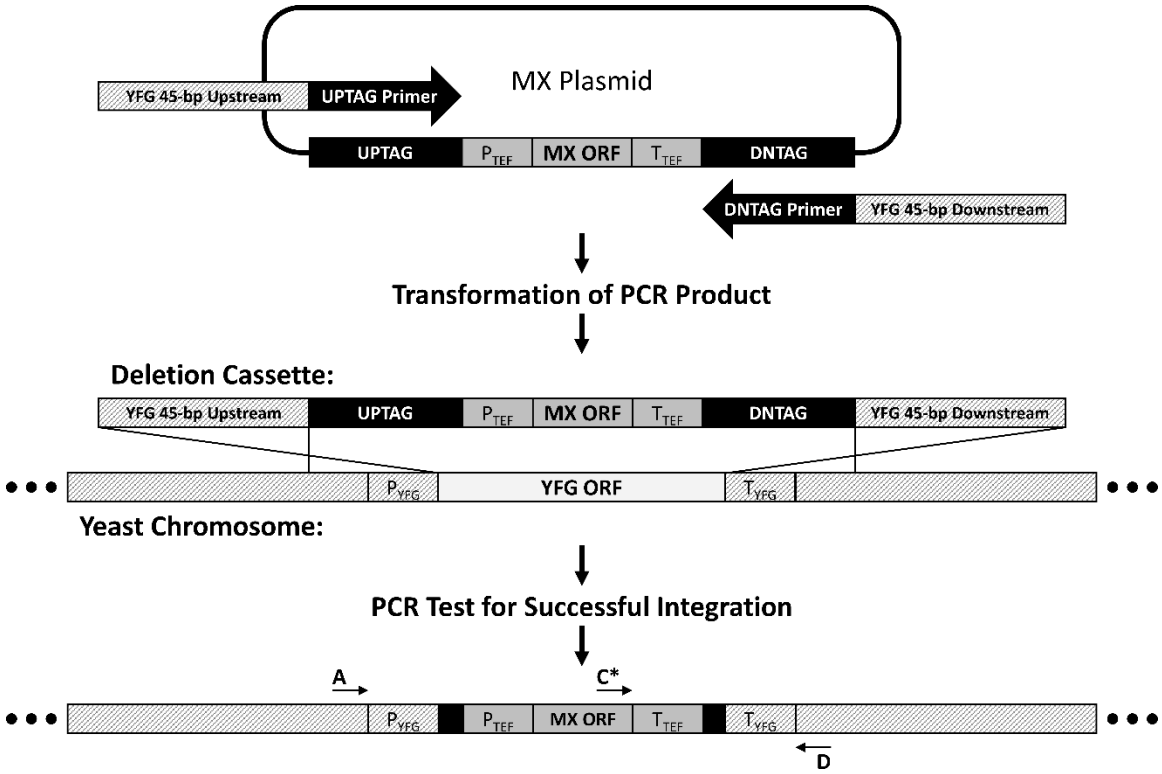

Figure 28. The yeast gene deletion strategy.

The mechanism of chromosomal integration by homologous recombination. 1) PCR is first performed in order to amplify and extend a plasmid-borne MX sequence with 45-bp upstream and downstream extensions flanking your favorite gene (YFG). These MX "expression" cassettes include the powerful, *always-on* promotor (P_{TEF}) and terminator (T_{TEF}) sequence derived from the heterologous *Ashbya gossyppi* TEF gene. 2) Yeast are then transformed with this deletion cassette. 3) Finally, PCR reactions using either the A or MX_C (*) upstream primers are performed that interrogate both the location and the size of the MX allele in the yeast genome. (DNA features are not drawn to scale.)

High-Fidelity PCR Amplification of [ecm14::KanMX] Cassette

The [ecm14::KanMX] deletion cassette was PCR amplified from an ecm14Δ knockout S288C-derived yeast strain (BY4741) obtained from the Yeast Knockout Collection (YKO) (Baker Brachmann et al., 1998). The high-fidelity Platinum SuperFi DNA Polymerase from Invitrogen was used to amplify the [ecm14::KanMX] deletion cassette. A PCR master mix was prepared for n+1 reactions (n=2) by combining 73 µl of sterile Milli-Q water, 30 µl of 5X SuperFi buffer, 15 µl of 2mM dNTP mix, 2.5 µl of ECM14_A primer, 2.5 µl of ECM14_D primer, and 2 µl of SuperFi DNA Polymerase (2 U/µl). Either 5 µl of wild-type Y1239 gDNA or ecm14Δ S288C gDNA was used as template DNA to a final volume of 50 µl per reaction. Forward and reverse primers were added to a final concentration of 0.5 µM. Each reaction tube was capped and carefully mixed before placing into a Bio-Rad T100 Thermal Cycler. Forty PCR cycles were performed using the following thermal profile: initial denaturation at 98°C for 30 seconds, thermocycling denaturation at 98°C for 10 seconds, thermocycling annealing at 63.6°C for 10 seconds, thermocycling extension at 72°C for 60 seconds, and final extension at 72°C for 5 minutes. Each PCR reaction was combined with 10 µl of New England Biolabs' Gel Loading Dye, Purple (6X) and ran in a 0.8% agarose gel (0.5 µg/ml EtBr) for visualization and gel extraction/cleanup.

Wild-type Y1239 gDNA was used as a control and produced a single, strong 1.9-kb wild-type ECM14 amplicon. The ecm14Δ knockout gDNA produced a single, strong 2.2-kb ecm14::KanMX amplicon (*see* ***Figure 29***).

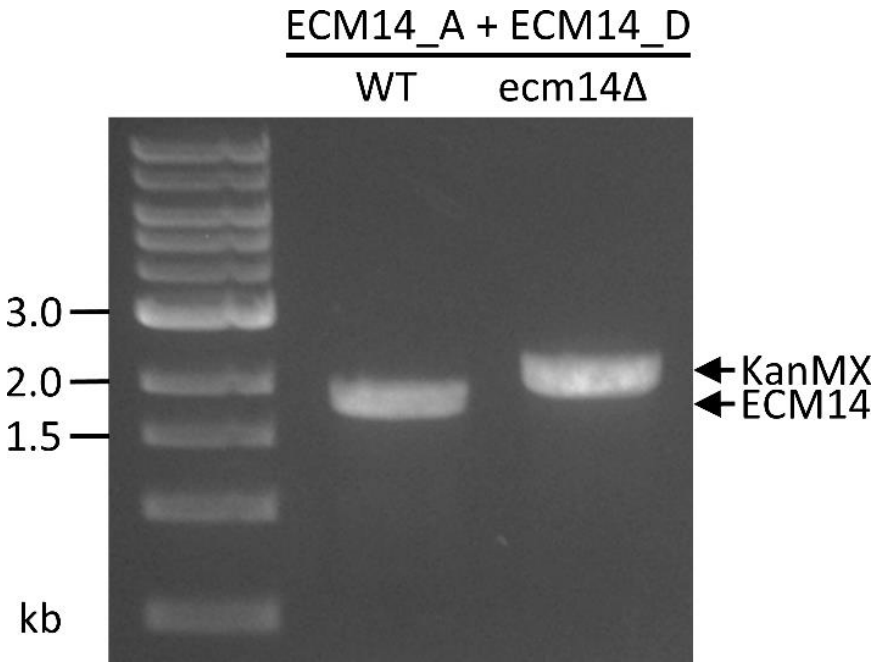

Figure 29. PCR Amplification of the ecm14::KanMX deletion cassette.

The ecm14::KanMX deletion cassette was PCR amplified from an ecm14Δ knockout S288C yeast strain obtained from the Yeast Knockout Collection (YKO). The high-fidelity Platinum SuperFi DNA Polymerase from Invitrogen was used to amplify the ecm14::KanMX deletion cassette (2.2 kb). Wild-type Y1239 yeast was also used as a control (1.9 kb).

Transformation of AF-1A/D Yeast with [ecm14::KanMX] Cassette

Fresh monoclonal colonies of either AF-1A or AF-1D (MATα and MATa, respectively) were used to inoculate 10 ml of YEPD media (*see* ***Table 4***). Liquid cultures were incubated overnight at 30°C with shaking at 250 RPM. Cultures were grown to 1-2x10^8 cells/ml, diluted to 2x10^6 cells/ml in 50 ml of fresh YEPD media, and returned to incubation at 30°C with shaking at 250 RPM for an additional 2-3 hours (≈2 cell divisions). These diluted cultures were grown to 0.7-1.5x10^7 cells/ml. Each culture was then transferred into a sterile 50-ml falcon tube and centrifuged for 5 minutes at 4,000 g. Each pellet was then decanted, resuspended in sterile water, and centrifuged for an additional 5 minutes.

Pellets were resuspended in 1 ml of filter-sterilized 1X TE/LiAc, transferred to a 1.5-ml microcentrifuge tube, centrifuged for 6 minutes at 4,000 g, and resuspended in 250 µl of 1X TE/LiAc. Double stranded salmon sperm DNA was denatured at 100°C for 5 minutes and placed on wet ice. Yeast transformation mixtures contained 5 µl of single-stranded salmon sperm DNA, 15 µl of [ecm14::KanMX] PCR product, 50 µl of yeast suspension, and 300 µl of 1X PEG/TE/LiAc (reagents were added in this order). Each transformation mixture was thoroughly vortexed, incubated at 30°C for 30 minutes, and heat-shocked at 42°C for 20 minutes. Each mixture was then briefly centrifuged at 21,130 g for 10 seconds, the PEG/TE/LiAc supernatant promptly removed, and the cells washed in 0.5 ml of sterile water. Pellets were resuspended in fresh YEPD media and incubated at 30°C for 2 hours to allow for adequate KanMX gene expression. These outgrowth cultures were then briefly centrifuged at 21,130 g for 10 seconds, the YEPD media removed, and

resuspended in 100 µl of sterile water. Finally, yeasts were plated on dry YEPD agar plates with Geneticin, incubated at 30°C for 2-3 days, and stored at 4°C.

Putative AF1-A/D [ecm14::KanMX] colonies from both transformation plates were streaked out for single colonies onto fresh YEPD agar plates with Geneticin and incubated at 30°C for 2-3 days. Putative monoclonal AF-1A/D [ecm14::KanMX] colonies were grown in 1 ml YEPD overnight to mid-log phase (1-5x10^7 cells/ml, OD_{600} = 0.4-1.7). Genomic DNA was extracted from 0.2 ml of each overnight culture. Frozen glycerol stocks were also made by combining 0.5 ml of each overnight culture with 0.5 ml of sterile glycerol (50% v/v) in 1-ml cryogenic storage tubes and stored at -80°C.

Low-fidelity PCR Genotyping of Putative AF-1A/D ecm14Δ Clones

Putative AF1-A/D [ecm14::KanMX] clones were screened for both the presence of the KanMX ORF and the absence of the ECM14 ORF using the low-fidelty DreamTaq™ DNA polymerase from Thermo Fisher Scientific. Either the upstream ECM14_A or Kan_C upstream primer was used with the downstream ECM14_D primer (*see* ***Table 1***). The Kan_C primer is specific for the KanMX ORF and is expected to produce an 888-bp amplicon when used with the downstream ECM14_D primer. However, the ECM14_A and ECM14_D primers are expected to produce either an 1,894-bp amplicon in wild-type yeast or an 2,185-bp amplicon in ecm14Δ yeast.

A PCR master mix was prepared for n+1 reactions (n=7) by combining 251.5 µl of sterile Milli-Q water, 35 µl of 10X DreamTaq buffer, 35 µl of 2mM dNTP mix, 8.75 µl of Kan_C primer, 8.75 µl of ECM14_D primer, and 3 µl of DreamTaq DNA polymerase (5 U/µl). An additional PCR master mix was prepared with ECM14_A and ECM14_D primers. Either 1 µl of gDNA from AF-1A_3, AF-1D_5, AF-1D_6, or AF-1D_7 was used as template DNA. Additionally, either 1 µl of gDNA from wild-type Y1239 or ecm14Δ S288C were used as controls. Forward and reverse primers were added to a final concentration of 0.2 µM. Each reaction tube was capped and carefully mixed before placing into a Bio-Rad T100 Thermal Cycler. Forty PCR cycles were performed using the following thermal profile: initial denaturation at 95°C for 30 seconds, thermocycling denaturation at 95°C for 30 seconds, thermocycling annealing at 57°C for 30 seconds, thermocycling extension at 72°C for 60 seconds, and final extension at 72°C for 5 minutes. Each PCR reaction was combined with 10 µl of New England Biolabs' Gel Loading Dye, Purple (6X) and ran in a 0.8% agarose gel (0.5 µg/ml EtBr) (*see* ***Figure 30***).

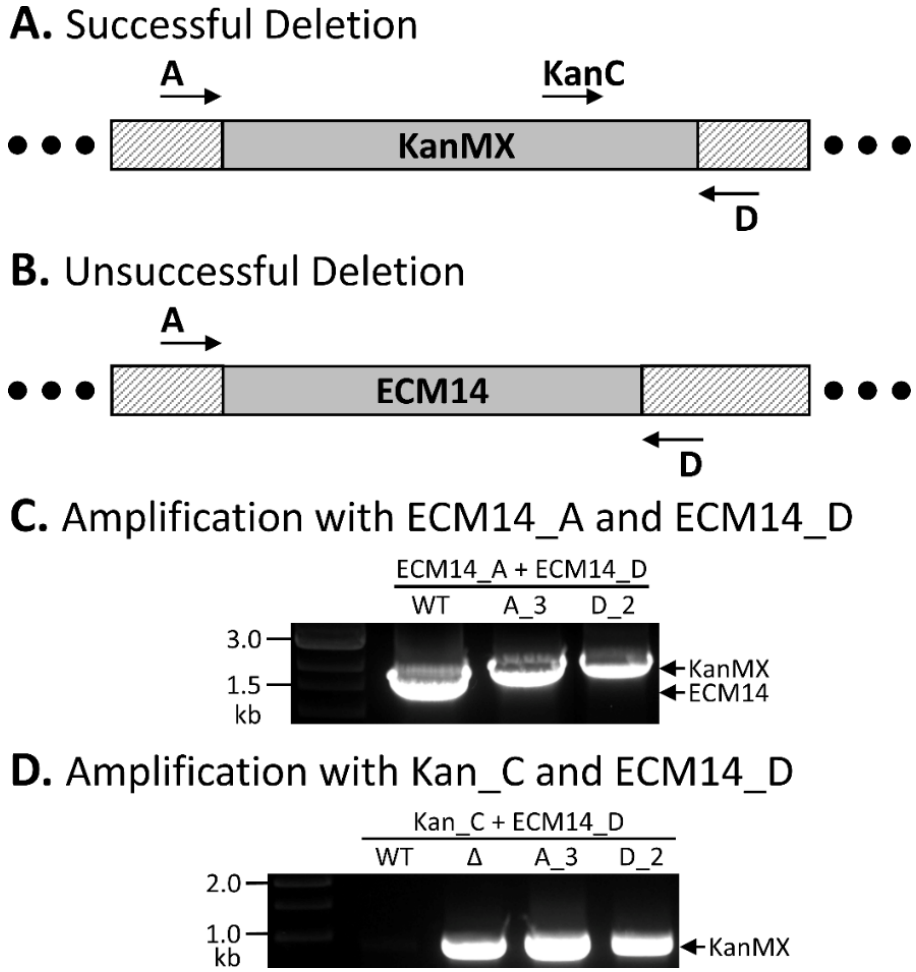

Figure 30. Knock-out yeast strains were constructed using the ecm14::KanMX deletion cassette and were verified using low-fidelity PCR.

Both mating types (AF-1A and AF1-D) were transformed using the [ecm14::KanMX] deletion cassette and verified using PCR reactions that interrogate both the size and the location of the KanMX allele. (**A**, **B**) Successful ECM14 deletion can be verified by screening for a 2.2-kb KanMX amplicon or a 1.9-kb ECM14 amplicon (**C**). Successful ECM14 deletion can also be verified by screening for a 0.9-kb Kan_C + ECM14_D amplicon (**A**, **D**). The YKO ecm14 knockout (**Δ**) was used as a Kan_C positive control. Wild-type Y1239 lack the KanMX allele (**B**, **D**).

Those successful AF-1A/D [ecm14::KanMX] clones were streaked for single colonies from glycerol stocks onto YEPD agar plates with Geneticin and incubated at 30°C for 2-3 days for future transformation reactions. The successful AF-1A_3 and AF-1D_2 clones were renamed to CM-1A_3 and CM-1D_2 respectively (*see* ***Table 4***).

Transformation of CM-1A/D ecm14Δ Yeast with pSLS1-ECM14

Fresh monoclonal colonies of either CM-1A_3 [ecm14::KanMX] or CM-1D_2 [ecm14::KanMX] (MATα and MATa, respectively) were used to inoculate 10 ml of YEPD media. Liquid cultures were incubated overnight at 30°C with shaking at 250 RPM. Cultures were grown to 1-2x10^8 cells/ml, diluted to 2x10^6 cells/ml in 50 ml of fresh YEPD media, and returned to incubation at 30°C with shaking at 250 RPM for an additional 2-3 hours (≈2 cell divisions). Diluted cultures were grown to 0.7-1.5x10^7 cells/ml. Each culture was transferred into a sterile 50-ml falcon tube and centrifuged for 5 minutes at 4,000 g. Each pellet was then decanted, resuspended in sterile water, and centrifuged for an additional 5 minutes.

Pellets were resuspended in 1 ml of filter-sterilized 1X TE/LiAc, transferred to a 1.5-ml microcentrifuge tube, centrifuged for 6 minutes at 4,000 g, and resuspended in 250 µl of 1X TE/LiAc. Double stranded salmon sperm DNA was denatured at 100°C for 5 minutes and placed on wet ice. Yeast transformation mixtures contained 5 µl of single-stranded salmon sperm DNA, 2 µl of either orientation of pSLS1-ECM14, 50 µl of yeast suspension, and 300 µl of 1X PEG/TE/LiAc (reagents were added in this order). Each transformation mixture was thoroughly vortexed, incubated at 30°C for 30 minutes, and heat-shocked at 42°C for 20 minutes. Each mixture was then briefly centrifuged at 21,130 g for 10 seconds, the PEG/TE/LiAc supernatant promptly removed, and washed in 0.5 ml of sterile water. Pellets were resuspended in fresh YEPD media and incubated at 30°C for 2 hours to allow for adequate KanMX gene expression. These outgrowth cultures were then briefly centrifuged at 21,130 g for 10 seconds, the YEPD media removed, and resuspended in 100 µl of sterile water.

Outgrowth cultures were plated on dry SC-Ura agar plates, incubated at 30°C for 3-4 days, and stored at 4°C. Minimal SC media lacking uracil (SC-Ura) was used to positively select for the pSLS1-ECM14 plasmid. Additionally, growth on both YEPD and YPGal were observed to demonstrate destabilization and loss of the plasmid when cultured under the influence of galactose (*see* ***Figure 31***).

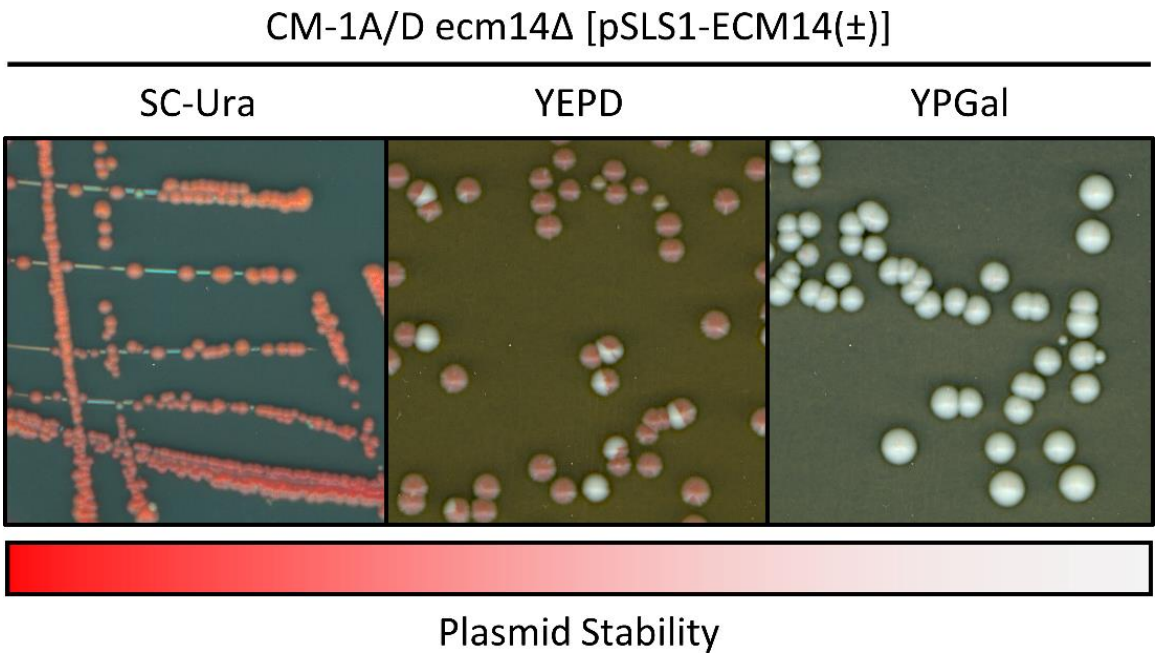

Figure 31. Growth of CM-1A/D ecm14Δ [pSLS1-ECM14(±)] on various growth media to demonstrate plasmid transformation and plasmid destabilization.

CM1-A/D ecm14Δ transformed with pSLS1-ECM14(±) was grown on SC-Ura, YEPD, and YPGal. Growth on SC-Ura positively selects for the pSLS1-ECM14(±) plasmid. Growth on YEPD and YPGal demonstrate plasmid destabilization prior to subsequent mutagenesis experiments. Growth on YPGal effectively destabilizes the plasmid and results in white colonies.

CHAPTER 5

THE SYNTHETIC LETHAL SCREEN METHODS AND RESULTS

EMS Mutagenesis in Yeast

Fresh monoclonal colonies of either CM-1A_2 [pSLS1-ECM14, ecm14::KanMX] or CM-1D_3 [pSLS1-ECM14, ecm14::KanMX] (MATα and MATa, respectively) were used to inoculate 10 ml of SC-Ura media. Routine growth of these pre-mutagenesis strains was performed in/on SC-Ura to ensure strong selective pressure for the pSLS1-ECM14 plasmid (*see above* ***Figure 31***). Overnight cultures were adjusted to $2x10^8$ cells/ml and 1 ml of each culture was pelleted in a clean 1.5-ml microcentrifuge at 21,130g for 10 seconds. Pellets were washed twice in 1 ml of sterile water. After the second washing, pellets were resuspended in 1.5 ml of sterile sodium phosphate buffer (0.1M, pH 7.0). EMS mutagenesis was performed by combining 0.7 ml of cell suspension, 1 ml of sodium phosphate buffer, and 50 µl of EMS in a sterile 15-ml polypropylene culture tube and incubating at 30°C for 45-60 minutes. The remaining cell suspensions were saved on wet ice as a pre-mutagenesis control. Mutagenesis was halted by adding 0.2 ml of EMS treated cells to a sterile culture tube containing 8 ml of sterile sodium thiosulfate solution (5% w/v). Nonmutagenized control cells were diluted equivalently as above by substituting EMS with sterile water. Both EMS-treated and untreated cell suspensions were serially diluted to 1-3 viable cells/µl for plating on various growth media (100-300 viable cells per plate at 100 µl plated volume). The EMS mutagenesis protocol used in this study was adapted from Winston (2008); and Barbour and Xiao (2006).

Diluted EMS-treated and untreated cell suspensions were plated on various growth media, incubated at 30°C for 2-4 days, and stored at 4°C for several days to facilitate

adequate color development. YPD/YEPD plates were used as controls to determine the effectiveness of EMS mutagenesis and to tune EMS exposure time in subsequent mutagenesis attempts. An ideal EMS exposure should cause ≈40% of the cells to be killed (Winston, 2008). An EMS exposure time of 45 minutes was determined to be effective at generating the ideal cell death (≈40%) in the strains used in this study (*see* ***Figure 32***).

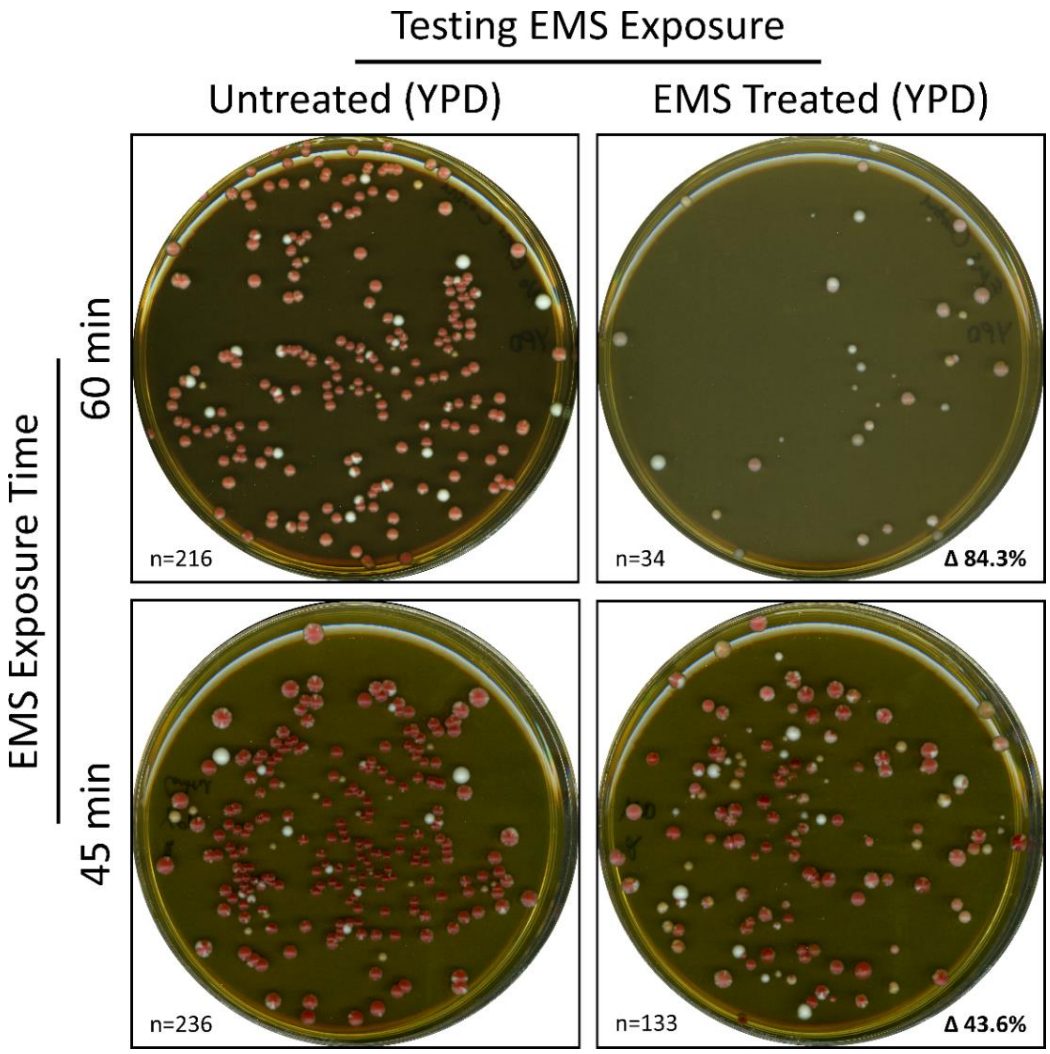

Figure 32. The effect of EMS exposure time on yeast cell viability.

EMS-treated and untreated cells were plated on YPD/YEPD media to test the effect of EMS exposure at both 60 minutes and 45 minutes. Sixty minutes of exposure was ineffective at producing enough viable yeast cells per plate (**84.3%** cell death). Reducing the EMS exposure to 45 minutes was determined to be more effective at generating a more ideal cell death (**43.6%**) in the yeast strains used in this study.

As was previously demonstrated in this study, growth on YPGal was used to effectively destabilize the pSLS1-ECM14 plasmid and to screen for novel plasmid dependencies (*see above* ***Figure 31***). However, because the pSLS1-ECM14 plasmid contains two additional yeast genes (*ADE3* and *URA3*) (*see above* ***Figure 11***), this approach is fundamentally unable to immediately identity false-positive synthetic lethal situations caused by a novel dependency on either the plasmid-borne *ADE3* or *URA3* genes.

Synthetic lethal situations likely require that all genes involved are transcriptionally active. Therefore, effectively silencing either the *ADE3* or *URA3* gene could potentially improve the sensitivity of the synthetic lethal screen and decrease the number of false-positives to be subsequently evaluated. Because an active, plasmid-borne *ADE3* gene is fundamentally required for the *ADE3*/*ADE2* color-colony assay essential to the synthetic lethal screen, any potential modifications to decrease *ADE3* false-positives were not immediately available for use in this study. However, because the *URA3* gene is not (and should not be) involved in the synthetic lethal screen and any subsequent color-colony assays, modifications were made in later mutagenesis attempts in order to potentially decrease these undesirable *URA3* false-positives. Because the *URA3* gene is effectively silenced when grown on media containing excess uracil, a modified YPGal media containing extra uracil (5 ml of sterile, 0.2% uracil per liter of media) was also used for screening EMS-treated cells.

In addition to the modification to potentially decrease *URA3* false-positives, a modified YPGal media containing 20 μg/ml, filter-sterilized calcofluor white (1000X stock prepared at 20mg/ml) was also used for screening EMS-treated cells. Because exposure to calcofluor white (CFW) amplifies the effect of mutations in genes involved in cell wall

architecture and remodeling (Lussier et al., 1997; Ram et al., 1994), screening these EMS-treated cells on media containing CFW was thought to increase the sensitivity of the synthetic lethal screen to resolving novel mutations in genes that are functionally related to ECM14. The concentration of CFW used in this modified YPGal media was based on the standard concentration used to generate CFW hypersensitive mutants by Lussier et al. (1997). Finally, a combination YPGal media containing both uracil and CFW was used in the later mutagenesis attempts in order to further encourage the generation of true ECM14 synthetic lethal colonies. In either case, plates containing EMS-treated cells were screened for non-sectored, red colonies of interest to be further studied (*see* **Figure *33***).

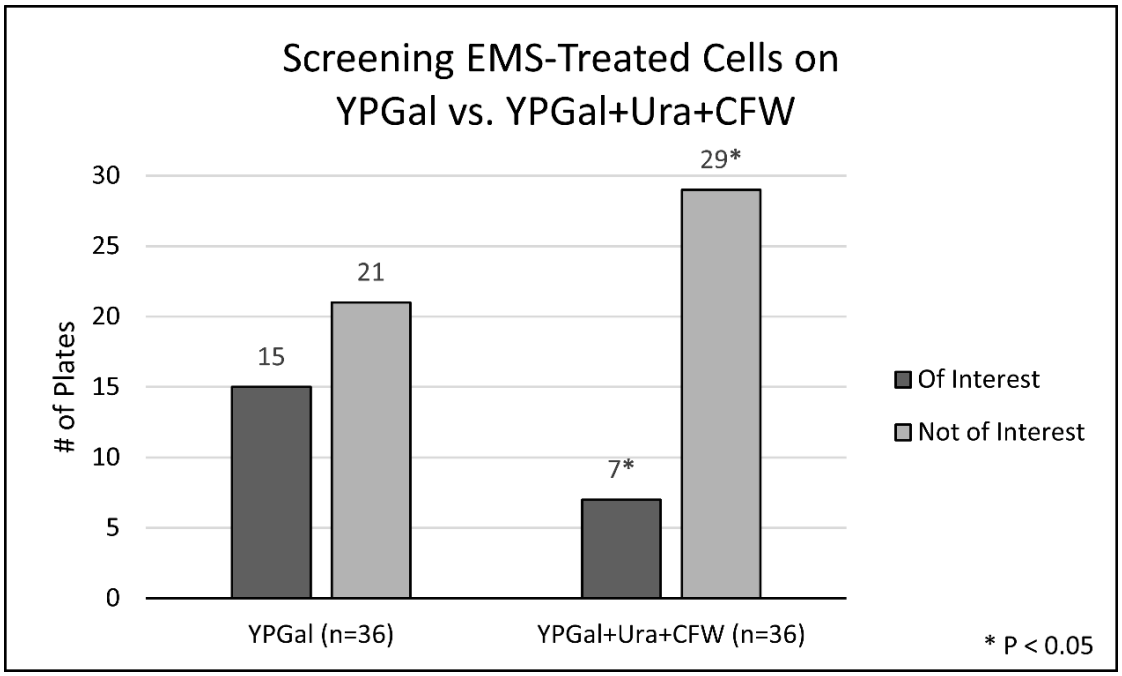

Figure 33. Screening EMS-treated cells on YPGal vs. YPGal+Ura+CFW.

YPGal was used to effectively destabilize the pSLS1-ECM14 plasmid and to screen for novel plasmid dependencies. **YPGal+Ura+CFW** was used to discourage the generation of *URA3* false-positives and to place additional pressure on genes thought to be functionally related to ECM14. All screening plates represented were plated with the same number of cells at the same time. Plates containing colonies of interest were those that contained at least one non-sectored red colony to be further evaluated as a putative synthetic lethal.

Screening Putative ECM14 Synthetic Lethal Mutants

Approximately 27,000 EMS-treated CM-1D ecm14Δ [pSLS1-ECM14(-)] yeast colonies were screened for novel plasmid dependencies on galactose media (either YPGal, YPGal+CFW, or YPGal+Ura+CFW). Three significant rounds of EMS mutagenesis were performed, with each yielding approximately 10,000 screenable EMS-treated yeast colonies. Colonies of interest were those that initially produced the desired non-sectored red phenotype on these media. These non-sectored red colonies of interest were noted and re-streaked for single colonies on fresh media to further evaluate the development and the maintenance of the desired non-sectored red phenotype. Colonies that initially appeared as white or sectored on galactose media were those that likely did not contain synthetic lethal mutations and, thus, were not considered for further evaluation.

As recommended by Barbour and Xiao (2006), additional rounds of growth tests on both SC-Ura and SC+5FOA were attempted in order to test for true synthetic lethal mutants caused by novel secondary mutations or for potential false-positives caused by reversions of marker genes. True synthetic lethal mutants fundamentally require the plasmid for survival. Therefore, counterselection against the plasmid-borne URA3 gene using 5-FOA should result in plasmid loss and, thus, cell death for true synthetic lethal mutants. This prescribed test was performed on the first five putative synthetic lethal mutants (SL1-5) produced from the first significant round of EMS mutagenesis performed.

A series of spotting assays was performed using serial 1:10 dilutions of these first five putative synthetic lethal mutants (SL1-5), one EMS-treated white non-synthetic lethal colony, and one untreated control strain. Although these five putative synthetic lethal mutants seemed to retain the desired phenotype when subsequently re-plated/spotted on

YPGal, all five were unable to appreciably grow on any SC-based media. However, because the EMS-treated white non-synthetic lethal colony (CM-1D ecm14Δ) and untreated control strain (CM-1D ecm14Δ [pSLS1-ECM14(-)]) were able to grow on SC-based media, this apparent lack of growth of these synthetic lethal mutants seemed to have nothing to do with either the presence or the absence of the ECM14 gene (*see* ***Figure 34***). Instead, this lack of growth seemed to be the result of the secondary mutations in these putative synthetic lethal mutants conferring significant fitness deficiencies on these SC-based media. Therefore, these apparent fitness deficiencies suggest that counterselection of the pSLS1-ECM14 plasmid using SC+5FOA is likely not an informative diagnostic when assessing the quality of the putative synthetic lethal mutants produced in this study.

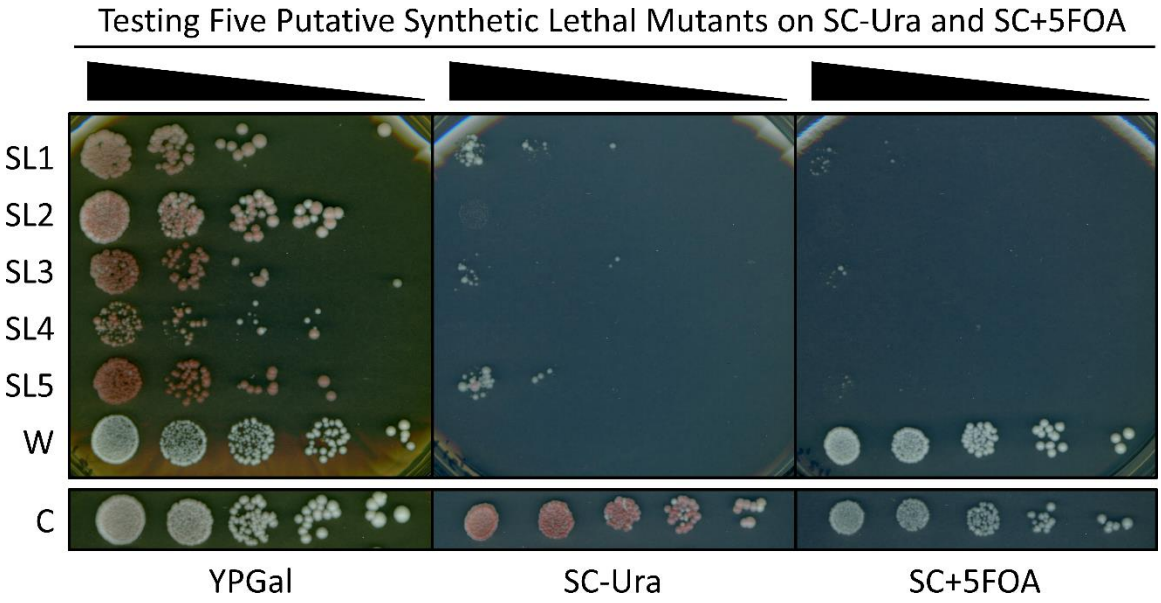

Figure 34. Testing five putative synthetic lethal mutants on SC-Ura and SC+5FOA.

Spotting assay on YPGal, SC-Ura, and SC+5FOA. Serial 1:10 dilutions of five putative synthetic lethal mutants (**SL1-5**), one EMS-treated white, non-synthetic lethal colony (**W**), and one untreated control strain (**C**). Plates were incubated at 30°C for ≈120 hours. Putative synthetic lethal mutants were derived from EMS-treated CM-1D ecm14Δ [pSLS1-ECM14(-)] colonies initially screened on YPGal.

All putative synthetic lethal mutants were re-streaked for single colonies on galactose media to further evaluate the development and maintenance of the desired phenotype. Those initial white non-synthetic lethal colonies always remained white when re-streaked on galactose media. However, many colonies of interest that initially produced the desired non-sectored red phenotype on these media following EMS treatment produced a variety of patterns (so-called *mosaics*) when re-streaked on galactose media (*see* ***Figure 35***).

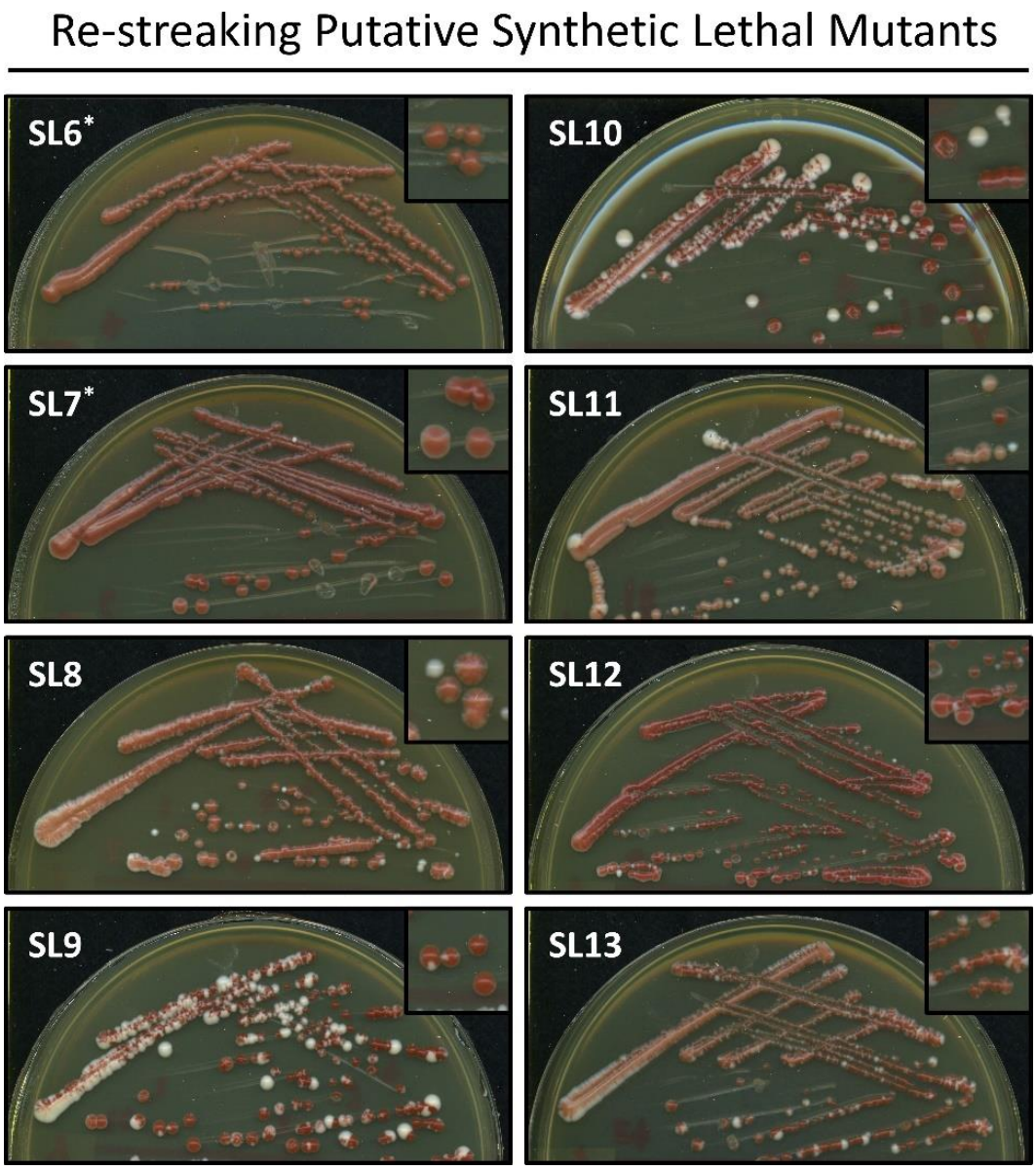

Figure 35. Re-streaking putative synthetic lethal mutants on galactose media.

Putative synthetic lethal mutants that initially produced the desired non-sectored red phenotype were re-streaked for single colonies on galactose media to evaluate the development and maintenance of the red phenotype. (*) Colonies that were screened on media containing Ura+CFW. A variety of patterns were observed.

Non-sectored red colonies from these re-streaked galactose plates were used to inoculate 2 ml of YPD liquid culture. These cultures were incubated overnight at 30°C with shaking at 250 RPM. Frozen glycerol stocks were prepared by combining 500 µl of 50% (v/v) sterile glycerol and 500 µl of overnight culture in a cryovial tube and were stored at -80°C. Several non-sectored red colonies were used to prepare frozen glycerol stocks for multiple clones from the same re-streaked plate (*see above* **Figure 35**).

The first round of EMS mutagenesis yielded **five** putative synthetic lethal mutants (**SL1-5**) derived from YPGal. The second round yielded **two** putative synthetic lethal mutants derived from YPGal+Ura+CFW (**SL6-7**) and **six** from YPGal (**SL8-13**). [**SL14** was discovered accidentally after rechecking EMS-screening plates that had been stored at 4°C for many weeks.] The final round yielded **eight** putative synthetic lethal mutants derived from YPGal+Ura and YPGal+Ura+CFW (**SL15-22**). Due to both timing and logistical constraints, only those eight putative synthetic lethal mutants from the second round (**SL6-13+14**) were able to be considered for future whole-genome sequencing and bioinformatic analysis. Three clones from each of these eight putative synthetic lethal mutants were re-streaked in mass on YPD, YPGal+Ura, and YPGal+Ura+CFW to further evaluate the development and maintenance of the desired phenotype prior to selection for whole-genome sequencing (*see* ***Figure 36*** *and* ***Figure 37***). Seven putative synthetic lethal mutants (SL6, SL7, SL8, SL11, SL12, SL13, and SL14) were selected for whole-genome sequencing Those remaining eight putative synthetic lethal mutants (**SL15-22**) were noted and stored for any future work relating to ECM14 using this synthetic lethal system.

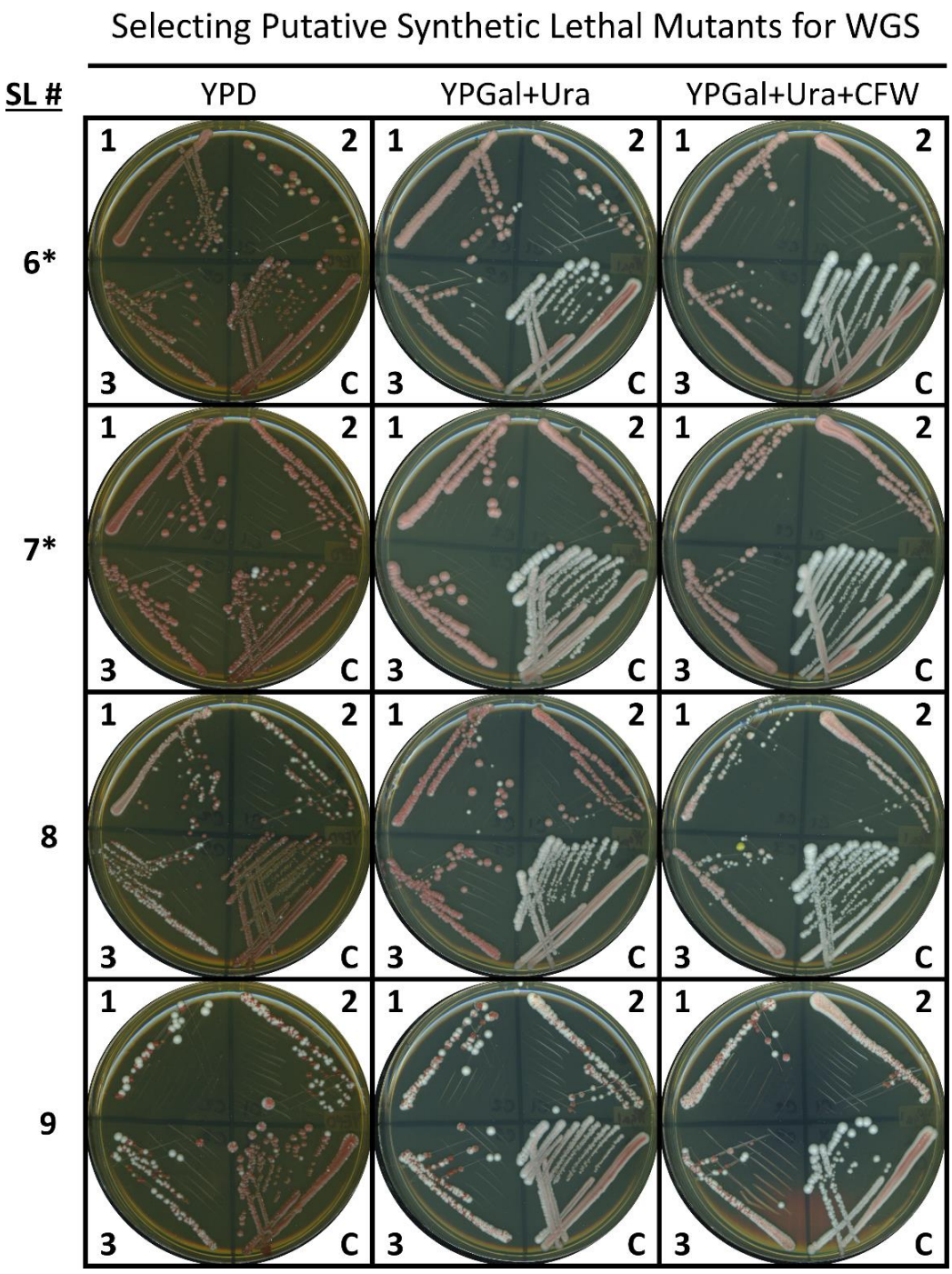

Figure 36. Re-streaking putative synthetic lethal mutants (SL6-9) for the selection of candidates for whole-genome sequencing.

Three clones (**1-3**) of each putative synthetic lethal mutant (SL6-9) were re-streaked on YPD, YPGal+Ura, and YPGal+Ura+CFW to further evaluate the development and maintenance of the desired phenotype prior to selection for whole-genome sequencing. An additional untreated control strain (**C**) (CM-1D ecm14Δ [pSLS1-ECM14(-)]) was also streaked for comparison.

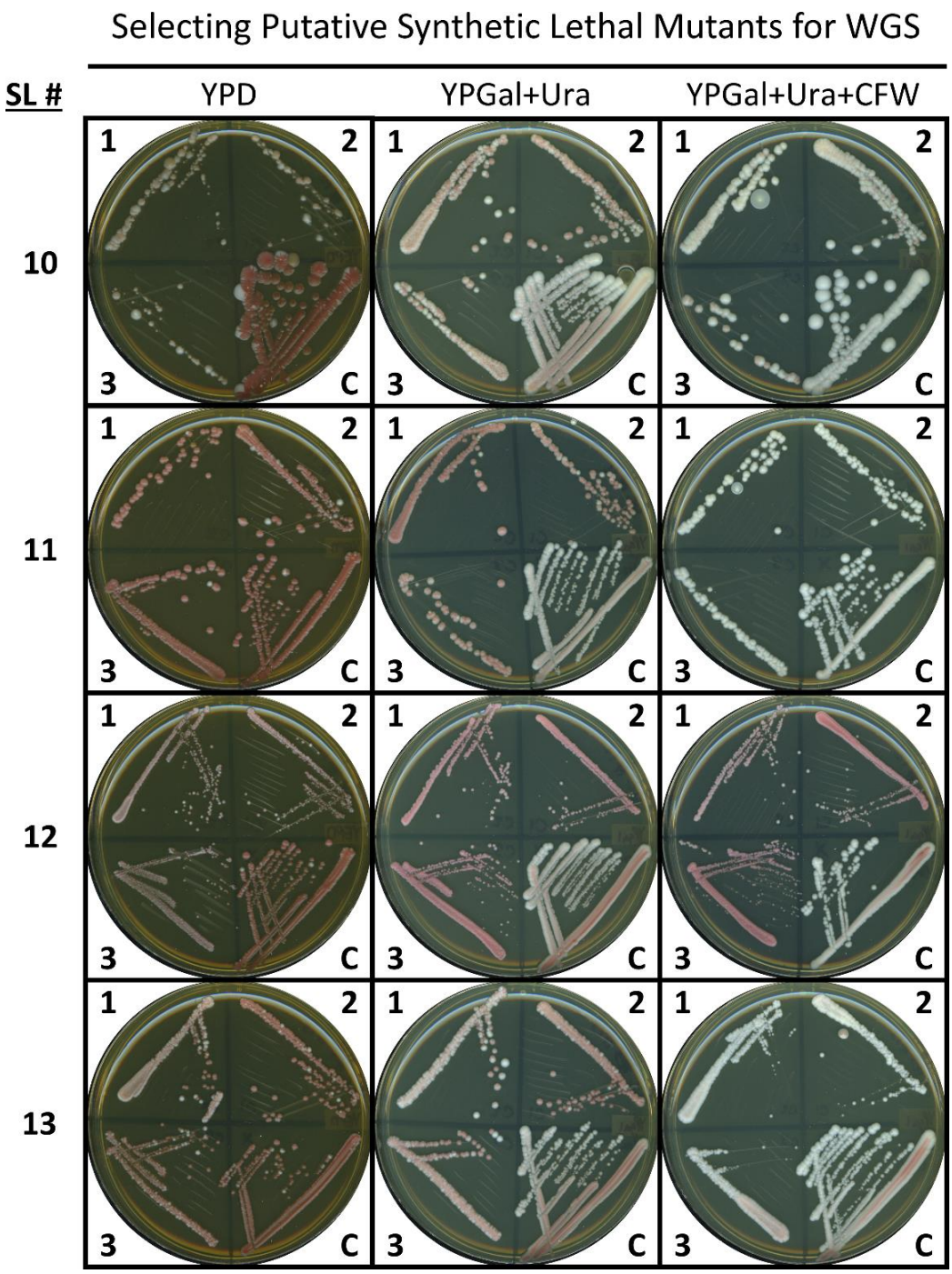

Figure 37. Re-streaking putative synthetic lethal mutants (SL10-13) for the selection of candidates for whole-genome sequencing.

Three clones (**1-3**) of each putative synthetic lethal mutant (SL10-13) were re-streaked on YPD, YPGal+Ura, and YPGal+Ura+CFW to further evaluate the development and maintenance of the desired phenotype prior to selection for whole-genome sequencing. An additional untreated control strain (**C**) (CM-1D ecm14Δ [pSLS1-ECM14(-)]) was also streaked for comparison.

CHAPTER 6

BIOINFORMATICS METHODS AND RESULTS

Whole-Genome Sequencing of Seven Putative Synthetic Lethal Yeast Strains

Fresh monoclonal colonies from seven putative synthetic lethal mutants (SL6, SL7, SL8, SL11, SL12, SL13, and SL14) and one wild-type, non-mutagenized yeast strain (all derived from CM-1D_3[ECM14(-)], *see* ***Table 4***), were cultured overnight at 30°C with shaking in 3 ml of YEPD supplemented with 25 µl of 0.5% (w/v) filter-sterilized adenine solution. Cultures were grown to saturation. Each overnight culture was entirely pelleted in a clean 1.5-ml microcentrifuge tube and washed twice with sterile water. Each of the eight microcentrifuge tubes containing yeast cell pellets were thoroughly sealed with Parafilm and shipped on ice overnight to The Sequencing Center in Fort Collins, Colorado.

Next-generation whole-genome sequencing of all eight samples was performed by The Sequencing Center (TSC) using the Illumina MiniSeq® platform with a targeted mean coverage depth of ≈40-50× (*see* ***Table 5***). The raw forward and reverse sequencing reads for each sample were returned from The Sequencing Center in the ubiquitous FASTQ file format for future bioinformatic analyses. [The FASTQ file format is the *de facto* standard for storing the output of high-throughput, next-generation sequencers. The FASTQ format stores both the biological sequence and the corresponding quality score for each called sequence letter.] The forward and reverse reads for each sample were paired to form contiguous reads using Geneious Prime 2019.2.3 (Kearse et al., 2012).

W303 Genome Annotation and Reference-guided Genome Assembly

The *Saccharomyces cerevisiae* yeast strains constructed in this study were based on the W303 lineage. Therefore, an annotated W303 genome was desired. [Genome annotation is the process of identifying various genomic elements (e.g. introns, exons, regulatory sequences, coding regions, etc.) and linking useful biological metadata to these elements.] Although S288C and W303 are closely related, they differ enough to render comparative analyses and variant calling inordinately noisy and problematic. Ralser et al. (2012) reported that the W303 and S288C genomes differ in 8,133 nucleotide positions and 799 open reading frames. Despite these notable differences and some telomeric rearrangements (notably found on chromosome XI, XIII, and XVI), the overall landscape and structure of the S288C and W303 genomes are very similar (Matheson et al., 2017).

The annotated S288C genome was downloaded from the NCBI Reference Sequence (RefSeq) database (Release 97, November 2019). The W303 genome has been thoroughly sequenced (Goodwin et al., 2015; Lang et al., 2013; Matheson et al., 2017; Ralser et al., 2012; Song et al., 2015). Recent high quality PacBio® W303 sequencing data (47.0× coverage) from Matheson et al. (2017) were downloaded from the NCBI GenBank database (LYZE01000001-21). Genome annotations were transferred from each S288C chromosome to their corresponding W303 chromosome using the BLAST-like "Annotate From…" alignment tool in Geneious Prime 2019.2.3 (*see* ***Figure 38***). In order to decrease the computational time required to align all S288C annotation sequences against the W303 genome, each W303 chromosome was annotated using only its corresponding S288C chromosome as the source database (instead of the entire S288C genome as the source database). Each annotated S288C chromosome was placed in separate folders that were

individually selected as the "Source" database when annotating each corresponding W303 chromosome. Annotations were matched using a generous similarity index of 80%. Otherwise, the default "Annotate From…" tool parameters were used.

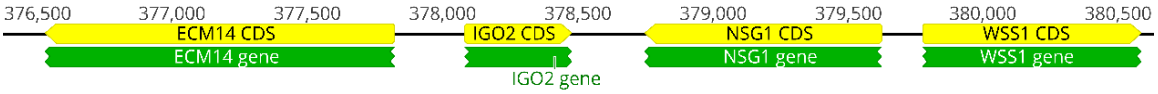

Figure 38. An annotated portion of W303 chromosome VIII.

An annotated portion of W303 chromosome VIII (376,500…380,500) that includes the ECM14 gene (green) and its corresponding coding sequence (yellow). The ECM14 CDS is found along the interval 377,832 to 376,531. Genome annotations from the S288C RefSeq genome were matched, transferred, and merged onto the W303 genome using Geneious Prime 2019.2.3.

Those matched annotations were subsequently merged onto the W303 genome. Because the Geneious Prime "Annotate From…" tool only transfers annotations (and their associated biological metadata), the underlying nucleotide sequences remained intact and unmodified. These annotated W303 chromosomes were combined and exported as a new Geneious sequence list representing the entire W303 genome (17 sequences: 16 chromosomes plus mitochondrial DNA) to be used in subsequent reference-guided genome assemblies and analyses.

The paired reads for each putative synthetic lethal mutant and the wild-type, non-mutagenized control yeast strain were mapped against this annotated W303 genome using the Geneious Prime "Map to Reference" tool to generate whole-chromosome assemblies for each sample (*see* ***Table 5***). The default "Map to Reference" tool parameters were used (Mapper: **Geneious Assembler**; Sensitivity: **Medium-Low Sensitivity / Fast**; and Fine Tuning: **Iterate up to 5 times**).

Variant Calling, Filtering, and Analysis

Genetic variants were called on each chromosome from each sequenced sample using the Geneious Prime "Annotate & Predict > Find Variations/SNPs…" tool (Minimum Coverage: **3**; Minimum Variant Frequency: **0.8**; Maximum Variant P-Value: $\mathbf{10^{-2}}$; and Find Variants: **Inside & Outside CDS**). Because each samples' paired reads were mapped to the same reference chromosomes and, thus, linked in the Geneious software to the same underlying reference sequence list in memory, Geneious automatically created unique variant tracks for each sample that could be compared, manipulated, and visualized simultaneously on a per-chromosome basis (*see **Figure 39***). Some manual curation of these automated variant calls was performed in order to eliminate either conserved variants (as compared to the WT control sequence), variants called in uninteresting genomic regions (e.g. between coding sequences or within telomeric sequences), or variants called in regions of generally poor sequencing coverage.

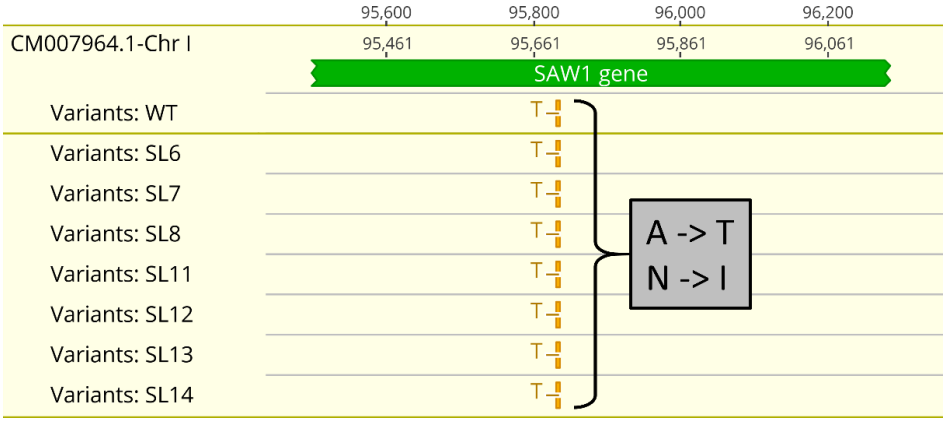

Figure 39. Visualizing variant calls across multiple samples simultaneously.

An example multi-track visualization of variant calls across multiple sequenced samples (WT, SL6, SL7, SL8, SL11, SL12, SL13 and SL14). This example highlights a small portion of chromosome I (95,300…96,300) that includes the SAW1 gene. [A conserved polymorphism was observed in this gene across all samples (interval 95,694; A -> T; N -> I; $P < 10^{-137}$).] In general, any conserved polymorphisms were noted and ultimately disregarded as true variants of interest.

The eight sequenced samples were originally mapped in error against the BY4741 genome (the YKO deletion strain derived from and closely related to S288C) by request to The Sequencing Center (TSC). Because the yeast strains used in this study were derived from the W303 strain, this initial bioinformatics work performed by TSC was fundamentally inadequate at producing any useful variant call data. However, this work did provide insights into the genomic divergence of these yeast strains from BY4741/S288C. The average number of variants per chromosome from SL6, SL7, SL8, SL11, SL12, SL13, and SL14 were calculated. Chromosome XI (with 1,739 variants) was the most divergent and chromosome XVI (with 131 variants) was the most similar (*see* ***Figure 40***). Ralser et al. (2012) reported similar results in their study on the ancestry and physiology of the W303 strain.

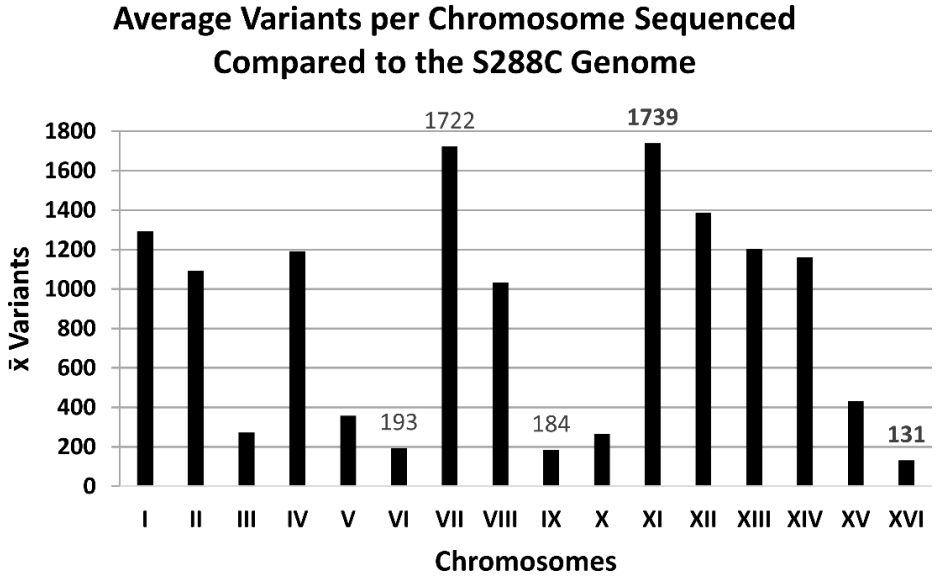

Figure 40. Comparing sequenced chromosome similarity with the S288C genome.

The average number of automated variant calls per chromosome were determined after mapping against the S288C genome. A significant number of automated variant calls (>8,000) were expected when mapping against the S288C genome. Chromosome XI (**1,739 variants**) was the most divergent and chromosome XVI (**131 variants**) was the most similar.

The seven sequenced putative synthetic lethal samples yielded a total of **594** variants of interest after manual curation of those automated variant calls when mapped against the correct reference genome (W303) (*see **Figure 41***). Mapping against the wrong genome (S288C) produced an untenable average number of variant calls (**13,659 ± 369**) (*see above **Figure 40***). However, mapping against the correct genome (W303) decreased the average number of automated variant calls significantly by ≈**18.9×** per sample (**723 ± 69**). Manual curation of these automated W303 variant calls further reduced this list (**85 ± 31**).

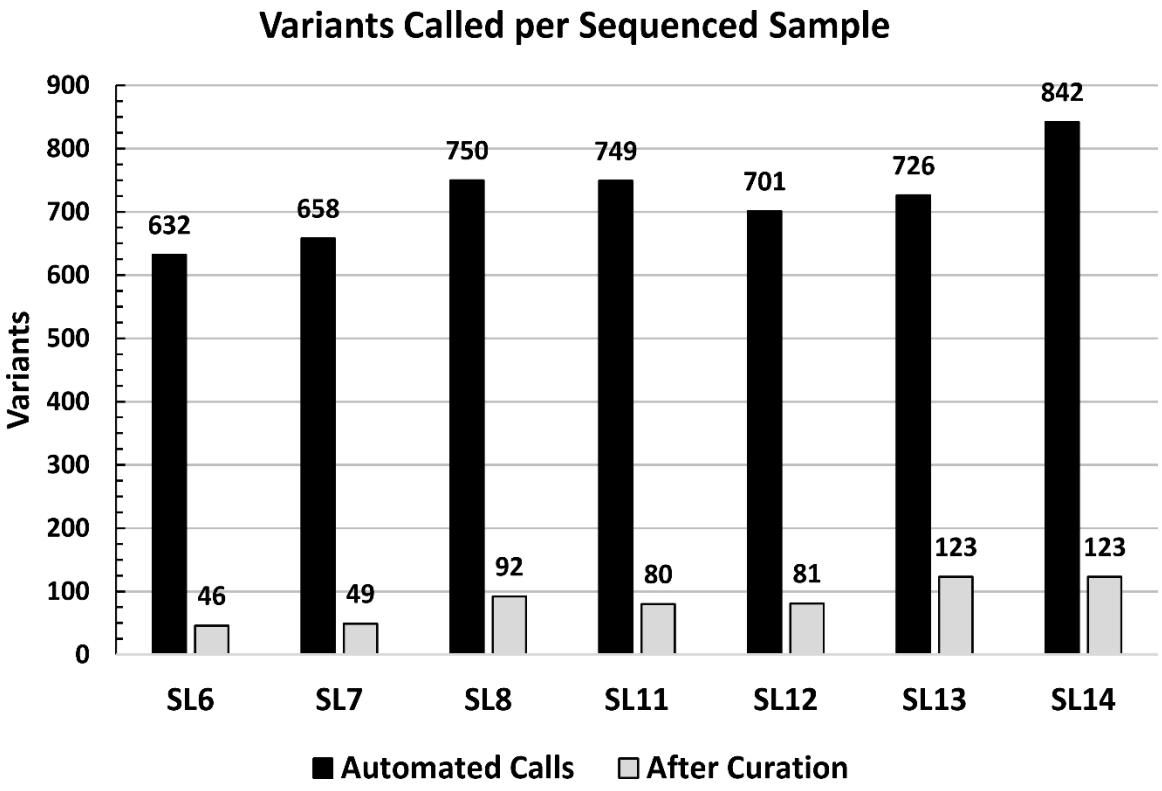

Figure 41. The number of variants called per sequenced sample.

Automated variant calls when mapped against the correct W303 genome (**723 ± 69**). Manual curation of these automated W303 variant calls yielded even fewer variants of interest per sequenced sample (**85 ± 31**). Total variants of interest identified after manual curation: **594**. **N=7**.

Mutant Protein Sequence Analysis Using PROVEAN

The list of identified variants was further evaluated using the public **Pro**tein **V**ariation **E**ffect **An**alyzer (PROVEAN) server hosted by the J. Craig Venter Institute (Choi & Chan, 2015). PROVEAN is a bioinformatics software tool for predicting whether an amino acid sequence variation impacts the underlying biological function of the mature protein. The automated PROVEAN workflow begins by first querying the NCBI non-redundant (nr) protein database for related sequences using BLASTp (e-value cutoff = 0.1) and then clustering these sequences using CD-HIT (75% global sequence identity). The top 30 CD-HIT clusters are used as the "supporting sequence set" for computing delta alignment scores for each single or multiple amino acid variation (i.e. substitutions, insertions, or deletions) for each related sequence. The final PROVEAN score is an average of the per-cluster average delta alignment scores with larger clusters being weighted more than smaller clusters.

The coding sequence (CDS) codon numbers and amino acid changes for each variant were exported from the Geneious assemblies and reformatted in the standard Human Genome Variation Society (HGVS) nomenclature (e.g. V251I, A174V, etc.). The wild-type amino acid sequence for each gene was obtained from SGD YeastMine and used as the query sequence for each PROVEAN job (Cherry et al., 2012). The predefined PROVEAN threshold score of -2.50 was used to initially qualify variants as either being deleterious ($\leq$-2.50) or neutral (>-2.50). Because this predefined PROVEAN threshold score was originally calibrated by Choi and Chan (2015) to resolve the maximum separation between deleterious and neutral variants in human protein variants, the distribution of PROVEAN scores of essential yeast gene variants identified in this study

was further examined. The list of essential yeast genes was obtained from SGD and used to filter the list of identified variants for only those essential genes. Finally, these essential gene mutants were visually mapped using the chromoMap R package (Anand & Rodriguez Lopez, 2020) (*see* ***Figure 42***). A total of **50** mutant essential genes were identified with PROVEAN scores less than the predefined PROVEAN baseline score (-2.50). Except for the wild-type, non-mutagenized control yeast strain (WT), all putative synthetic lethal mutants (SL6, SL7, SL8, SL11, SL12, SL13 and SL14) contained several mutant essential genes (**7 ± 3**) with PROVEAN scores at or below the baseline score (*see* ***Table 6***).

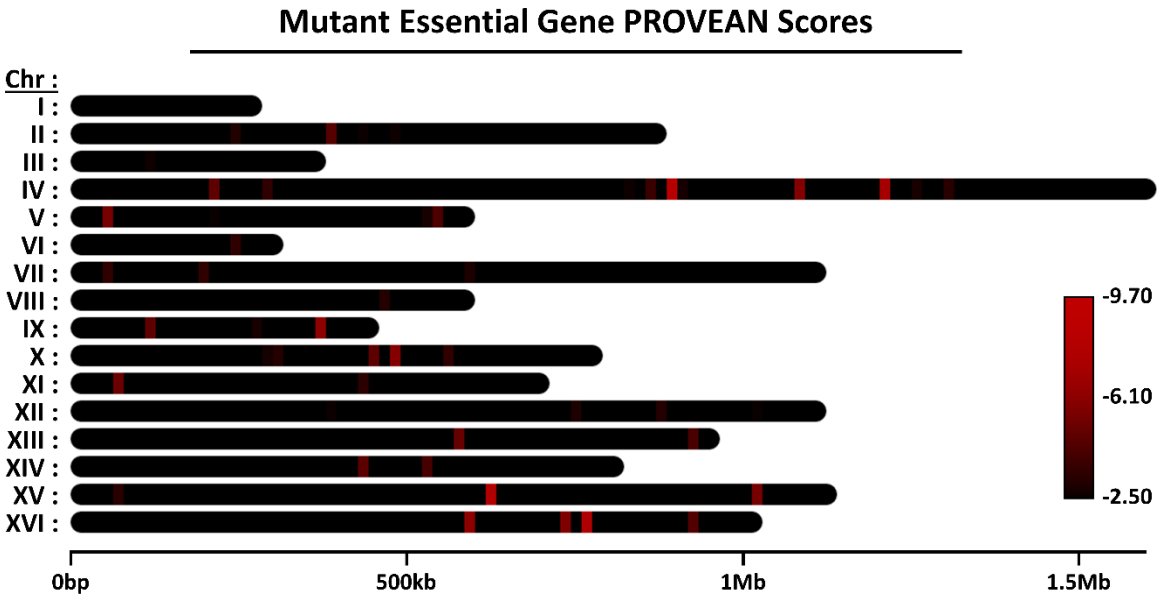

Figure 42. Mutant essential gene PROVEAN scores.

PROVEAN scores for all mutant essential genes identified in this study (N=50) were visualized as a heatmap applied to each chromosome using the chromoMap package for the R programming language (Anand, 2019). Shades of red indicates a PROVEAN score lower (i.e. more "deleterious') than the baseline score (-2.50).

Gene Ontology (GO) Enrichment Analysis

The curated lists of putative ecm14 synthetic lethal variants identified by whole-genome sequencing for each sequenced sample (SL6, SL7, SL8, SL11, SL12, SL13, and SL14) were further filtered across a range of PROVEAN thresholds starting at -2.50 and decreasing in increments of 0.75 down to a final threshold of -6.25. Both the 'GO Process' and 'GO Function' ontology aspects were queried using the *Saccharomyces* genome database (SGD) *GO Term Finder* (v0.86, July 2020) and the resulting terms merged per sequenced sample (Ashburner et al., 2000; Boyle et al., 2004). Because several essential genes were found to be mutated at or below the baseline threshold of -2.50, a more stringent PROVEAN cutoff of -4.00 was chosen for the final GO analysis (*see above* ***Figure 42****, see* ***Table 6*****)**. However, *GO Term Finder* queries were performed across the entire range of PROVEAN thresholds. Additionally, the *GO Term Finder* input query form only supports selecting an enrichment p-value threshold below 0.01, 0.05, or 0.1 (the default is 0.01). However, because this is limitation solely in the *GO Term Finder* user interface and not in the underlying data processing pipeline, simply replacing '`&pvalue=0.1`' with '`&pvalue=1`' in the resulting URL in the address bar forces the search to include all resulting GO ontology terms. These data were recorded in a master GO spreadsheet for easy filtering and sorting.

Genes annotated by the SGD as interacting either physically or genetically with *ADE3*, *ECM14*, or *URA3* were also used to query the SGD *Go Term Finder*. As before, both the 'GO Process' and 'GO Function' ontology aspects were queried and merged. These data were also recorded in the master GO spreadsheet. These additional GO analyses were performed in order to compare these interacting genes with those mutated in the

synthetic lethal assays, to further determine any likely relevance with these three plasmid-borne fungal genes (*ADE3*, *ECM14*, or *URA3*).

Each merged GO dataset was further filtered for only those GO ontology terms with enrichment p-values ≤ 0.5. This further trimmed the list of ontology terms to be considered in a single-page visualization. The resulting terms were then grouped according to their association with *ADE3*, *ECM14*, and *URA3* and then sorted within each group by ascending p-values. Any ontology terms not associated with these three genes were classified as ND, not determined, and sorted alphabetically. Finally, enrichment p-values were converted to $-LOG_{10}$ values and these values plotted as a heatmap using conditional formatting across a linear red-white, largest-to-smallest color gradient (*see* ***Figure 43***).

Little overlap was observed between GO terms for *ADE3*, *ECM14*, and *URA3*, suggesting that these genes have very distinct biochemical and biological roles. For *ECM14*, the top reported GO function and process terms, found to be enriched in this gene list compared with a list of all yeast genes ($p<0.05$), included terms such as vesicle-mediated transport (GO:0016192), vesicle fusion (GO:0006906), lipid binding (GO:0008289), and organelle organization (GO:0006996) (*see* ***Figure 43***). These terms are consistent with a secreted or secretory role for Ecm14 in cell wall maintenance.

For all putative synthetic lethal mutants except SL7 and SL13, some enrichment (although not always reaching significance, $p < 0.05$) was found for GO terms associated with *ADE3* interacting genes. Very few of the GO terms identified for the remaining synthetic lethal gene mutations reached a $p<0.05$ level of significance. In some cases, mutations were identified in genes known to be synthetic lethal with *ade3*. For example, both SL6 and SL7 presented very consistently red synthetic lethal colonies; a mutation in

the *SHM2* gene resulting in a D219N amino acid change was identified in both. Likewise, SL11 and SL14 contained mutations in the *SHM2* gene, while SL13 contained a mutation in the *SHM1* gene. Both the *SHM1* and *SHM2* genes are reported in the SGD to be synthetic lethals of the *ade3* gene.

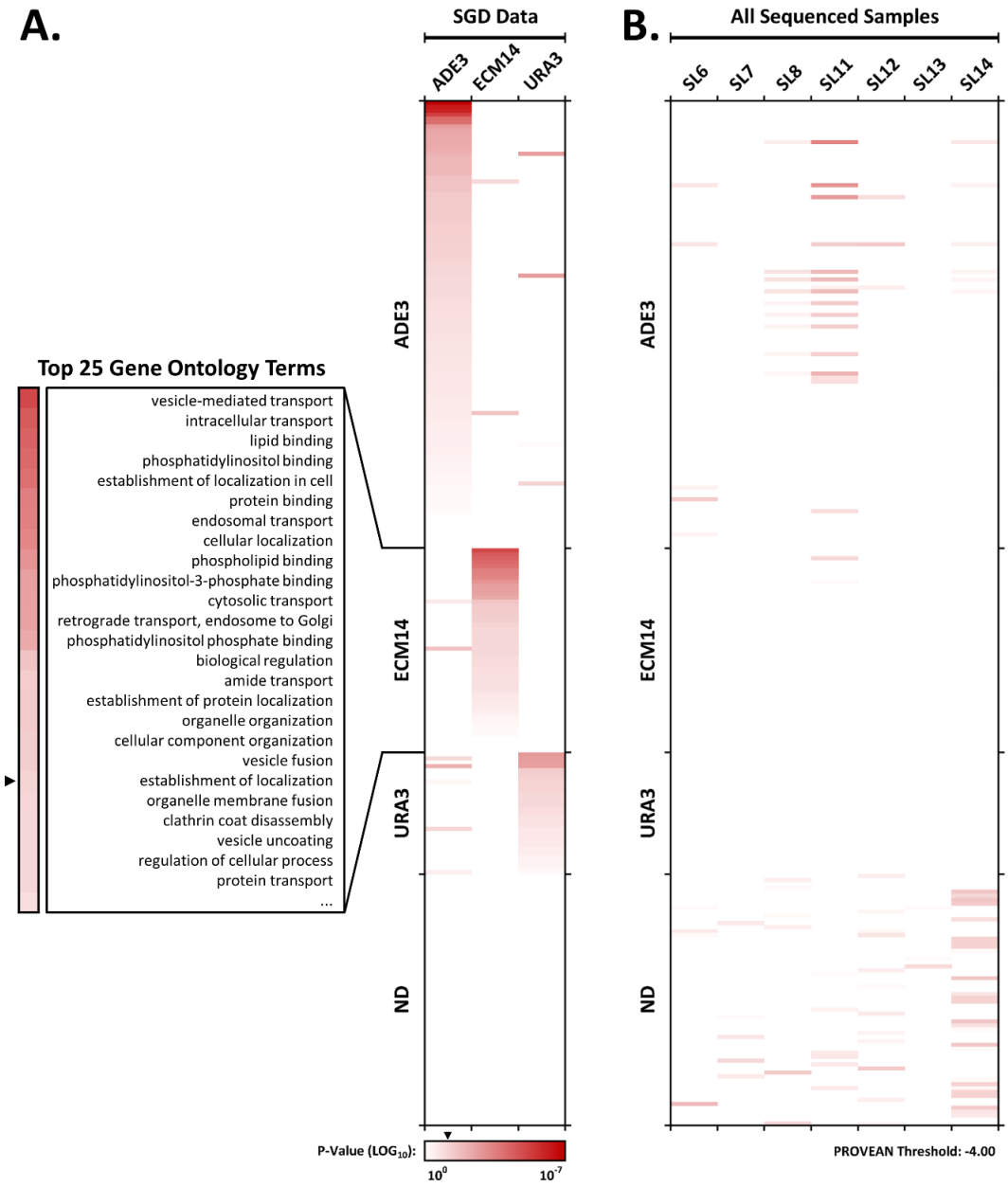

Figure 43. Gene Ontology (GO) functional enrichment analysis of interactors.

(**A**) Genes annotated by the Saccharomyces genome database (SGD) as interacting either physically or genetically with *ADE3*, *ECM14*, or *URA3* were analyzed by Gene Ontology (GO). Little overlap is observed between GO terms for these three genes. The heatmap represents P-values for enrichment of genes within particular GO classifications. Arrowhead next to heatmap scale bar indicates 0.05. Top 25 GO terms are shown for *ECM14* only. (**B**) GO analysis for genes identified by whole-genome sequencing as being mutated in the indicated synthetic lethal, and for which PROVEAN analysis surpassed a threshold of -4.0. GO terms are mapped to those identified in the SGD for *ADE3*, *ECM14*, and *URA3* interactors. All others are listed at ND, not determined.

CHAPTER 6

DISCUSSION

Fungi necessarily rely on a wide variety of hydrolases to break down and absorb the materials on which they grow. Few studies have examined carboxypeptidases expressed by fungi. Only two ascomycete metallocarboxypeptidases have been previously identified and characterized. Several publications have examined one CPA from *Metarhizium anisopliae* (Austin et al., 2011; Austin & Waugh, 2012; Joshi & St Leger, 1999), while a second related CPA from *Trichophyton rubrum* had also been partially characterized (Zaugg et al., 2008). Both enzymes exhibited typical CPA-like substrate specificity.

It is likely that Ecm14 is an inactive enzyme, or pseudoenzyme. Although it is difficult, if not impossible, to prove the absence of a function, previous studies have been unable to detect carboxypeptidase activity in *S. cerevisiae* Ecm14 using standard chromogenic assays, and this is consistent with the substitution of critical residues within the pseudoactive site of Ecm14.

ECM14 was previously identified through transposon mutagenesis as a gene providing resistance to calcofluor white and thus implicated in cell wall maintenance (Lussier et al., 1997). This ECM14 mutant, in addition to being sensitive to calcofluor white, exhibited an increased mannose:glucose ratio and was sensitive to the effects of the aminoglycoside antibiotic hygromycin B and the glycolipid papulacandin B, thought to inhibit β1,3-glucan synthesis (Lussier et al., 1997). In this study of cell wall maintenance, two yeast strains with differing sensitivity to calcofluor white were used, and the ecm14 mutant was only identified in one of these two screens. This differing sensitivity is

consistent with our observation of strain-specific sensitivities, as mild calcofluor white sensitivity was only detected in BY4741a cells grown long-term in one of two labs.

An understanding of the subcellular or extracellular location of a protein such as Ecm14 can provide useful information regarding the biological function of the protein. Although Schott (2015) was unable to identify Ecm14 in the liquid growth media of *S. cerevisiae* cells expressing Ecm14-His6, this polyhistidine-tagged Ecm14 was secreted when expressed heterologously from Sf9 insect cells. Others have shown differing results regarding localization. One large-scale analysis of yeast GFP-fusion proteins found that Ecm14-GFP was localized to the yeast vacuole (Huh et al., 2003). A proteomic analysis of *S. cerevisiae* grown in biofilm-like mats identified Ecm14 as one of many proteins found in the extracellular matrix, but not secreted into the medium of a liquid culture (Faria-Oliveira et al., 2015). These results suggest that Ecm14 likely functions within the extracellular matrix of yeast, maybe primarily under biofilm-like conditions, or possibly within the vacuole for the intracellular regulation of cell wall and extracellular matrix components.

Ecm14 has been identified in studies of several pathogenic fungi in recent years, and these studies continue to point toward a function in cell wall/extracellular matrix remodeling. A *C. albicans* screen for genes involved in aggregate invasive growth, a process related to biofilm formation, identified an ECM14 mutant showing fewer aggregates than the wild-type (Chow et al., 2019). This hints at a role in aggregate invasion, although the equivalent experiment performed with an *S. cerevisiae* ecm14Δ mutant did not result in a phenotype (Chow et al., 2019), suggesting either a false positive in the first screen or the use of inappropriate conditions for this assay with *S. cerevisiae*. A

transcriptomics study of the plant pathogen *V. dahliae* showed that the expression of Ecm14, along with a variety of cell wall degrading hydrolases, was upregulated in response to exudates from *V. dahliae*-susceptible cotton roots (Zhang et al., 2020). This again points to a possible role for Ecm14 in invasion. Finally, a recent analysis of glycosylation in the plant pathogen *M. oryzae* showed that the glycosylation of Ecm14 is upregulated primarily during mycelial growth as compared with the conidia and appressoria stages (Chen et al., 2020). It was proposed that many proteins differentially glycosylated during mycelial growth were involved in cell wall biogenesis and nutrient utilization (Chen et al., 2020). Schott (2015) found that Ecm14 expressed in *S. cerevisiae* was also glycosylated, although the function of this modification was not further examined.

The above studies indicating a role for Ecm14 in fungal growth and pathogenesis lend support to this study in which Ecm14 is shown to be highly conserved across the ascomycete lineage (*see* **Figure *45***). These aspects of fungal growth are important to understand as we try to counteract the pathogenic effects of many fungi. An analysis of all reported interactions with ECM14 supports the likelihood that a functional secretory pathway and other aspects of organellar trafficking are required for Ecm14 function. Unfortunately, attempts to further identify redundant pathways through a synthetic lethal approach suggested that redundant pathways were many, such that a single-gene synthetic lethal clone from the analysis of nearly 30,000 mutants was unable to be isolated, or that Ecm14 was not required in the conditions and cells examined during these assays.

It could be that a clear phenotype for *S. cerevisiae* ecm14 mutants has not yet been identified because most yeast that are used in the lab are unable to effectively form aggregates. Scientists have effectively selected against this phenotype to make yeast

genetics simpler. The data suggests that, while Ecm14 may not be required in many laboratory strains of *S. cerevisiae*, Ecm14 is a conserved and important pseudoenzyme within many fungal species that exists in zymogen form until activated by a trypsin-like endopeptidase, whereupon it likely functions in cell wall maintenance and/or fungal invasion.

APPENDIX A

TABLES

Table 1. PCR primers used in this study.

Name	nt	GC %	Sequence (5' to 3')
T7	20	40.0	TAATACGACTCACTATAGGG
T3	21	42.9	GCAATTAACCCTCACTAAAGG
ECM14_A	25	36.0	ATTTTAGTTCTTGTGGTGACAGCTT
ECM14_D	25	36.0	AATCACTAAAAAGGAGATGAAGCCT
ECM14_A-*Bam*HI	35	37.1	(TATA)[GGATCC]ATTTTAGTTCTTGTGGTGACAGCTT
ECM14_D-*Bam*HI	35	37.1	(TATA)[GGATCC]AATCACTAAAAAGGAGATGAAGCCT
ECM14_Internal	20	50.0	GCAGATAAATGATGGGCAGG
Kan_C (KanC)	22	36.4	TGATTTTGATGACGAGCGTAAT
ECM14_H69A-for	28	57.1	CACAGGTGGTATA{GC}TGCCAGAGAGTGG
ECM14_H69A-rev	28	57.1	CCACTCTCTGGCA{GC}TATACCACCTGTG
ECM14_H145A-for	30	43.3	GTATCGACATTGAT{GC}TTCCTTTGGTTTCC
ECM14_H145A-rev	30	43.3	GGAAACCAAAGGAA{GC}ATCAATGTCGATAC
ECM14_K270A-for	30	53.3	GGGCATTCCAGTTG{GC}ATTACGGGATACAG
ECM14_K270A-rev	30	53.3	CTGTATCCCGTAAT{GC}CAACTGGAATGCCC
ECM14_K270E-for	30	50.0	CTGTATCCCGTAATT{C}CAACTGGAATGCCC
ECM14_K270E-rev	30	50.0	GGGCATTCCAGTTG{G}AATTACGGGATACAG

() denote restriction enzyme sitting sequence.
[] denote restriction enzyme recognition sequence.
{} denote mutagenic substitution sites.

Table 2. Plasmids constructed and used in this study.

Name	Kb	Insert	Features	Source	Reference
pBluescript II SK+	3.0	--	Amp^R, LacZα, ORI	Peter J. Lyons	Alting-Mees and Short, 1989
pRS316	4.9	--	Amp^R, CEN4/ARS1, LacZα, ORI, URA3	Melanie J. Dobson	Sikorski and Hieter, 1989
pAG25	3.7	--	Amp^R, NatMX, ORI	Peter J. Lyons	Goldstein et al., 1999
pSLS1	10.8	--	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	Melanie J. Dobson	Barbour et al., 2000
pRS316-NatMX	6.2	NatMX	Amp^R, CEN4/ARS1, LacZα, ORI, URA3	This study	--
pBluescript-ECM14[1]	4.8	ECM14	Amp^R, LacZα, ORI	This study	--
-H69A[1]	4.8	-H69A	Amp^R, LacZα, ORI	This study	--
-H145A[1]	4.8	-H145A	Amp^R, LacZα, ORI	This study	--
-K270A[1]	4.8	-K270A	Amp^R, LacZα, ORI	This study	--
-K270E[1]	4.8	-K270E	Amp^R, LacZα, ORI	This study	--
pSLS1-ECM14[1]	12.7	ECM14	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	This study	--
-H69A[1]	12.7	-H69A	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	This study	--
-H145A[1]	12.7	-H145A	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	This study	--
-K270A[1]	12.7	-K270A	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	This study	--
-K270E[1]	12.7	-K270E	ADE3, Amp^R, CEN4/ARS1, ORI, URA3, pGAL1	This study	--

[1] Both insert orientations were constructed. The minus orientation was used in this study.

Table 3. Three conserved positions mutated in ECM14 using QuikChange SDM.

Role (Position[1])	**M14 MCPs**	**ECM14**	***-H69A***	***-H145A***	***-K270E***	***-K270A***
Zinc Binding (**69**)	His/H, (Positive)	His/H, Positive	**Ala/A, (Nonpolar)**			
Substrate Binding (**145**)	Arg/R, Positive	His/H, Positive		**Ala/A, (Nonpolar)**		
Catalysis (**270**)	Glu/E, Negative	Lys/K, Positive			**Glu/E, (Negative)**	**Ala/A, (Nonpolar)**

[1] Relative to the mature polypeptide sequence of bovine CPA1.

Table 4. Yeast strains used in this study.

Name	Background	Genotype	Source
Y1239	--	MATa {ura3-52 lys2-801 ade2-101 trpl-1 his3-200 TUB4::3XHA, URA3}	Peter J. Lyons
YHR132CΔ	BY4741/2	MATa/α {leu2Δ met15Δ ura3Δ ecm14::KanMX}	Peter J. Lyons
AF-1A	W303	MATα {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ}	Melanie J. Dobson
AF-1D	W303	MATa {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ}	Melanie J. Dobson
CM-1A_2	AF-1A (W303)	MATα {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ ecm14::KanMX}	This study
CM-1D_3	AF-1D (W303)	MATa {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ ecm14::KanMX}	This study
CM-1A_2[ECM14][1]	AF-1A (W303)	MATα {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ ecm14::KanMX} [pSLS1-ECM14]*	This study
CM-1D_3[ECM14][1]	AF-1D (W303)	MATa {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ ecm14::KanMX} [pSLS1-ECM14]*	This study
CM-1A[NAT]	AF-1A (W303)	MATα {ade2-1 ura3-1 leu2-3,-112 his3-11,-15 trp1-1 ade3Δ} [pRS316-NatMX]	This study

[1] Both insert orientations were constructed. The minus orientation was used in this study.
{ } denote genomic-based features.
[ ] denote plasmid-based features.

Table 5. Whole-genome sequencing using the Illumina MiniSeq® platform.

Sample	Bases Read	Paired Reads	Matched Reads	Read Length $\bar{x}$ (σ)	Coverage $\bar{x}$	GC %
WT	878,377,752	5,997,874	4,899,419 (81.7%)	146.4 (15.7)	56.8	41.0
SL6	581,444,805	3,959,960	3,216,531 (80.6%)	146.8 (15.7)	36.9	41.0
SL7	494,730,687	3,451,230	2,746,783 (79.6%)	143.3 (19.6)	31.6	40.8
SL8	634,384,283	4,321,856	3,626,739 (83.9%)	146.8 (15.3)	42.5	40.5
SL11	633,162,416	4,309,744	3,721,436 (86.3%)	146.9 (15.2)	43.5	40.4
SL12	743,482,836	5,087,184	4,611498 (90.6%)	146.1 (16.2)	53.3	40.4
SL13	618,472,666	4,215,912	3,744,904 (88.8%)	146.7 (16.1)	44.3	40.5
SL14	561,659,729	3,869,674	3,400,631 (87.9%)	145.1 (17.9)	39.2	40.6
Σ	**5,145,715,174**	**35,213,434**	**29,967,941 (85.1%)**			

Table 6. Mutant essential genes identified in this study.

Gene/ORF Name	SL	Variant	PROVEAN Score[1]	Gene/ORF Name	SL	Variant	PROVEAN Score[1]
ARC19 (YKL013C)	7	**E37K**	-3.838	**PRP8** (YHR165C)	6	**E204K**	-3.563
ARH1 (YDR376W)	14	**S119N**	-3.000	**PSE1** (YMR308C)	14	**D324N**	-5.000
ARP5 (YNL059C)	13	**G129R**	-4.940	**REA1** (YLR106C)	6	**L2361F**	-2.526
BRR2 (YER172C)	12	**A2113T**	-3.080	**RET3** (YPL010W)	12	**G19R**	-7.279
BRR2 (YER172C)	8	**S51F**	-5.146	**ROM2** (YLR371W)	13	**A1293T**	-3.540
CFT1 (YDR301W)	13	**N1072Y**	-7.500	**RPA135** (YPR010C)	12	**P951L**	-9.250
COP1 (YDL145C)	11	**G54S**	-5.978	**RPB2** (YOR151C)	12	**C829Y**	-9.421
CYR1 (YJL005W)	12	**G1471D**	-5.887	**RPG1** (YBR079C)	6	**A582T**	-2.555
EXO84 (YBR102C)	14	**P92S**	-2.652	**SEC7** (YDR170C)	8	**S322F**	-2.719
FHL1 (YPR104C)	11	**T647M**	-5.400	**SLY1** (YDR189W)	13	**S91F**	-3.537
GCD6 (YDR211W)	14	**R120K**	-2.980	**SLY1** (YDR189W)	13	**S130L**	-5.365
GUS1 (YGL245W)	8	**A222T**	-4.000	**SMC2** (YFR031C)	7	**A168T**	-3.996
ILV3 (YJR016C)	12	**P148L**	-9.701	**SMC3** (YJL074C)	14	**A1141T**	-3.967
ILV3 (YJR016C)	7	**G454D**	-5.304	**SOG2** (YOR353C)	13	**G227E**	-6.911
KIN28 (YDL108W)	13	**E84K**	-3.989	**SSC1** (YJR045C)	14	**A281T**	-3.970
MET30 (YIL046W)	11	**G590S**	-3.015	**STH1** (YIL126W)	13	**T408I**	-5.967
MOT1 (YPL082C)	11	**P1265S**	-7.933	**STT4** (YLR305C)	14	**A662T**	-3.311
MSS4 (YDR208W)	11	**P432L**	-9.567	**SUA5** (YGL169W)	12	**A208V**	-4.000
NET1 (YJL076W)	8	**E107K**	-3.133	**TAM41** (YGR046W)	8	**G129S**	-3.200
NFS1 (YCL017C)	12	**V401M**	-2.647	**TIF34** (YMR146C)	12	**P143L**	-6.270
NOP15 (YNL110C)	14	**S131F**	-5.750	**TOR2** (YKL203C)	11	**P434L**	-6.408
NOP8 (YOL144W)	13	**R174S**	-3.733	**TRL1** (YJL087C)	14	**R352K**	-3.000
PAN1 (YIR006C)	8	**R635C**	-7.952	**TRR1** (YDR353W)	12	**G130C**	-8.827
PCM1 (YEL058W)	11	**G233D**	-6.965	**TRS120** (YDR407C)	14	**P935S**	-3.715
PRP6 (YBR055C)	8	**P734S**	-5.639	**UTP20** (YBL004W)	13	**D517N**	-3.428

[1] -2.50 PROVEAN threshold was used to qualify deleterious (≤) or neutral (>) variants.

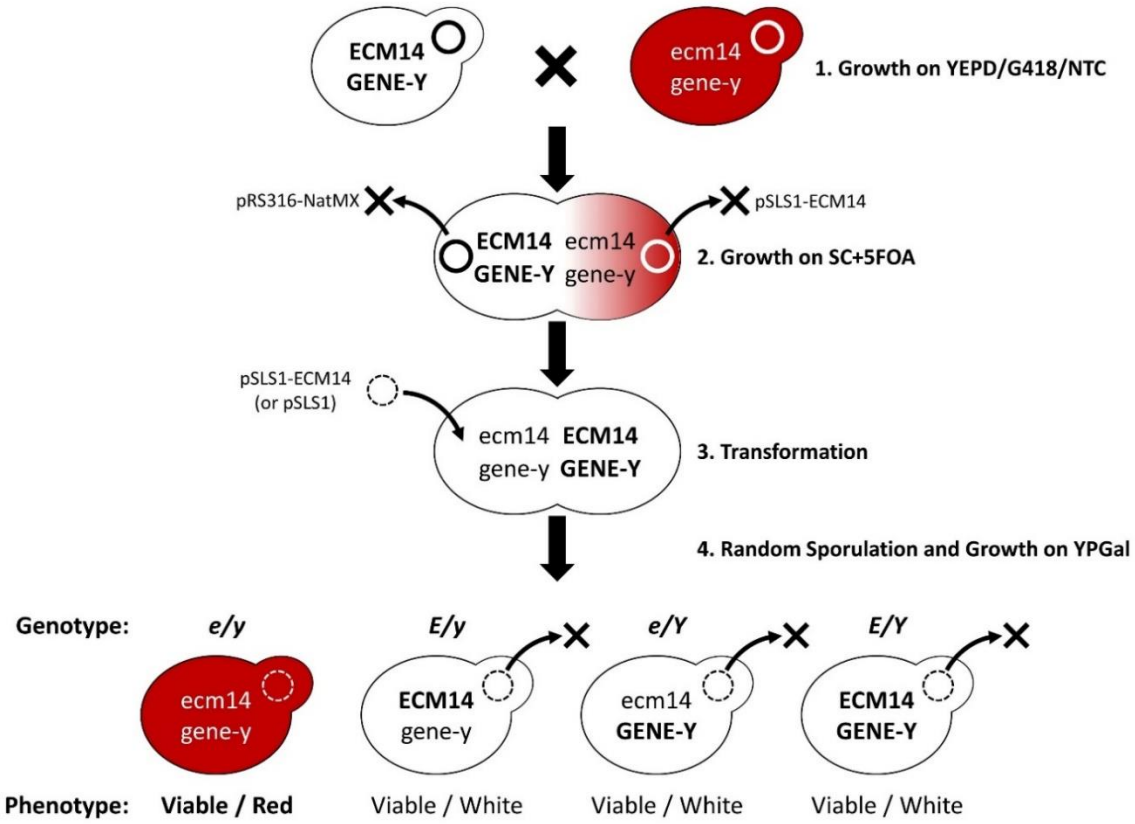

Figure 44. The functional rescue assay workflow.

(**1**) A diploid yeast strain is produced by crossing a wild-type haploid [pRS316-NatMX] strain with a haploid synthetically lethal strain (gene-y, ecm14::KanMX). Diploids are selected for by growth on YPD/G418/NTC. (**2**) Growth on SC+5FOA allows for selection against the now unnecessary pRS316-NatMX and pSLS1-ECM14 plasmids. (**3**) Cells are transformed with the pSLS1 or pSLS1-ECM14 plasmids followed by (**4**) random sporulation and germination of these haploid spores on YPGal.

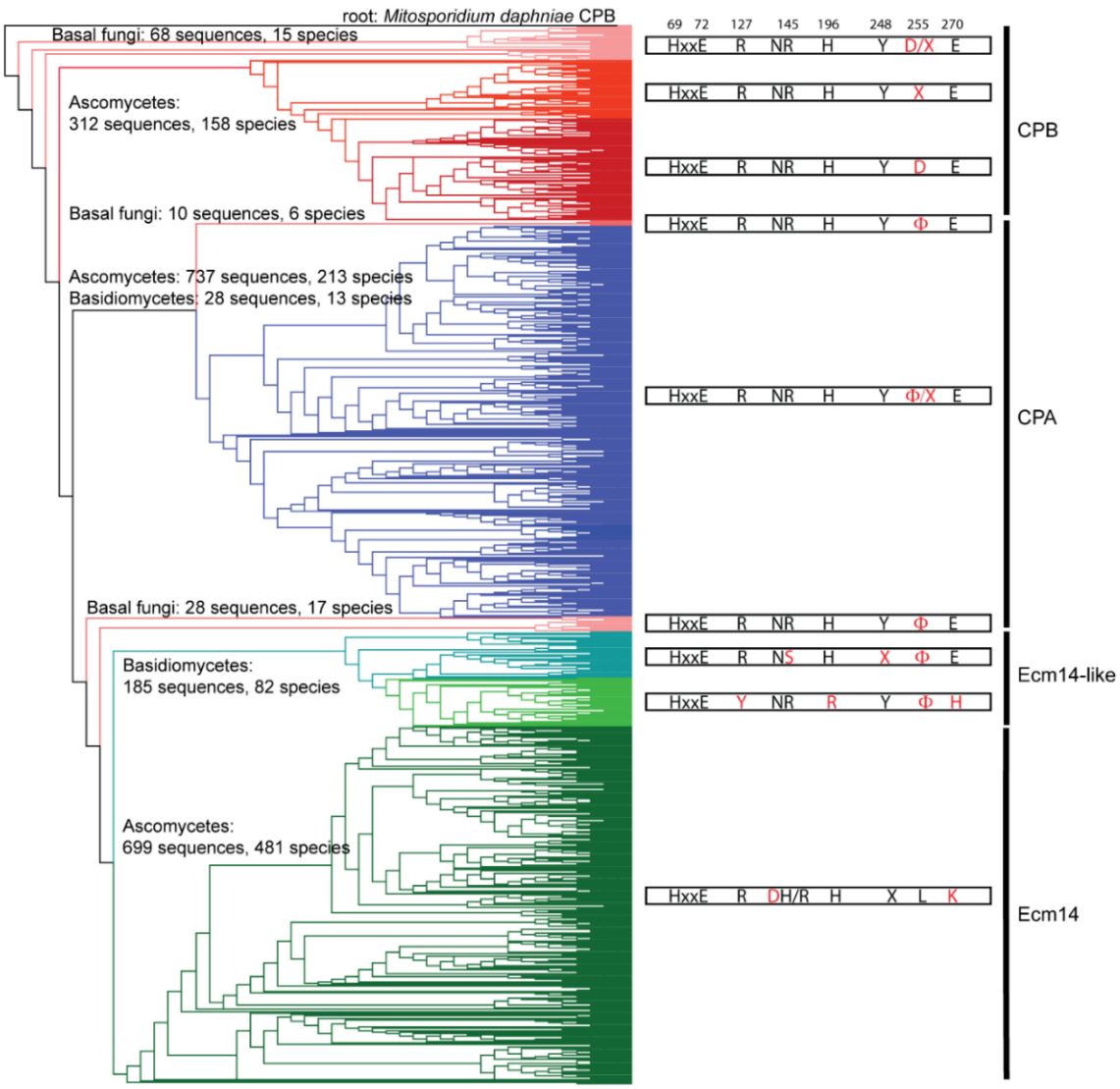

Figure 45. Ecm14 is conserved in Ascomycota, but not in other fungal phyla.

BLASTp was performed with bovine CPA1 as the query sequence and all fungal sequences as the search set. 2068 sequences of 150 amino acids or longer that contained the majority of the carboxypeptidase domain were identified and aligned using Clustal Omega. The resulting tree is shown, containing three predominant groups consistent with CPB-like, CPA-like, and Ecm14-like properties. Sequences from basal fungi are shown in pink. CPB sequences from ascomycetes are shown in dark red, with co-clustering sequences lacking the required Asp at position 255 shown in a lighter red. CPA-like sequences, mostly from ascomycetes, are shown in blue. Ecm14 sequences, all from ascomycetes, are shown in dark green, with a sister clade of Ecm14-like proteins from basidiomycetes, and with unique and conserved 'active' site residues, shown in cyan and light green. Ecm14 proteins are defined by the presence of N144D and E270K substitutions.

APPENDIX B

DIRECT SELECTION OF YEAST DIPLOIDS

Subcloning the NatMX Allele From pAG25 into pRS316

The original yeast strains used in this study are isogenic and, thus, lack complementary recessive auxotrophic markers typically used for the routine selection of diploids from mating reactions (*see above* ***Table 4***). Furthermore, these autotrophies can impair yeast growth and interfere with downstream methods like sporulation (Sadowski et al., 2008). Thus, alternative methods for selecting diploids can be desirable. In order to facilitate both the direct selection and stable propagation of diploids strains produced in this study, an additional yeast plasmid harboring a unique drug-resistance marker was constructed. By mating haploid yeast strains that each harbor a unique drug-resistance marker (e.g. NatMX and KanMX), it is possible to directly select for diploids on media containing a combination of these two drugs (e.g. NTC and G418). Only those diploids containing both markers can survive on this selection media (*see* **Figure *46***).

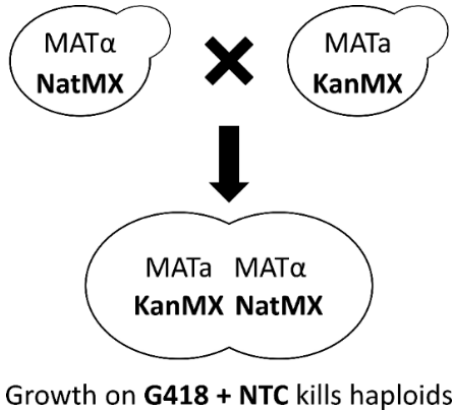

Figure 46. The direct selection of diploids using multiple drug-resistance markers.

Mating haploid yeast strains that each harbor a unique drug-resistance marker produces diploids that can grow on media containing both drugs. Because the parental haploids each lack one of these drug-resistance markers, these haploids are unable to survive on this media.

The dominant NatMX allele from the pAG25 plasmid was subcloned into the yeast shuttle vector pRS316 (*see above* ***Table 2***). The pAG25 plasmid alone is only useful as a template for PCR to create gene deletion cassettes and, thus, lacks the CEN4/ARS1 sequences required for replication in yeast. An ample amount of both pAG25 and pRS316 were thoroughly digested with *Bam*HI and *Eco*RI to ensure adequate recovery yields for future downstream applications. 15 µl of either pAG25 or pRS316 was combined with 28 µl of sterile Milli-Q water, 5 µl of CutSmart Buffer,1 µl of *Bam*HI-HF enzyme, and 1 µl of *Eco*RI-HF enzyme. Both digestion reactions were allowed to incubate for 2 hours at 37°C. Both digestion reactions were subsequently electrophoresed in a 0.8% agarose gel (0.5 µg/ml EtBr) for 60 minutes to allow for separation of the pAG25 (2,465 bp) and NatMX (1,239 bp) bands.

The pRS316 and NatMX bands were both gel extracted using the modified protocol for large DNA fragments, ran in an additional 0.8% agarose gel (0.5 µg/ml EtBr), and quantified using ImageJ. The pRS316-NatMX ligation reaction was performed overnight at 16°C with a 1:3 vector-to-insert ratio. This ligation reaction contained 0.5 µl of pRS316-*Bam*HI-*Eco*RI (44.62 ng/µl, 0.014 pmol), 13 µl of NatMX-*Bam*HI-*Eco*RI (2.68 ng/µl, 0.043 pmol), 3.5 µl of sterile Milli-Q water, 2 µl of T4 DNA Ligase Buffer, and 1 µl of T4 DNA Ligase.

One 50-µl aliquot of Invitrogen Max Efficiency DH5α was slowly thawed on wet ice and gently mixed with 5 µl of the overnight pRS316-NatMX ligation reaction. This transformation mixture was incubated on wet ice for 30 minutes, heat shocked at 42°C for 45 seconds, and returned to wet ice for 2 minutes. The transformation mixture was incubated for 60 minutes at 37°C in 450 µl of fresh SOC without ampicillin to allow for

adequate expression of the ampicillin resistance gene prior to plating. This outgrowth culture was subsequently plated on dry IPTG/X-Gal LB agar plates (250 µl/plate) with ampicillin and incubated at 37°C overnight for >20 hours.

Four putative pRS316-NatMX colonies (white) were cultured overnight and miniprepped using the NID method and digested with *Bam*HI-HF and *Eco*RI-HF at 37°C for 60 minutes. Each of these diagnostic digests included 5 µl of plasmid DNA, 12 µl of sterile Milli-Q water, 2 µl of CutSmart Buffer, and 1 µl of *Bam*HI-HF enzyme. Each digest was ran in a 0.8% agarose gel (0.5 µg/ml EtBr) and imaged using the AlphaImager HP gel imaging system (*see* ***Error! Reference source not found.*****)**. Only those digests that clearly l iberated both a pRS316 band (4.9-kb) and a NatMX band (1,239-bp) were considered for further interrogation. A map of the pRS316-NatMX plasmid was constructed using SnapGene (*see* **Figure *48***). Diploid selection using the pRS316-NatMX plasmid was demonstrated using the yeast strains relevant to this study (*see* **Figure *49***).

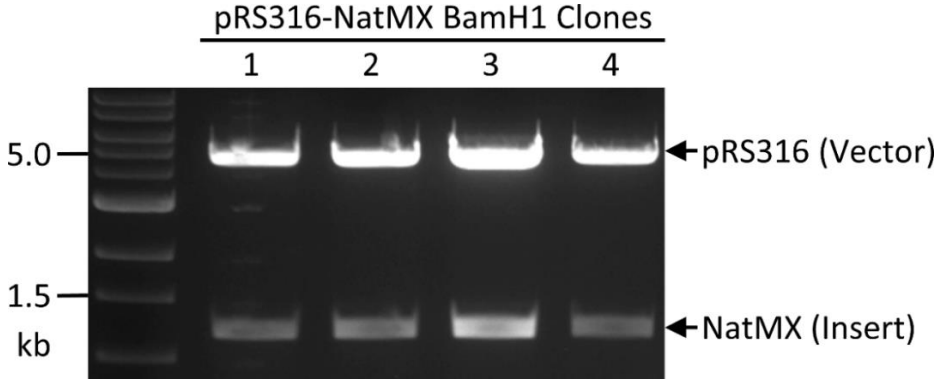

Figure 47. Subcloning of NatMX from pAG25 into pRS316.

The NatMX gene was subcloned from the pAG25 (3.7 kb) parent plasmid into the pRS316 (4.9 kb) destination vector. The pAG25 parent plasmid and pRS316 destination vector were double digested with BamH1 and EcoRI. T4 DNA ligase was used to ligate the NatMX insert into pRS316. Four putative pRS316-NatMX preparations were digested with *Bam*HI and *Eco*RI. Only those digests that clearly liberated both a pRS316 band (4.9 kb) and a NatMX band (1.2 kb) were considered for further interrogation.

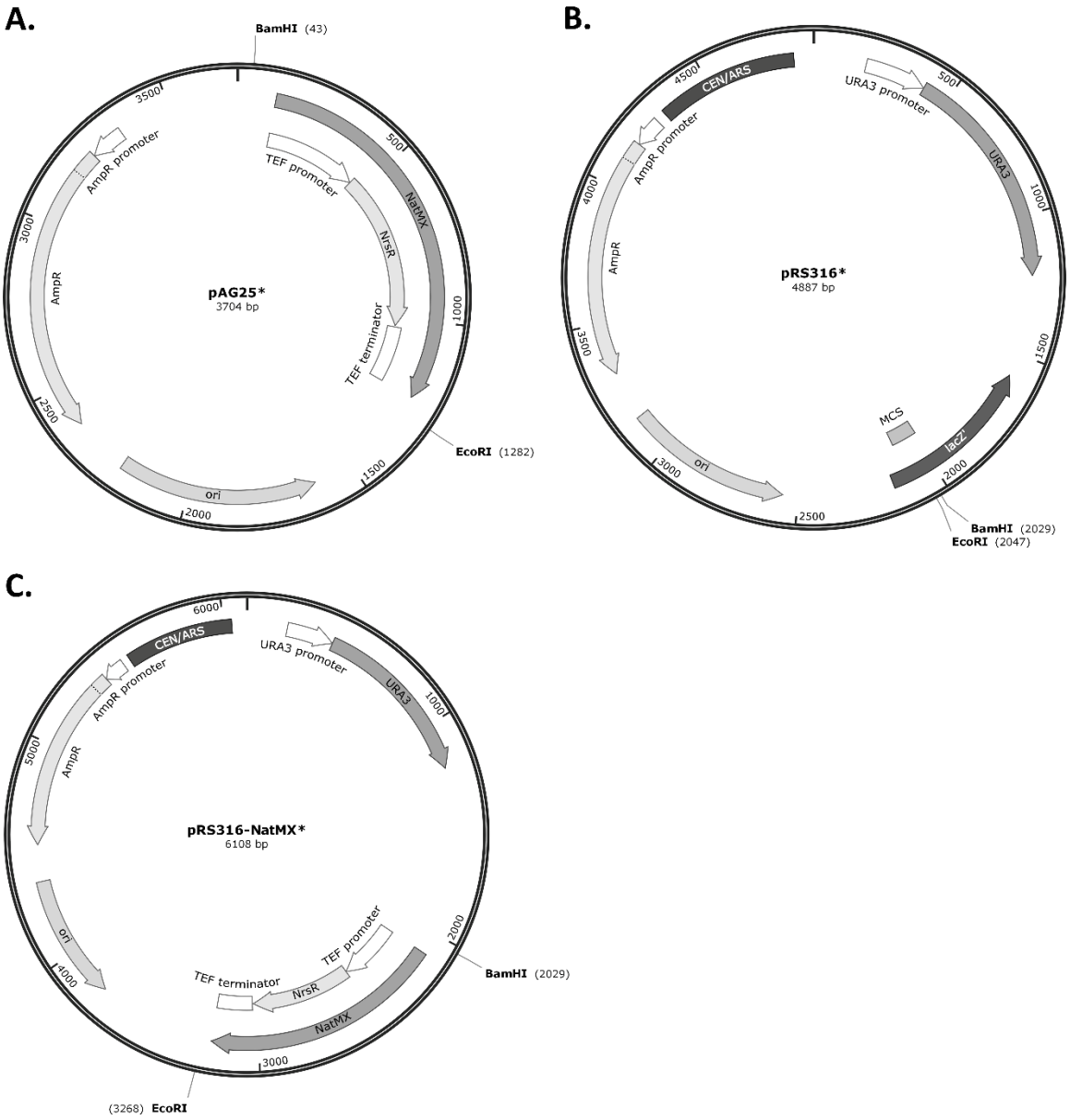

Figure 48. Diagrams of the pAG25 plasmid, the pRS316 shuttle vector, and the recombinant pRS316-NatMX plasmid.

(**A**) The pAG25 plasmid was the source of the NatMX expression cassette. (**B**) The pRS316 plasmid is able to replicate in both bacteria (high-copy) and yeast (single-copy). (**C**) The recombinant pRS316-NatMX plasmid. Only plasmid features relevant to this study are shown (*).

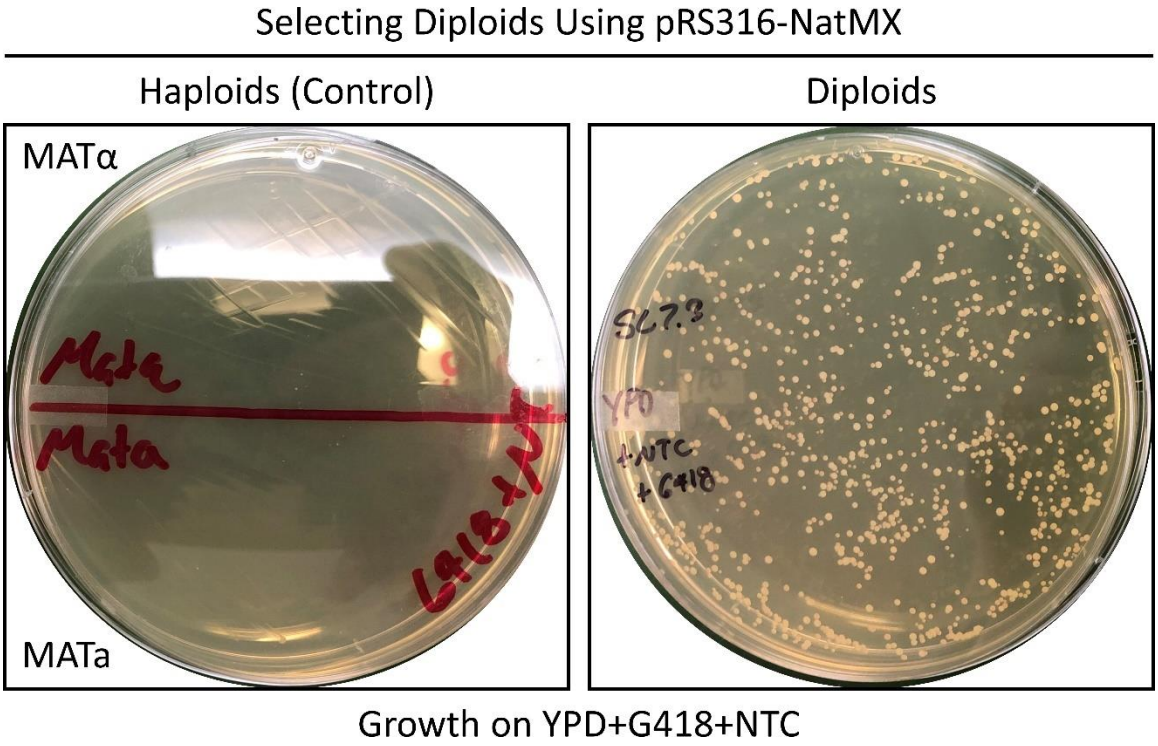

Figure 49. Selecting diploids using the pRS316-NatMX plasmid.

A basic proof-of-concept demonstrating the direct selection of yeast diploids using two unique drug-resistance markers and selective growth media containing a combination of two drugs. Haploid control yeast (MATα and MATa) were streaked on this YPD+G418+NTC media. Diploids were mated by mixing both mating types on a YPD plate (without drugs) and incubating overnight at 30°C. This overnight mating reaction was resuspended in 1ml of sterile water, diluted 100X, and plated on this YPD+G418+NTC media.

REFERENCE LIST

Adrain, C., & Freeman, M. (2012). New lives for old: evolution of pseudoenzyme function illustrated by iRhoms. *Nat Rev Mol Cell Biol, 13*(8), 489-498. doi:10.1038/nrm3392

Aguiar, T. Q., Silva, R., & Domingues, L. (2015). Ashbya gossypii beyond industrial riboflavin production: A historical perspective and emerging biotechnological applications. *Biotechnol Adv, 33*(8), 1774-1786. doi:10.1016/j.biotechadv.2015.10.001

Aguilar-Uscanga, B., & Francois, J. M. (2003). A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. *Lett Appl Microbiol, 37*(3), 268-274. doi:10.1046/j.1472-765x.2003.01394.x

Aliouat-Denis, C. M., Chabe, M., Demanche, C., Aliouat el, M., Viscogliosi, E., Guillot, J., Delhaes, L., & Dei-Cas, E. (2008). Pneumocystis species, co-evolution and pathogenic power. *Infect Genet Evol, 8*(5), 708-726. doi:10.1016/j.meegid.2008.05.001

Alsters, S. I. M., Goldstone, A. P., Buxton, J. L., Zekavati, A., Sosinsky, A., Yiorkas, A. M., Holder, S., Klaber, R. E., Bridges, N., van Haelst, M. M., le Roux, C. W., Walley, A. J., Walters, R. G., Mueller, M., & Blakemore, A. I. F. (2015). Truncating Homozygous Mutation of Carboxypeptidase E (CPE) in a Morbidly Obese Female with Type 2 Diabetes Mellitus, Intellectual Disability and Hypogonadotrophic Hypogonadism. *PLOS ONE, 10*(6), e0131417. doi:10.1371/journal.pone.0131417

Anand, L., & Rodriguez Lopez, C. M. (2020). chromoMap: An R package for Interactive Visualization and Annotation of Chromosomes. *bioRxiv*, 605600. doi:10.1101/605600

Arolas, J. L., Vendrell, J., Aviles, F. X., & Fricker, L. D. (2007). Metallocarboxypeptidases: emerging drug targets in biomedicine. *Curr Pharm Des, 13*(4), 349-366. doi:10.2174/138161207780162980

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29. doi:10.1038/75556

Auld, D. (2004). Catalytic mechanisms of metallopeptidases. *Handbook of Proteolytic Enzymes*, 268-289.

Austin, B. P., Tozser, J., Bagossi, P., Tropea, J. E., & Waugh, D. S. (2011). The substrate specificity of Metarhizium anisopliae and Bos taurus carboxypeptidases A: insights into their use as tools for the removal of affinity tags. *Protein Expr Purif, 77*(1), 53-61. doi:10.1016/j.pep.2010.11.005

Austin, B. P., & Waugh, D. S. (2012). Isolation of Metarhizium anisopliae carboxypeptidase A with native disulfide bonds from the cytosol of Escherichia coli BL21(DE3). *Protein Expr Purif, 82*(1), 116-124. doi:10.1016/j.pep.2011.11.015

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

Barbour, L., & Xiao, W. (2006). Synthetic lethal screen. *Methods Mol Biol, 313*, 161-169. doi:10.1385/1-59259-958-3:161

Barbour, L., Zhu, Y., & Xiao, W. (2000). Improving synthetic lethal screens by regulating the yeast centromere sequence. *Genome, 43*(5), 910-917. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11081983

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., & Cullin, C. (1993). A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. *Nucleic acids research, 21*(14), 3329-3330. doi:10.1093/nar/21.14.3329

Bergthorsson, U., Andersson, D. I., & Roth, J. R. (2007). Ohno's dilemma: evolution of new genes under continuous selection. *Proc Natl Acad Sci U S A, 104*(43), 17004-17009. doi:10.1073/pnas.0707158104

Birdsell, J., & Wills, C. (2003). The Evolutionary Origin and Maintenance of Sexual Recombination: A Review of Contemporary Models. *Evol. Biol., 33*, 27-138. doi:10.1007/978-1-4757-5190-1_2

Blobel, C. P., Carpenter, G., & Freeman, M. (2009). The role of protease activity in ErbB biology. *Exp Cell Res, 315*(4), 671-682. doi:10.1016/j.yexcr.2008.10.011

Bordallo, J., & Suarez-Rendueles, P. (1993). Control of Saccharomyces cerevisiae carboxypeptidase S (CPS1) gene expression under nutrient limitation. *Yeast, 9*(4), 339-349. doi:10.1002/yea.320090404

Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., & Sherlock, G. (2004). GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics, 20*(18), 3710-3715. doi:10.1093/bioinformatics/bth456

Boynton, P. J., & Greig, D. (2014). The ecology and evolution of non-domesticated Saccharomyces species. *Yeast, 31*(12), 449-462. doi:10.1002/yea.3040

Brookes, P., & Lawley, P. D. (1961). The reaction of mono- and di-functional alkylating agents with nucleic acids. *Biochem J, 80*(3), 496-503. doi:10.1042/bj0800496

Chen, X.-L., Liu, C., Tang, B., Ren, Z., Wang, G.-L., & Liu, W. (2020). Quantitative proteomics analysis reveals important roles of N-glycosylation on ER quality control system for development and pathogenesis in Magnaporthe oryzae. *PLOS Pathogens, 16*(2), e1008355. doi:10.1371/journal.ppat.1008355

Cherry, J. M., Hong, E. L., Amundsen, C., Balakrishnan, R., Binkley, G., Chan, E. T., Christie, K. R., Costanzo, M. C., Dwight, S. S., Engel, S. R., Fisk, D. G., Hirschman, J. E., Hitz, B. C., Karra, K., Krieger, C. J., Miyasato, S. R., Nash, R. S., Park, J., Skrzypek, M. S., Simison, M., Weng, S., & Wong, E. D. (2012). Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res, 40*(Database issue), D700-705. doi:10.1093/nar/gkr1029

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics (Oxford, England), 31*(16), 2745-2747. doi:10.1093/bioinformatics/btv195

Chow, J., Dionne, H. M., Prabhakar, A., Mehrotra, A., Somboonthum, J., Gonzalez, B., Edgerton, M., & Cullen, P. J. (2019). Aggregate Filamentous Growth Responses in Yeast. *mSphere, 4*(2). doi:10.1128/mSphere.00702-18

Chuang, J. S., & Schekman, R. W. (1996). Differential trafficking and timed localization of two chitin synthase proteins, Chs2p and Chs3p. *J Cell Biol, 135*(3), 597-610. doi:10.1083/jcb.135.3.597

Clarke, L., & Carbon, J. (1980). Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature, 287*(5782), 504-509. doi:10.1038/287504a0

Dobzhansky, T. (1946). Genetics of natural populations; recombination and variability in populations of Drosophila pseudoobscura. *Genetics, 31*, 269-290. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20985721

Duina, A. A., Miller, M. E., & Keeney, J. B. (2014). Budding yeast for budding geneticists: a primer on the Saccharomyces cerevisiae model system. *Genetics, 197*(1), 33-48. doi:10.1534/genetics.114.163188

Dunham, M. J., & Louis, E. J. (2011). Yeast evolution and ecology meet genomics. *EMBO Rep, 12*(1), 8-10. doi:10.1038/embor.2010.204

Dutt, A., Canevascini, S., Froehli-Hoier, E., & Hajnal, A. (2004). EGF signal propagation during C. elegans vulval development mediated by ROM-1 rhomboid. *PLoS Biol, 2*(11), e334. doi:10.1371/journal.pbio.0020334

Eaton, D. L., Malloy, B. E., Tsai, S. P., Henzel, W., & Drayna, D. (1991). Isolation, molecular cloning, and partial characterization of a novel carboxypeptidase B from human plasma. *J Biol Chem, 266*(32), 21833-21838. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/1939207

Engel, S. R., Dietrich, F. S., Fisk, D. G., Binkley, G., Balakrishnan, R., Costanzo, M. C., Dwight, S. S., Hitz, B. C., Karra, K., Nash, R. S., Weng, S., Wong, E. D., Lloyd, P., Skrzypek, M. S., Miyasato, S. R., Simison, M., & Cherry, J. M. (2014). The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3 (Bethesda), 4*(3), 389-398. doi:10.1534/g3.113.008995

Eriksson, O. E., & Winka, K. (1997). Supraordinal taxa of Ascomycota. *Myconet, 1*, 16.

Eyers, P. A., & Murphy, J. M. (2013). Dawn of the dead: protein pseudokinases signal new adventures in cell biology. *Biochem Soc Trans, 41*(4), 969-974. doi:10.1042/BST20130115

Eyers, P. A., & Murphy, J. M. (2016). The evolving world of pseudoenzymes: proteins, prejudice and zombies. *BMC Biol, 14*(1), 98. doi:10.1186/s12915-016-0322-x

Ezihe-Ejiofor, J. A., & Hutchinson, N. (2013). Anticlotting mechanisms 1: physiology and pathology. *Continuing Education in Anaesthesia Critical Care & Pain, 13*(3), 87-92. doi:10.1093/bjaceaccp/mks061

Faria-Oliveira, F., Carvalho, J., Ferreira, C., Hernáez, M. L., Gil, C., & Lucas, C. (2015). Quantitative differential proteomics of yeast extracellular matrix: there is more to it than meets the eye. *BMC microbiology, 15*, 271-271. doi:10.1186/s12866-015-0550-1

Ferguson, B. A., Dreisbach, T. A., Parks, C. G., Filip, G. M., & Schmitt, C. L. (2003). Coarse-scale population structure of pathogenic Armillaria species in a mixed-conifer forest in the Blue Mountains of northeast Oregon. *Canadian Journal of Forest Research, 33*(4), 612-623. doi:10.1139/x03-065

Fesel, P. H., & Zuccaro, A. (2016). β-glucan: Crucial component of the fungal cell wall and elusive MAMP in plants. *Fungal Genetics and Biology, 90,* 53-60. doi:https://doi.org/10.1016/j.fgb.2015.12.004

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics, 151*(4), 1531-1545. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10101175

Fricker, L. D. (2018). Carboxypeptidase E and the Identification of Novel Neuropeptides as Potential Therapeutic Targets. *Advances in pharmacology (San Diego, Calif.), 82*, 85-102. doi:10.1016/bs.apha.2017.09.001

Galagan, J. E., Calvo, S. E., Borkovich, K. A., Selker, E. U., Read, N. D., Jaffe, D., FitzHugh, W., Ma, L. J., Smirnov, S., Purcell, S., Rehman, B., Elkins, T., Engels, R., Wang, S., Nielsen, C. B., Butler, J., Endrizzi, M., Qui, D., Ianakiev, P., Bell-Pedersen, D., Nelson, M. A., Werner-Washburne, M., Selitrennikoff, C. P., Kinsey, J. A., Braun, E. L., Zelter, A., Schulte, U., Kothe, G. O., Jedd, G., Mewes, W., Staben, C., Marcotte, E., Greenberg, D., Roy, A., Foley, K., Naylor, J., Stange-Thomann, N., Barrett, R., Gnerre, S., Kamal, M., Kamvysselis, M., Mauceli, E., Bielke, C., Rudd, S., Frishman, D., Krystofova, S., Rasmussen, C., Metzenberg, R. L., Perkins, D. D., Kroken, S., Cogoni, C., Macino, G., Catcheside, D., Li, W., Pratt, R. J., Osmani, S. A., DeSouza, C. P., Glass, L., Orbach, M. J., Berglund, J. A., Voelker, R., Yarden, O., Plamann, M., Seiler, S., Dunlap, J., Radford, A., Aramayo, R., Natvig, D. O., Alex, L. A., Mannhaupt, G., Ebbole, D. J., Freitag, M., Paulsen, I., Sachs, M. S., Lander, E. S., Nusbaum, C., & Birren, B. (2003). The genome sequence of the filamentous fungus Neurospora crassa. *Nature, 422*(6934), 859-868. doi:10.1038/nature01554

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., & Oliver, S. G. (1996). Life with 6000 genes. *Science, 274*(5287), 546, 563-547. doi:10.1126/science.274.5287.546

Gomis-Ruth, F. X. (2008). Structure and mechanism of metallocarboxypeptidases. *Crit Rev Biochem Mol Biol, 43*(5), 319-345. doi:10.1080/10409230802376375

Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C., & McCombie, W. R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Res, 25*(11), 1750-1756. doi:10.1101/gr.191395.115

Greenblatt, E. J., Olzmann, J. A., & Kopito, R. R. (2011). Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum. *Nat Struct Mol Biol, 18*(10), 1147-1152. doi:10.1038/nsmb.2111

Greig, D., & Leu, J. Y. (2009). Natural history of budding yeast. *Curr Biol, 19*(19), R886-890. doi:10.1016/j.cub.2009.07.037

Hancock, J. M., & Bishop, M. J. (2014). EMBOSS (The European Molecular Biology Open Software Suite). *Dictionary of Bioinformatics and Computational Biology*. doi:doi:10.1002/9780471650126.dob0206.pub2

Harrington, B. J., & Raper, K. B. (1968). Use of a fluorescent brightener to demonstrate cellulose in the cellular slime molds. *Applied microbiology, 16*(1), 106-113. Retrieved from https://pubmed.ncbi.nlm.nih.gov/5688830

Hawksworth, D., & Lücking, R. (2017). Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiology Spectrum, 5*. doi:10.1128/microbiolspec.FUNK-0052-2016

Herth, W., & Schnepf, E. (1980). The fluorochrome, calcofluor white, binds oriented to structural polysaccharide fibrils. *Protoplasma, 105*(1), 129-133. doi:10.1007/BF01279855

Hillenmeyer, M. E., Fung, E., Wildenhain, J., Pierce, S. E., Hoon, S., Lee, W., Proctor, M., St Onge, R. P., Tyers, M., Koller, D., Altman, R. B., Davis, R. W., Nislow, C., & Giaever, G. (2008). The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science, 320*(5874), 362-365. doi:10.1126/science.1150021

Hinnen, A., Hicks, J. B., & Fink, G. R. (1978). Transformation of yeast. *Proc Natl Acad Sci U S A, 75*(4), 1929-1933. doi:10.1073/pnas.75.4.1929

Hoffman, C. S., Wood, V., & Fantes, P. A. (2015). An Ancient Yeast for Young Geneticists: A Primer on the Schizosaccharomyces pombe Model System. *Genetics, 201*(2), 403-423. doi:10.1534/genetics.115.181503

Houbraken, J., Frisvad, J. C., & Samson, R. A. (2011). Fleming's penicillin producing strain is not Penicillium chrysogenum but P. rubens. *IMA Fungus, 2*(1), 87-95. doi:10.5598/imafungus.2011.02.01.12

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., & O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. *Nature, 425*(6959), 686-691. doi:10.1038/nature02026

Idowu, T. A., Schott, M. J., & Lyons, P. J. (2016). Characterization of Ecm14, a Fungal Pseudopeptidase. *The FASEB Journal, 30*(S1), 596.597-596.597. doi:10.1096/fasebj.30.1_supplement.596.7

Igarashi, R., Suzuki, M., Nogami, S., & Ohya, Y. (2005). Molecular dissection of ARP1 regions required for nuclear migration and cell wall integrity checkpoint functions in Saccharomyces cerevisiae. *Cell Struct Funct, 30*(2), 57-67. doi:10.1247/csf.30.57

Jeffery, C. J. (2003). Moonlighting proteins: old proteins learning new tricks. *Trends Genet, 19*(8), 415-417. doi:10.1016/S0168-9525(03)00167-7

Joshi, L., & St Leger, R. J. (1999). Cloning, expression, and substrate specificity of MeCPA, a zinc carboxypeptidase that is secreted into infected tissues by the fungal entomopathogen Metarhizium anisopliae. *J Biol Chem, 274*(14), 9803-9811. doi:10.1074/jbc.274.14.9803

Jouhten, P., Ponomarova, O., Gonzalez, R., & Patil, K. R. (2016). Saccharomyces cerevisiae metabolism in ecological context. *FEMS Yeast Res, 16*(7). doi:10.1093/femsyr/fow080

Kalinina, E., Biswas, R., Berezniuk, I., Hermoso, A., Aviles, F. X., & Flicker, L. D. (2007). A novel subfamily of mouse cytosolic carboxypeptidases. *The FASEB Journal, 21*(3), 836-850. doi:10.1096/fj.06-7329com

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics, 28*(12), 1647-1649. doi:10.1093/bioinformatics/bts199

Kirk, P. M., Cannon, P. F., David, J. C., & Stalpers, J. A. (2001). *Ainsworth and Bisby's dictionary of the fungi: 9th edition*. Wallingford: CABI Publishing.

Klis, F. M., Boorsma, A., & De Groot, P. W. (2006). Cell wall construction in Saccharomyces cerevisiae. *Yeast, 23*(3), 185-202. doi:10.1002/yea.1349

Koshland, D., Kent, J. C., & Hartwell, L. H. (1985). Genetic analysis of the mitotic transmission of minichromosomes. *Cell, 40*(2), 393-403. doi:10.1016/0092-8674(85)90153-9

Kwon, A., Scott, S., Taujale, R., Yeung, W., Kochut, K. J., Eyers, P. A., & Kannan, N. (2019). Tracing the origin and evolution of pseudokinases across the tree of life. *Sci Signal, 12*(578). doi:10.1126/scisignal.aav3810

Landvik, S., Schumacher, T., Eriksson, O., & Moss, S. (2003). Morphology and ultrastructure of Neolecta species. *Mycological research, 107*, 1021-1031. doi:10.1017/S0953756203008219

Lang, G. I., Parsons, L., & Gammie, A. E. (2013). Mutation rates, spectra, and genome-wide distribution of spontaneous mutations in mismatch repair deficient yeast. *G3 (Bethesda, Md.), 3*(9), 1453-1465. doi:10.1534/g3.113.006429

Laurent, J. M., Young, J. H., Kachroo, A. H., & Marcotte, E. M. (2016). Efforts to make and apply humanized yeast. *Brief Funct Genomics, 15*(2), 155-163. doi:10.1093/bfgp/elv041

Leducq, J. B., Charron, G., Samani, P., Dube, A. K., Sylvester, K., James, B., Almeida, P., Sampaio, J. P., Hittinger, C. T., Bell, G., & Landry, C. R. (2014). Local climatic adaptation in a widespread microorganism. *Proc Biol Sci, 281*(1777), 20132472. doi:10.1098/rspb.2013.2472

Lemberg, M. K. (2011). Intramembrane proteolysis in regulated protein trafficking. *Traffic, 12*(9), 1109-1118. doi:10.1111/j.1600-0854.2011.01219.x

Lesage, G., & Bussey, H. (2006). Cell wall assembly in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev, 70*(2), 317-343. doi:10.1128/MMBR.00038-05

Levine, B., Mizushima, N., & Virgin, H. W. (2011). Autophagy in immunity and inflammation. *Nature, 469*(7330), 323-335. doi:10.1038/nature09782

Lezin, G., Kosaka, Y., Yost, H. J., Kuehn, M. R., & Brunelli, L. (2011). A one-step miniprep for the isolation of plasmid DNA and lambda phage particles. *PLOS ONE, 6*(8), e23457. doi:10.1371/journal.pone.0023457

Liu, W., Li, L., Ye, H., Chen, H., Shen, W., Zhong, Y., Tian, T., & He, H. (2017). From Saccharomyces cerevisiae to human: The important gene co-expression modules. *Biomed Rep, 7*(2), 153-158. doi:10.3892/br.2017.941

Looke, M., Kristjuhan, K., & Kristjuhan, A. (2011). Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques, 50*(5), 325-328. doi:10.2144/000113672

Lussier, M., White, A. M., Sheraton, J., di Paolo, T., Treadwell, J., Southard, S. B., Horenstein, C. I., Chen-Weiner, J., Ram, A. F., Kapteyn, J. C., Roemer, T. W., Vo, D. H., Bondoc, D. C., Hall, J., Zhong, W. W., Sdicu, A. M., Davies, J., Klis, F. M., Robbins, P. W., & Bussey, H. (1997). Large scale identification of genes involved in cell surface biosynthesis and architecture in Saccharomyces cerevisiae. *Genetics, 147*(2), 435-450. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/9335584

Lyons, P. J., & Fricker, L. D. (2011). Carboxypeptidase O is a glycosylphosphatidylinositol-anchored intestinal peptidase with acidic amino acid specificity. *The Journal of biological chemistry, 286*(45), 39023-39032. doi:10.1074/jbc.M111.265819

Magnelli, P., Cipollo, J. F., & Abeijon, C. (2002). A refined method for the determination of Saccharomyces cerevisiae cell wall composition and beta-1,6-glucan fine structure. *Anal Biochem, 301*(1), 136-150. doi:10.1006/abio.2001.5473

Martins, N., Ferreira, I. C., Barros, L., Silva, S., & Henriques, M. (2014). Candidiasis: predisposing factors, prevention, diagnosis and alternative treatment. *Mycopathologia, 177*(5-6), 223-240. doi:10.1007/s11046-014-9749-1

Matheson, K., Parsons, L., & Gammie, A. (2017). Whole-Genome Sequence and Variant Analysis of W303, a Widely-Used Strain of Saccharomyces cerevisiae. *G3 (Bethesda), 7*(7), 2219-2226. doi:10.1534/g3.117.040022

Nagashima, M., Yin, Z. F., Broze, G. J., Jr., & Morser, J. (2002). Thrombin-activatable fibrinolysis inhibitor (TAFI) deficient mice. *Front Biosci, 7*, d556-568. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11815293

Nagata, T., & Takebe, I. (1970). Cell wall regeneration and cell division in isolated tobacco mesophyll protoplasts. *Planta, 92*(4), 301-308. doi:10.1007/BF00385097

Naggert, J. K., Fricker, L. D., Varlamov, O., Nishina, P. M., Rouille, Y., Steiner, D. F., Carroll, R. J., Paigen, B. J., & Leiter, E. H. (1995). Hyperproinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity. *Nature Genetics, 10*(2), 135-142. doi:10.1038/ng0695-135

Neiman, A. M. (2005). Ascospore formation in the yeast Saccharomyces cerevisiae. *Microbiol Mol Biol Rev, 69*(4), 565-584. doi:10.1128/MMBR.69.4.565-584.2005

Neiman, A. M. (2011). Sporulation in the budding yeast Saccharomyces cerevisiae. *Genetics, 189*(3), 737-765. doi:10.1534/genetics.111.127126

Nguyen, T. H., Fleet, G. H., & Rogers, P. L. (1998). Composition of the cell walls of several yeast species. *Applied Microbiology and Biotechnology, 50*(2), 206-212. doi:10.1007/s002530051278

Nijman, S. M. (2011). Synthetic lethality: general principles, utility and detection using genetic screens in human cells. *FEBS Lett, 585*(1), 1-6. doi:10.1016/j.febslet.2010.11.024

Noda-Garcia, L., & Barona-Gomez, F. (2013). Enzyme evolution beyond gene duplication: A model for incorporating horizontal gene transfer. *Mob Genet Elements, 3*(5), e26439. doi:10.4161/mge.26439

Ohno, S. (1970). *Evolution by gene duplication*. London, New York,: Allen & Unwin; Springer-Verlag.

Okonechnikov, K., Golosova, O., Fursov, M., & the, U. t. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics, 28*(8), 1166-1167. doi:10.1093/bioinformatics/bts091

Oliver, S. G., van der Aart, Q. J., Agostoni-Carbone, M. L., Aigle, M., Alberghina, L., Alexandraki, D., Antoine, G., Anwar, R., Ballesta, J. P., Benit, P., & et al. (1992). The complete DNA sequence of yeast chromosome III. *Nature, 357*(6373), 38-46. doi:10.1038/357038a0

Otazo, M. R. d. l. V., Lorenzo, J., Tort, O., Avilés, F. X., & Bautista, J. M. (2013). Functional segregation and emerging role of cilia-related cytosolic carboxypeptidases (CCPs). *The FASEB Journal, 27*(2), 424-431. doi:10.1096/fj.12-209080

Parzych, K. R., Ariosa, A., Mari, M., & Klionsky, D. J. (2018). A newly characterized vacuolar serine carboxypeptidase, Atg42/Ybr139w, is required for normal vacuole function and the terminal steps of autophagy in the yeast Saccharomyces cerevisiae. *Mol Biol Cell, 29*(9), 1089-1099. doi:10.1091/mbc.E17-08-0516

Pascual, R., Burgos, F. J., Salva, M., Soriano, F., Mendez, E., & Aviles, F. X. (1989). Purification and properties of five different forms of human procarboxypeptidases. *Eur J Biochem, 179*(3), 609-616. doi:10.1111/j.1432-1033.1989.tb14590.x

Peter, J., De Chiara, M., Friedrich, A., Yue, J. X., Pflieger, D., Bergstrom, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., Cruaud, C., Labadie, K., Aury, J. M., Istace, B., Lebrigand, K., Barbry, P., Engelen, S., Lemainque, A., Wincker, P., Liti, G., & Schacherer, J. (2018). Genome evolution across 1,011 Saccharomyces cerevisiae isolates. *Nature, 556*(7701), 339-344. doi:10.1038/s41586-018-0030-5

Petrera, A., Lai, Z. W., & Schilling, O. (2014). Carboxyterminal protein processing in health and disease: key actors and emerging technologies. *J Proteome Res, 13*(11), 4497-4504. doi:10.1021/pr5005746

Pfaller, M. A., & Diekema, D. J. (2007). Epidemiology of invasive candidiasis: a persistent public health problem. *Clinical microbiology reviews, 20*(1), 133-163. doi:10.1128/CMR.00029-06

Pils, B., & Schultz, J. (2004). Inactive enzyme-homologues find new function in regulatory processes. *J Mol Biol, 340*(3), 399-404. doi:10.1016/j.jmb.2004.04.063

Pink, R. C., Wicks, K., Caley, D. P., Punch, E. K., Jacobs, L., & Carter, D. R. (2011). Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA, 17*(5), 792-798. doi:10.1261/rna.2658311

Popolo, L., Gualtieri, T., & Ragni, E. (2001). The yeast cell-wall salvage pathway. *Med Mycol, 39 Suppl 1*, 111-121. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11800265

Ralser, M., Kuhl, H., Ralser, M., Werber, M., Lehrach, H., Breitenbach, M., & Timmermann, B. (2012). The Saccharomyces cerevisiae W303-K6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt. *Open biology, 2*(8), 120093-120093. doi:10.1098/rsob.120093

Ram, A. F. J., Wolters, A., Hoopen, R. T., & Klis, F. M. (1994). A new approach for isolating cell wall mutants in Saccharomyces cerevisiae by screening for hypersensitivity to calcofluor white. *Yeast, 10*(8), 1019-1030. doi:10.1002/yea.320100804

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., & Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res, 46*(D1), D624-D632. doi:10.1093/nar/gkx1134

Reiterer, V., Eyers, P. A., & Farhan, H. (2014). Day of the dead: pseudokinases and pseudophosphatases in physiology and disease. *Trends Cell Biol, 24*(9), 489-505. doi:10.1016/j.tcb.2014.03.008

Reznik, S. E., & Fricker, L. D. (2001). Carboxypeptidases from A to z: implications in embryonic development and Wnt binding. *Cell Mol Life Sci, 58*(12-13), 1790-1804. doi:10.1007/PL00000819

Ribeiro, A. J. M., Das, S., Dawson, N., Zaru, R., Orchard, S., Thornton, J. M., Orengo, C., Zeqiraj, E., Murphy, J. M., & Eyers, P. A. (2019). Emerging concepts in pseudoenzyme classification, evolution, and signaling. *Sci Signal, 12*(594). doi:10.1126/scisignal.aat9797

Roncero, C., & Duran, A. (1985). Effect of Calcofluor white and Congo red on fungal cell wall morphogenesis: in vivo activation of chitin polymerization. *J Bacteriol, 163*(3), 1180-1185. doi:10.1128/JB.163.3.1180-1185.1985

Ruderfer, D. M., Pratt, S. C., Seidel, H. S., & Kruglyak, L. (2006). Population genomic analysis of outcrossing and recombination in yeast. *Nat Genet, 38*(9), 1077-1081. doi:10.1038/ng1859

Rueden, C. T., Schindelin, J., Hiner, M. C., DeZonia, B. E., Walter, A. E., Arena, E. T., & Eliceiri, K. W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics, 18*(1), 529. doi:10.1186/s12859-017-1934-z

Sadowski, I., Lourenco, P., & Parent, J. (2008). Dominant marker vectors for selecting yeast mating products. *Yeast, 25*(8), 595-599. doi:10.1002/yea.1604

Schoner, D., Kalisch, M., Leisner, C., Meier, L., Sohrmann, M., Faty, M., Barral, Y., Peter, M., Gruissem, W., & Buhlmann, P. (2008). Annotating novel genes by integrating synthetic lethals and genomic information. *BMC Syst Biol, 2*, 3. doi:10.1186/1752-0509-2-3

Schott, M. J. (2015). Expression and Characterization of ECM14, a Metallocarboxypeptidase from Saccharomyces Cerevisiae. *Master's Theses, 88*. Retrieved from https://digitalcommons.andrews.edu/theses/88/

Smith, M. L., Bruhn, J. N., & Anderson, J. B. (1992). The fungus Armillaria bulbosa is among the largest and oldest living organisms. *Nature, 356*(6368), 428-431. doi:10.1038/356428a0

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S., & Youngs, H. (2004). Toward a systems approach to understanding plant cell walls. *Science, 306*(5705), 2206-2211. doi:10.1126/science.1102765

Song, G., Dickins, B. J. A., Demeter, J., Engel, S., Gallagher, J., Choe, K., Dunn, B., Snyder, M., & Cherry, J. M. (2015). Correction: AGAPE (Automated Genome Analysis PipelinE) for Pan-Genome Analysis of Saccharomyces cerevisiae. *PLOS ONE, 10*(5), e0129184-e0129184. doi:10.1371/journal.pone.0129184

Spatafora, J., Sung, G.-H., Johnson, D., Hesse, C., O'Rourke, B., Serdani, M., Spotts, R., Lutzoni, F., Valerie, H., Miadlikowska, J., Reeb, V., Gueidan, C., Fraker, E., Lumbsch, T., Lücking, R., Schmitt, I., Hosaka, K., Aptroot, A., Roux, C., & Schoch, C. (2006). A five-gene phylogeny of Pezizomycotina. *Mycologia, 98*, 1018-1028. doi:10.3852/mycologia.98.6.1018

Sugiyama, J., Hosaka, K., & Suh, S. O. (2006). Early diverging Ascomycota: phylogenetic divergence and related evolutionary enigmas. *Mycologia, 98*(6), 996-1005. doi:10.3852/mycologia.98.6.996

Tan, F., Morris, P., Skidgel, R., & Erdös, E. (1993). Sequencing and cloning of human prolylcarboxypeptidase (Angiotensinase C). Similarity to both serine carboxypeptidase and prolylendopeptidase families. *The Journal of biological chemistry, 268*, 16631-16638.

Wach, A., Brachat, A., Pöhlmann, R., & Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. *Yeast, 10*(13), 1793-1808. doi:10.1002/yea.320101310

Wenger, R. M. (1984). Synthesis of cyclosporine. Total syntheses of 'cyclosporin A' and 'cyclosporin H', two fungal metabolites isolated from the species Tolypocladium inflatum GAMS. *Helvetica Chimica Acta, 67*(2), 502-525. doi:10.1002/hlca.19840670220

Wiederhold, E., Gandhi, T., Permentier, H. P., Breitling, R., Poolman, B., & Slotboom, D. J. (2009). The yeast vacuolar membrane proteome. *Mol Cell Proteomics, 8*(2), 380-392. doi:10.1074/mcp.M800372-MCP200

Winston, F. (2008). EMS and UV Mutagenesis in Yeast. *Current Protocols in Molecular Biology, 82*(1), 13.13B.11-13.13B.15. doi:10.1002/0471142727.mb1303bs82

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H., Chu, A. M., Connelly, C., Davis, K., Dietrich, F., Dow, S. W., El Bakkoury, M., Foury, F., Friend, S. H., Gentalen, E., Giaever, G., Hegemann, J. H., Jones, T., Laub, M., Liao, H., Liebundguth, N., Lockhart, D. J., Lucau-Danila, A., Lussier, M., M'Rabet, N., Menard, P., Mittmann, M., Pai, C., Rebischung, C., Revuelta, J. L., Riles, L., Roberts, C. J., Ross-MacDonald, P., Scherens, B., Snyder, M., Sookhai-Mahadeo, S., Storms, R. K., Veronneau, S., Voet, M., Volckaert, G., Ward, T. R., Wysocki, R., Yen, G. S., Yu, K., Zimmermann, K., Philippsen, P., Johnston, M., & Davis, R. W. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science, 285*(5429), 901-906. doi:10.1126/science.285.5429.901

Yang, Z., Huang, J., Geng, J., Nair, U., & Klionsky, D. J. (2006). Atg22 recycles amino acids to link the degradative and recycling functions of autophagy. *Mol Biol Cell, 17*(12), 5094-5104. doi:10.1091/mbc.e06-06-0479

Zaugg, C., Jousson, O., Lechenne, B., Staib, P., & Monod, M. (2008). Trichophyton rubrum secreted and membrane-associated carboxypeptidases. *Int J Med Microbiol, 298*(7-8), 669-682. doi:10.1016/j.ijmm.2007.11.005

Zettl, M., Adrain, C., Strisovsky, K., Lastun, V., & Freeman, M. (2011). Rhomboid family pseudoproteases use the ER quality control machinery to regulate intercellular signaling. *Cell, 145*(1), 79-91. doi:10.1016/j.cell.2011.02.047

Zhang, X., Cheng, W., Feng, Z., Zhu, Q., Sun, Y., Li, Y., & Sun, J. (2020). Transcriptomic analysis of gene expression of Verticillium dahliae upon treatment of the cotton root exudates. *BMC Genomics, 21*(1), 155. doi:10.1186/s12864-020-6448-9

Zörgö, E., Chwialkowska, K., Gjuvsland, A. B., Garré, E., Sunnerhagen, P., Liti, G., Blomberg, A., Omholt, S. W., & Warringer, J. (2013). Ancient Evolutionary Trade-Offs between Yeast Ploidy States. *PLOS Genetics, 9*(3), e1003388. doi:10.1371/journal.pgen.1003388