Humanized Yeast as a Platform to Measure the Functional Impact of Human Genetic Variation

Farhat Zafar

A

Thesis

In the Department of Biology

Presented in Partial Fulfillment of the Requirements for the Degree of

Master of Science (Biology) at

Concordia University

Montreal, Quebec, Canada

March 2021

School of Graduate Studies

This is to certify that the thesis prepared

By: **Farhat Zafar**

Entitled: **Humanized Yeast as a Platform to Measure the Functional Impact of Human Genetic Variation**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ External Examiner

Dr. Christopher Brett

_____Examiner

Dr. Michael Hallett

_____ Examiner, Chair

Dr. Laurent Potvin-Trottier

_____ Supervisor

Dr. Aashiq Kachroo

Approved by

_____Graduate Program Director

Dr. Robert Weladji

_____Dean of Faculty

Selvadurai Dayanandan

2021

Abstract

Humanized Yeast as a Platform to Measure the Impact of Human Genetic Variation

Farhat Zafar

Data from thousands of human genomes is currently available, revealing several thousand variations in our genes. A grand challenge is to determine what, if any, impact each missense mutation has on human health. The fact that many human genes can functionally replace their corresponding yeast equivalents provides an opportunity to use humanized yeast to model human genetic variation. Swapping essential yeast genes with their human counterparts allows us to link the fitness of the human protein variant with the fitness of the yeast cell. Therefore, mutations that have functional consequences will show phenotypic differences in the humanized yeast, such as slower growth rate.

We demonstrate this approach on the rare human genetic disorder known as mevalonate kinase deficiency by devising a strategy to conditionally replace yeast Sc*ERG12* gene with the orthologous human mevalonate kinase gene, Hs-*MVK*. Next, we demonstrate that yeast growth acts as an easily measured proxy for the proper functioning of the human gene. By humanizing yeast with gene variants followed by sequencing as a readout to deconvolute each strain's relative growth, we can score the functional impact of human gene variation. Thus, neutral variants are easily distinguished from deleterious variants. This work, in addition to previous similar strategies, establishes a platform for using humanized yeast to model human genetic variation at scale.

Acknowledgements

I would like to thank Dr. Aashiq Kachroo for his guidance over the past couple of years. Thank you for always being available to discuss science and answer my questions. And for encouraging me during times when I was discouraged from experiments not working as expected. This research would not have been possible without your help.

I would also like to thank my committee members Dr. Michael Hallett and Dr. Laurent Potvin-Trottier for providing guidance and feedback. And Dr. Christopher Brett for acting as external examiner.

Thank you to our collaborators Dr. Frederick Roth and Nishka Kishore from the Roth Lab at the University of Toronto.

A thank you to the Genome Foundry Platform Coordinators Dr. Smitha Amarnath and Nicholas Gold for their help using the foundry equipment.

Thank you to Dr. Chris Law for help using the Leica DM6000 Microscope.

Thank you to YuXiang Ren and Courtney Gamache for contributing to this work by creating mutant variants through site-directed mutagenesis.

A special thank you to all the Kachroo Lab members for your friendship, guidance, and support. There were many times when we brainstormed ideas or just wanted to vent our frustration over experiments that were not working. Those moments made my experience as a graduate student better. Thank you as well to all my friends from other labs at the Center for Applied Synthetic Biology.

Thank you to Dr. Orly Weinberg and Dr. Malcolm Whiteway for the opportunities they have provided through the NSERC-CREATE Applied Synthetic Biology Program.

Dedications

I would like to dedicate this thesis to my family. Thank you to my parents for always supporting me. I never felt any pressure from you about my studies. Thank you to my brother Ammad for always being there to listen to me talk about my research. I really would not have been able to complete this thesis on time if it were not for you driving me to the University during the pandemic so I could avoid taking public transport.

Table of Contents

List of Figures

List of Tables

Chapter 1: Background

1.1 Using Yeast to Measure the Functional Impact of Human Genetic Variation

In the past decade, sequencing technology has rapidly improved to increase in speed and precision while decreasing in cost. Currently, a human genome can be sequenced for around $1000[1]. From the bacteriophage Phi-X174 genome, which was the first genome to be sequenced, to the Human Genome Project's successful sequencing of the entire human genome in 2001, our knowledge of genomes continues to expand [2,3]. Next Generation Sequencing (NGS) based on the parallelization of the sequencing process has been a huge leap in this field[4]. These advances allow for the generation of massive amounts of sequencing data. A grand challenge of this field is to accurately annotate that information and determine what, if any, impact each sequence change has on human health.

Differences in our genomic sequences allow for human genetic variation. These differences can be categorised as normal polymorphism that do not have a significant effect on phenotype or they can be disease-causing. Identifying variants that are pathogenic can help assess disease severity and provide targeted therapy[5]. The extensive amount of variant data requires highthroughput pipelines for testing the functions of variants. One such pipeline, known as deep mutational scanning (DMS), can measure the impact of nearly every amino change in a protein on function[6]. Such studies are important to determine the impact of variants on human health even when they are yet to be identified in the human population.

The American College of Medical Genetics and Genomics (ACMG) standard is commonly used to categorize variants as benign, likely pathogenic, pathogenic, or variant of uncertain significance (VUS)[7]. The majority of variants in databases such as ClinVar, are

characterized as VUS[6]. This is primarily due to a lack of information about the functional impact of these variants. While computational models that use the structural information of the protein may be used to predict the functional impact of a variant, however, computational predictions are only one-third as accurate as functional assays at predicting the severity of a mutant[8]. A direct assay for the function of a protein in a living cell, such as complementation assays, could allow for a functional readout of a human protein and each of its variant's phenotype. Such assays can even be applied to genes of two diverged species, such as humans and yeast. In fact, our laboratory has already tested many human genes (one-at-a-time) to define a nearly complete set of humanizable essential yeast genes[9].

Complementation between human and yeast genes has been studied as early as 1985, when Kataoka et al. sought to replace yeast RAS proteins with homologous RAS proteins from humans [10]. In this study, they took advantage of the fact that at least one functional RAS gene is needed for growth of haploid yeast spores. Chimeric proteins were created in which part of the human RAS amino acid sequence was fused with the yeast RAS2 protein. Expression of this construct was controlled by the galactose-inducible promoter GAL10. Growth of spores on media containing galactose indicated successful complementation. The loss of endogenous RAS function was rescued by the chimeric yeast/human protein[10].

One such study by Sun et al. in 2016 sought to rescue temperature-sensitive mutations in 125 yeast essential genes by replacing them with 139 orthologous human genes [8]. A human gene successfully complementing a yeast gene would allow for growth of the mutant yeast at the non-permissive temperature. Failure to grow indicated a failure to complement. In this study, complementation occurred for 26 human disease genes. From these, 101 disease-associated and 78 non-disease-associated variants were constructed and tested. Functional complementation by

spotting and by growth curves in liquid media were compared and found to have high correlation[8]. Although my research focuses on complementation by spotting, it is worth noting that liquid complementation assays are also effective.

Despite a billion years of divergence between humans and yeast, we share thousands of orthologous genes with yeast[11]. About 961 human genes have identifiable yeast orthologs (**Figure 1**). Orthologs are genes that diverged due to a speciation event, whereas paralogs are genes that diverged due to duplication[12]. The corresponding yeast orthologs (702) are essential in yeast (absolutely required for cellular viability). Our laboratory, together with several groups, have systematically tested the replaceability of nearly every essential gene in yeast with their human equivalents identifying many human genes that complement the corresponding yeast gene function [8,9].This observation provides an opportunity to use humanized yeast to model human genetic variation.

The key idea is that genetic variants (missense and nonsense mutations) with functional consequences will show phenotypic changes in the humanized yeast, such as diminished fitness for loss-of-function variants. Such a trend has previously been exploited in yeast [6,9]. Taken together, these studies strongly support using humanized yeast growth fitness as an easily measurable proxy for the proper functioning of the human gene.

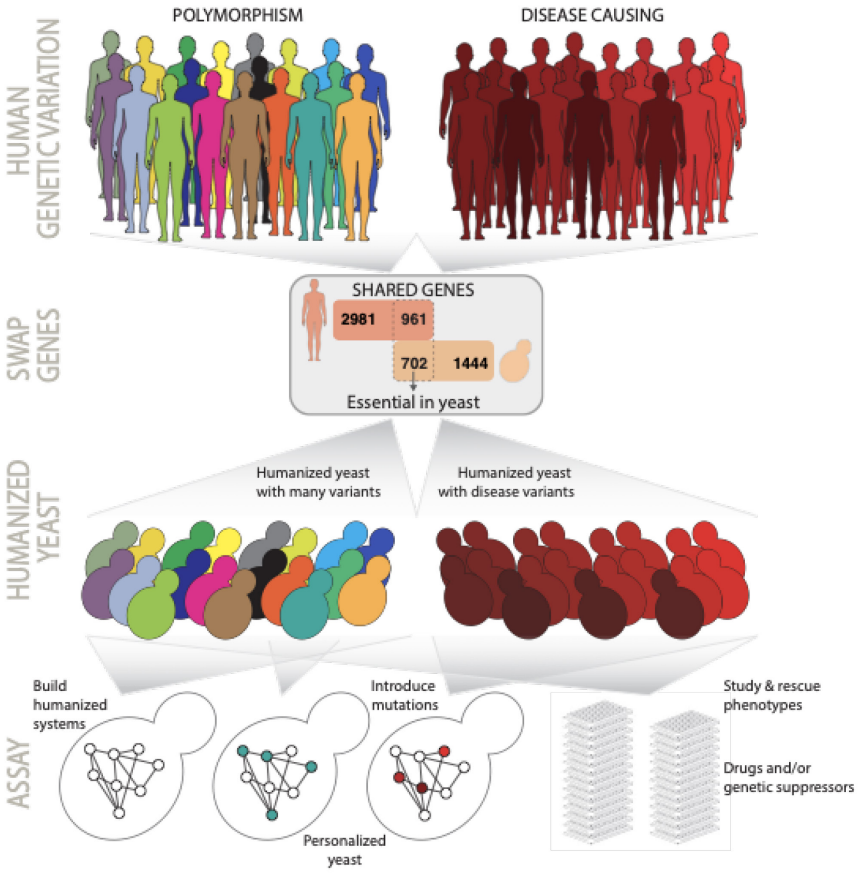

Figure 1. Humanization of Yeast to Study Variant Effect. Humanized yeast will act as test tubes to study human genetic variation and discover drugs or suppressor mutations that alleviate the defects in these genes providing a comprehensive understanding of the effect of human genetic variation.

Previously, our laboratory has shown that replaceable human genes show a strong trend for genes encoding proteins in the same pathway or complex to be similarly replaceable. Notably, we found that all genes except two of the cholesterol biosynthesis genes can be humanized in yeast[9] (**Figure 2**).

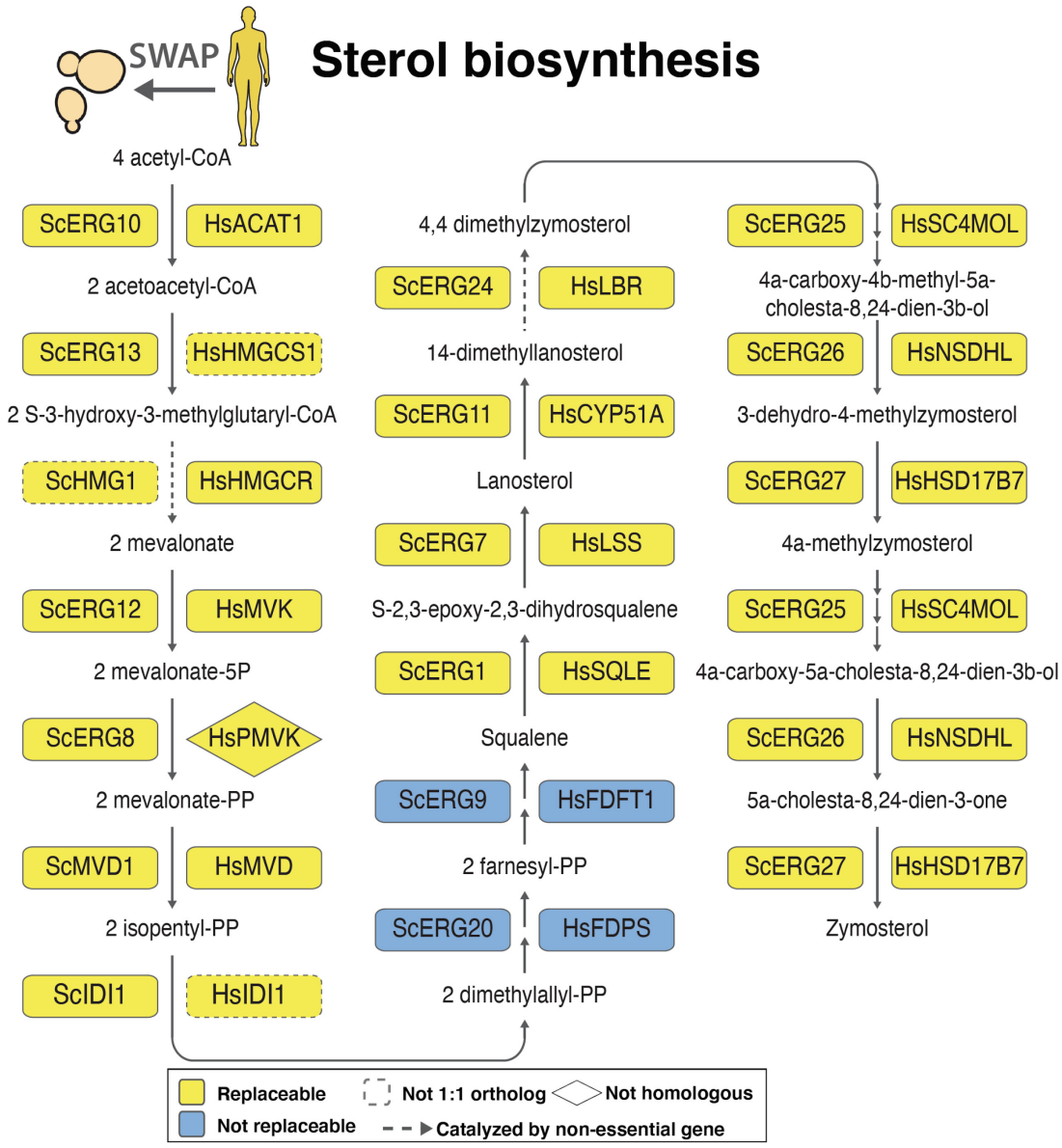

Figure 2. Functional Replaceability of Sterol Biosynthesis Pathway genes from Humans to Yeast. 17 of 19 sterol biosynthesis yeast genes were replaceable by their human counterparts. Genes highlighted in yellow represent successful replaceability. Genes highlighted in blue represent human genes that could not replace the yeast genes [9]

1.2 Human Mevalonate Kinase & Associated Diseases

Cholesterol is an important component of cell membranes. Cholesterol is obtained through diet and synthesized in the endoplasmic reticulum of liver cells [13]. This biochemical

process is carried out by the sterol biosynthesis pathway. Mutations in the genes of this pathway can lead to diverse human disease[14]. MVK (Human Mevalonate Kinase) can functionally replace the yeast ortholog (*ERG12*). My project focuses on characterizing the impact of mutations using a humanized yeast model.

MVK protein exists as a tetramer, with 3 ATP binding sites at positions 13, 55 and 135, two active sites at position 146 and 204 and two metal binding sites for magnesium at positions 146 and 193 (**Figure 3**)[15]. Mevalonate Kinase phosphorylates mevalonic acid to make mevalonate phosphate, which leads to the production of isopentenyl pyrophosphate (IPP), non-sterol isoprenoids and cholesterol.

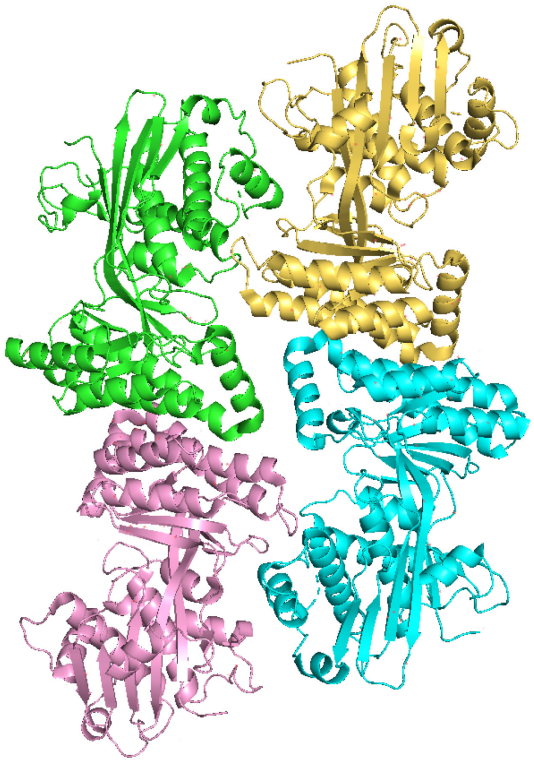

Figure 3. Visualization of Mevalonate Kinase Protein Structure.
Functional human MVK protein is a homo tetramer (PDB number 2R3V)[16]. Each monomer consists of 396 amino acids. Protein structure has been rendered using Pymol software. Each monomer is represented by a different color.

Mutations in human MVK that result in a decreased kinase activity lead to an accumulation of mevalonic acid and decreased production of IPP and non-sterol isoprenoids[17]. Decrease in the production of IPP can have several consequences since IPP in involved in the production of the isoprenoids farnesyl pyrophosphate (FPP) and geranylgeranyl pyrophosphate, which allow protein prenylation to help anchor proteins to the cell membrane [18]. One such protein, the GTPase RhoA remains unprenylated due to decreased IPP and leads to increased secretion of proinflammatory cytokines such as IL-1β [19]. These cytokines cause autoinflammatory symptoms. IPP is also involved in tRNA modification and synthesis of heme, ubiquinone and cholesterol[18]. A functional cholesterol biosynthesis pathway is necessary for human health.

Many mutations reported in the MVK gene lead to mevalonate kinase deficiency (MKD). This can clinically manifest as either hyperimmunoglobulinemia D and periodic fever syndrome (HIDS) or mevalonic aciduria (MA)[20]. These are rare autoinflammatory and autosomal recessive disorders. HIDS gets its name from the increased IgD levels that can be found in patients with MKD[21]. However, this is not an accurate diagnostic tool because IgD levels can be normal in up to 20% of patients[22]. Instead, mevalonic acid concentration is a better measure of the severity of the disease. HIDS patients may have moderately increased levels of mevalonic acid in their urine, whereas MA patients will have high concentration of mevalonic acid in their plasma and urine[23]. HIDS is the common form and usually associated with less severe symptoms, whereas MA is rare and associated with severe symptoms[17]. One of the characteristic symptoms of both HIDS and MA are recurrent fevers that can exceed 40°C[17]. Most patients experience their first inflammatory attack within the first year of their life. For

HIDS patients, these typically last 3-7 days[24]. Symptoms of HIDS include: fever, gastrointestinal pain, diarrhea, skin lesions, and lymphadenopathy[23]. Symptoms of MA include fever, gastrointestinal discomfort, skin lesions, psychomotor retardation, and facial dysmorphia[23]. While there is overlap in the symptoms, the main distinction for MA is that it involves the central nervous system[25]. In severe cases of MA, it can even lead to early childhood death[18]. There is currently no approved treatment for MKD[20]. While MKD is rare, most of the cases have been found in the Netherlands. The founder mutation Val377IIle (V377I) is thought to have originated in the Dutch population[26]. A founder mutation refers to a high frequency of alleles for a particular population, usually due to that population starting off as a small group with low genetic variation[27]. V377I is the most common variant found in MKD patients and is known to be pathogenic. It occurs on the C-terminal of the MVK gene and is associated with HIDS[18]. Although Dutch cases of MKD are the most prevalent, this deficiency can affect people of different ethnicities. A recent study of MKD in India described the symptoms of ten children with HIDS[28]. Five of those patients carried the V377I mutation in a compound heterozygote state[28]. In fact, most clinical cases of MKD have been identified as compound heterozygotes[23]. Compound heterozygous state refers to two different mutant alleles of the gene in diploid state. It is not clear if the compound heterozygous nature of the MVK alleles contributes to the disease phenotype.

Due to the wide range of MKD's phenotypic severity, MVK is a promising candidate for studying human disease in a simplified organism such as yeast. The difference in phenotypic symptoms is linked to the reduced mevalonate kinase activity[18]. The enzymatic activity will differ depending on the variant. MVK has a range of variants that can exhibit different

phenotypes. Human MVK can functionally replace the essential yeast ortholog, *ERG12 [9]*. Knocking out *ERG12* prevents growth in yeast cells. Replacing the function of the knocked out *ERG12* with a wild-type copy of human MVK allows for restoration of growth. By swapping essential *ERG12* yeast gene with human MVK, we would couple the fitness of that particular yeast gene with the fitness of the yeast cell.Therefore, mutations that have functional consequences will show phenotypic differences in the humanized yeast, such as slower growth rate. My project focuses on this important problem. Is it possible to use growth as a readout for pathogenicity in humanized yeast models? This is important since the deluge of data in terms of human genomes sequenced has resulted in an exponential increase in the observation of human gene variants in MVK. Nearly 3000 missense, nonsense and other coding sequence mutations have been deposited in the NCBI Variant Viewer database [29]. Except for a few that are reported as pathogenic (31) or likely pathogenic (10), it is unknown what each of these variants contributes to the function of MVK.

1.3 Humanized Yeast as a Model for Studying the Impact of Genetic Variation

Baker's yeast, *Saccharomyces cerevisiae*, continues to be an exceptional model organism [30]. Its genome has been sequenced since 1996 and consists of ~6000 genes[31]. Several knockout collections are commercially available, allowing for genetic studies [32]. It can exist in a haploid or diploid state. Mating between A and α types can produce diploids[33]. While sporulation under nutrient poor conditions can lead to the production of tetrads consisting of four haploid cells[33]. Yeast's proficiency with homologous recombination is one of the reasons it is useful for genetic manipulation. By adding homology arms to a gene of interest, it can integrate into the targeted part of the genome[34]. Additionally, studies have demonstrated that yeast can

express foreign DNA from organisms such as bacteria, plants, and humans ([35]). While it may seem like a better idea to work with mammalian cells when studying human disease, maintaining mammalian cell cultures is far more challenging than yeast. Yeast has a short generation with a doubling time of 90 minutes[33]. Cell cultures can easily be stored at -80°C in glycerol for long term storage. As eukaryotes, yeast cells also have protein folding, proteolytic processing and glycosylation, similar to mammalian cells[34]. In combination with low cost experimentation, rapid growth, many selectable genetic markers, and simple conversion between haploid and diploid forms, yeast is an ideal system for *in vivo* chromosome-scale manipulation and editing. The introduction of human genes is simple and efficient, and allows us to bring all of the remarkable benefits of using yeast cells to the study of human genetics.

Our preliminary data confirm that some clinical variants of the human mevalonate kinase (MVK) linked to autoinflammatory diseases [36–38] fail to complement the loss of yeast *ERG12* as well as high-frequency human alleles of MVK ([38]) (**Figure 4**). However, a few critical variants, such as V377I and H20P, showed no marked difference in their growth phenotype, when compared to the wild-type MVK gene. Our previous assays for testing human gene variants used a galactose-inducible promoter to control the expression of human MVK using a 2-micron high copy plasmid while using doxycycline to control yeast gene expression (Tet-Off system) [39]. Galactose promoters can induce expression by 1000-fold under high galactose conditions, making it useful for overexpression assays[40]. However, galactose promoters are repressed by dextrose[41]. Yeast cells must be grown without dextrose if we want to avoid a reduction in expression. Changing the carbon source for yeast cells may have an effect on cellular metabolism, and it is hard to determine whether phenotypic changes are due to this

change[40]. Moreover, galactose promoters confer a population effect whereby not all the cells in the population show expression at the same level. Under certain conditions (such as a mixture of glucose and galactose), some cells may have their expression completely turned off, while others are highly expressed. It is, therefore, important to use a more tunable control of expression.

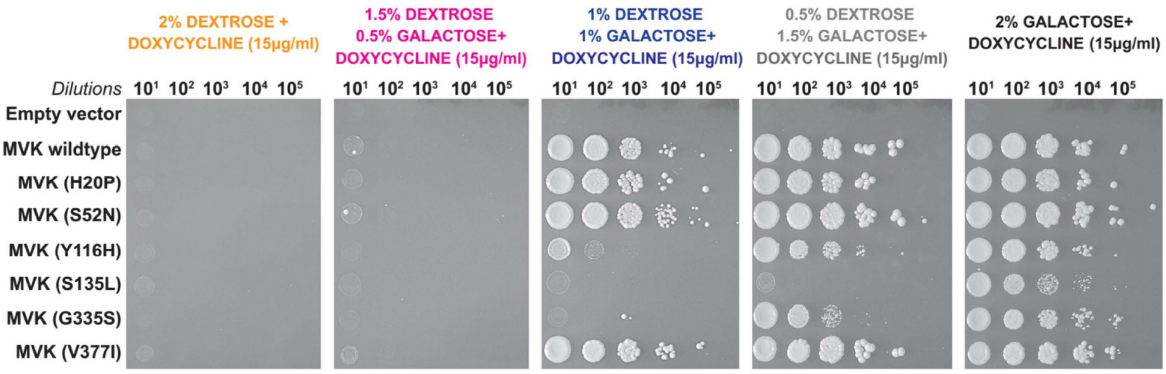

Figure 4. Complementation Assay of Several Human MVK Variants in Yeast. The yeast ortholog *ERG12* is controlled by pTET promoter that is turned off using 15µg/mL doxycycline. Expression of human MVK (pGAL-MVK) is controlled by galactose concentrations in the growth media. Galactose concentration is increasing in the columns to the right. The empty vector control which lacks the human MVK gene, never shows growth since both human and yeast orthologs are being repressed. The column in the center has a concentration of 1% dextrose, 1% galactose. In this condition, we can see that three of the mutants: Y116H, S135L and G335S exhibit a severe growth defect when replacing the *ERG12* gene. These mutations are linked to severe forms of MKD disease in humans [36,37]. The column to the very right represents an overexpression condition, in which a high concentration of galactose leads to overexpression of the MVK gene. In this condition, nearly all mutants allow for growth due to expression levels being high.

Although this experiment provides valuable information on the severity of MVK variants, control of the human gene by galactose may lead to some issues, specifically the non-uniform expression at the population level. Therefore, in my research, the mode of control will be switched. Galactose is used to control the yeast gene, since it only needs to be completely expressed or repressed. While doxycycline is used to control the human gene, in which expression can be controlled linearly by adjusting the concentration of the antibiotic.

1.4 Tetracycline-Regulated Systems

The tetracycline system used in this study is based on the Tn10 tetracycline resistance operon found in *E. coli*. Although the Tet repressor (TetR) is natively a repressor in bacteria, it has been modified to induce expression in eukaryotic cells[34]. By fusing the TetR to the VP16 transactivator of herpes simplex virus, a Tet activator (tTA) now serves as transcriptional activator[40]. Tetracycline-regulated systems can exist as a Tet-Off system or a Tet-On system. If the Tet transactivator drives expression in the presence of tetracycline, it is referred to as a Tet-On system[42]. If the Tet transactiator drives expression in the absence of tetracycline, it is referred to as a Tet-Off system[42]. In my research, I work with a Tet-Off system. Instead of tetracycline, its more stable derivative doxycycline is used (**Figure 5**).

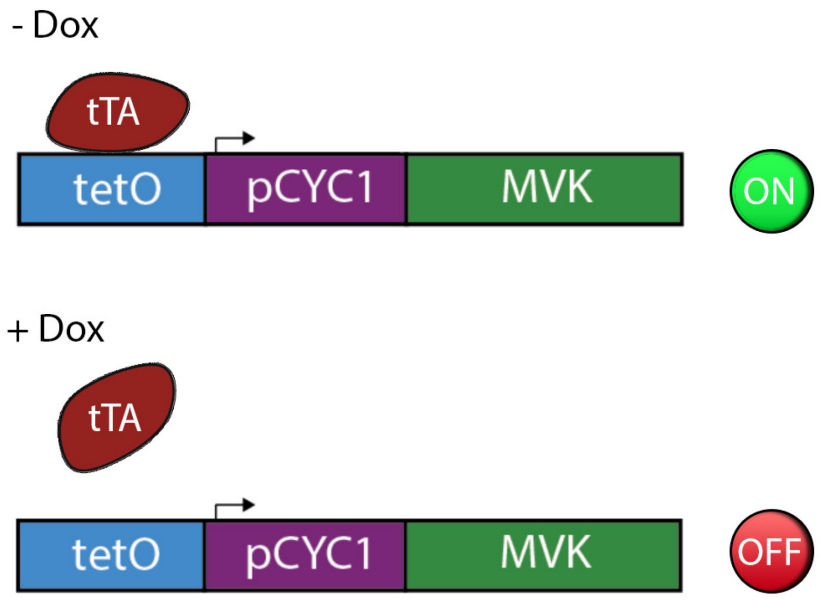

Figure 5. Overview of Tet-Off System for Human Gene Expression in Yeast. A Tetracycline-regulated system will be constructed that can turn the human gene expression ON or OFF in the absence or presence of doxycycline respectively. Interaction between the tetracycline transactivator and tetracycline operator drives expression of the human Mevalonate Kinase gene. Without this interaction, the minimal pCYC1 promoter will express the gene at low levels. Different amounts of doxycycline allows linear control of transcription from this promoter [40].

In the presence of doxycycline, the conformation of the transactivator is changed so it can no longer bind to a Tet Operator (tetO) and drive expression[34]. In this condition, our gene of interest is essentially "turned OFF". When doxycycline is absent, the transactivator can bind to the tetO and induce expression. In this condition, our gene of interest is "turned ON" (**Figure 5**). In addition to the wild-type tetO (O stands for an Operator), there also exist several mutants. The affinity of the binding site can be affected by the mutations. A previous study found that the wild-type tetO is the most efficient at driving expression[34]. The tetO I work with in my research consists of two repeats of the wild-type tetO.

1.5 Galactose-Inducible Promoters

Dextrose is the preferred carbon source for yeast under laboratory conditions[43]. In the presence of dextrose, yeast genes involved in the metabolism of other carbon sources are repressed such as genes involved in galactose metabolism[41]. The galactose promoter used in this study is GAL1. This promoter comes from the GAL1 gene, which is part of the galactose metabolism process, one of the processes repressed by dextrose. The carbon source raffinose does not repress the GAL1 promoter, but it does not induce expression[44]. Therefore, expression is above basal level in the presence of raffinose. Galactose is used to induce expression of the GAL1 promoter and can increase expression by more than 1000-fold compared to expression in the presence of dextrose[40]

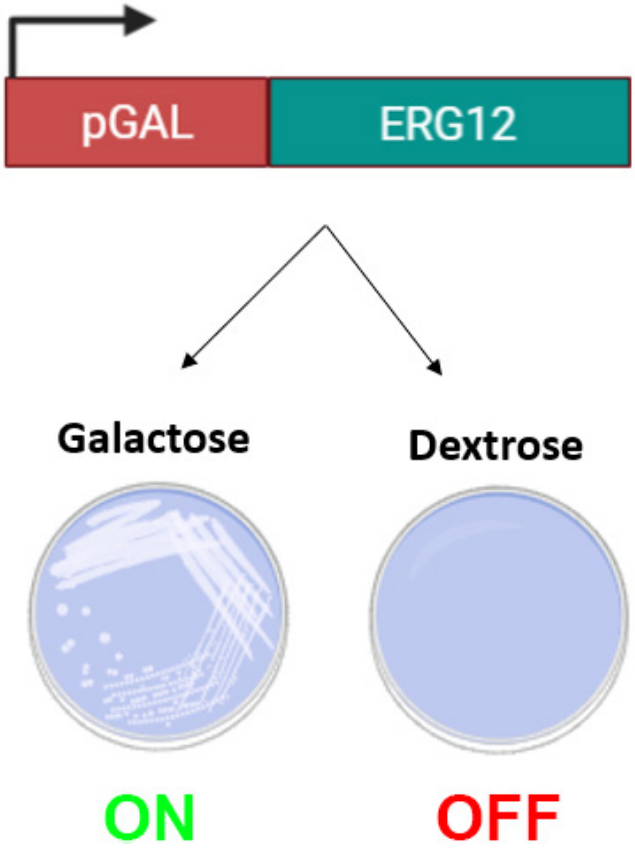

Figure 6. Overview of Galactose-Induced Control of *ERG12* Gene.

Yeast *ERG12* gene is essential for the viability of the cell. If the native promoter for the yeast *ERG12* gene is replaced by a GAL promoter, the gene can be turned ON or OFF by changing the carbon source in the medium. In the presence of galactose, the gene can be turned ON resulting in the growth of the yeast whereas in the presence of Dextrose, pGAL promoter is turned OFF resulting in the absence of growth.

In this study, the GAL1 promoter is used to control expression of the yeast ortholog *ERG12*. By having high concentration such as 2% dextrose in the media, the *ERG12* gene is fully repressed. Since *ERG12* is an essential gene, the presence of dextrose should result in the lack of growth on media. However, in the presence of 2% galactose (or mixture of galactose and dextrose), the *ERG12* gene is highly expressed and should allow yeast growth.

Chapter 2: Designing the human and yeast gene ON/OFF system in yeast

2.1 Yeast Strain harboring *ERG12* Gene under Galactose-Inducible Promoter

We wanted to build a system where expression of the yeast gene *ERG12* can be controlled by galactose. To achieve this, the native promoter of *ERG12* was replaced with a galactose-inducible promoter. Since ERG12 is an essential gene, this would only allow the modified strain to grow on galactose whereas no growth should be observed with dextrose in the medium. Using this strategy, an ON/OFF switch for the *ERG12* gene was created (**Figure 6**). We used conventional homologous recombination to insert a galactose-inducible promoter close to the start codon of *ERG12* (**Figure 7A**). Furthermore, for the selection of the engineered strain, the insert also contains a kanamycin cassette which allows for growth on the antibiotic G418 (**Figure 7A**). By inserting the galactose-inducible promoter and kanamycin cassette between the 5'UTR and coding region of *ERG12*, the native promoter is disrupted and expression of *ERG12* is controlled by galactose.

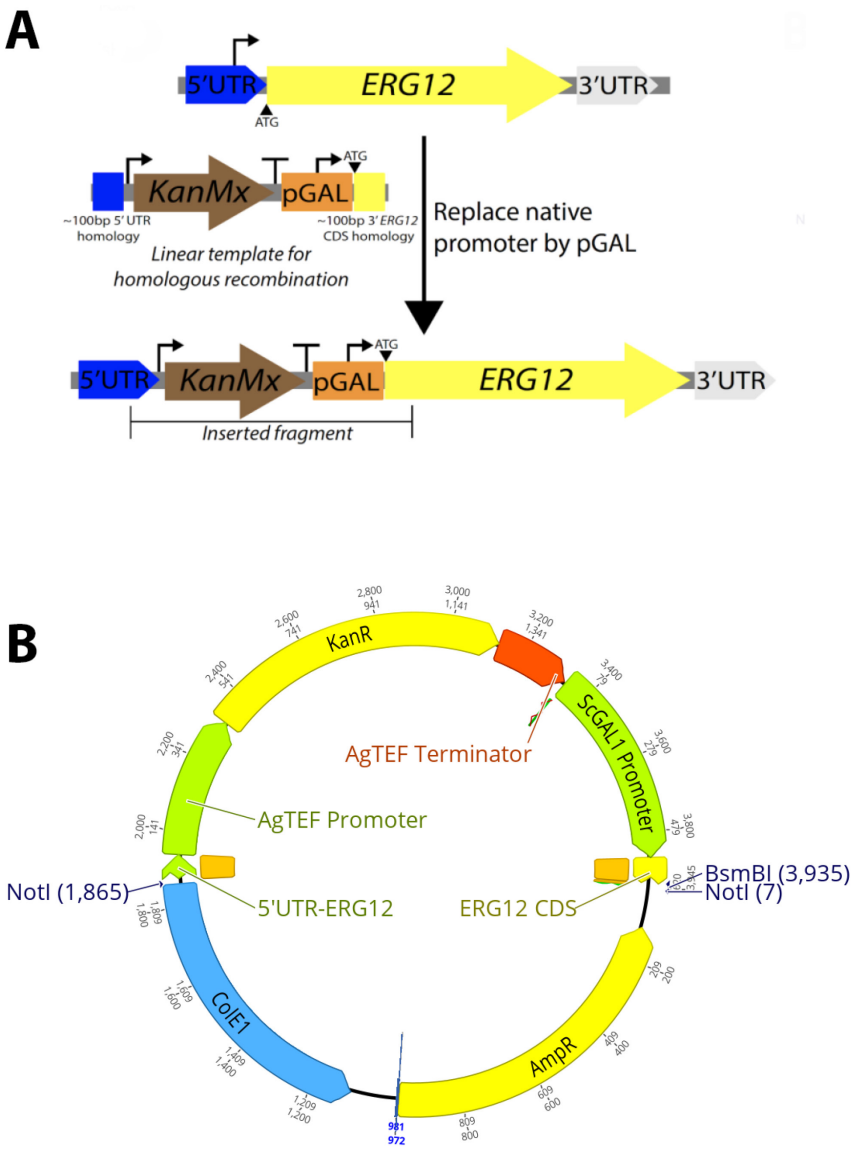

Figure 7. Insertion of Galactose Promoter with Kanamycin selection to replace the native *ERG12* promoter in yeast. (A) Top panel shows the native *ERG12* yeast locus with 5' and 3' untranslated regions (UTRs). The linear fragment of DNA to be inserted at the *ERG12* locus harbors homology to the 5' UTR (left) and the *ERG12* CDS (coding sequence) (right). The fragment inserts a pGAL promoter right next to the yeast gene. The cassette also contains an expression unit for *KAN* resistance gene. The successfully engineered strains will,therefore, grow on media containing G418. **(B)** The assembly of the template for homologous recombination at the *ERG12* locus was carried out using Golden Gate strategy. The fragments were obtained from MoClo toolkit [45]. The linear insert can be generated using NotI and BsmbI restriction enzymes. This linear fragment is used for homologous recombination of the native yeast *ERG12* locus. Additional homology to 5'UTR of *ERG12* (left end) and 3' *ERG12* CDS regions (right end) were added using PCR (~100bp sequence homology at either end).

The pGAL1 was amplified from pYTK030 and kanamycin cassette from pYTK077 of the Deuber Golden Gate Collection [45]. The kanamycin cassette consists of a kanamycin resistance gene, as well as a promoter pAgTEF and a terminator tAgTEF (these sequences have no homology to any yeast locus). Homology for the *ERG12* locus was added to the 3' and 5' ends of this construct and it was assembled into the pYTK095 vector (**Figure 7B**). Homology for 5'UTR-*ERG12* was added to the forward primer used to amplify the kanamycin cassette. Homology for 3'UTR-*ERG12* was added to the reverse primer used to amplify the pGAL1 promoter. At first, the homology region consisted of 65bp. After assembly into pYTK095, a new set of primers was designed to add additional 32bp of homology on each end. The added homology led to successful insertion at the native yeast locus.

Cut sites for NotI and BsmBI flanked the homology sequences, allowing for digestion and release of a linear insert (**Figure 7B**). After digestion, the DNA was added to competent cells from a BY4741 (Haploid Mat A) strain of yeast and transformed using the EZ Yeast Transformation Kit. A recovery step of growth in YPG (Yeast Extract, Peptone & Galactose) for several hours was added to the protocol. The cells were then plated on YPG media containing G418, allowing for selection of colonies that could only grow if they acquired the pGAL-*ERG12* promoter as well as the kanamycin cassette. Confirmation PCR was performed on the resulting colonies and they were sent for sequencing (**Figure S3**). Sequencing results confirmed that the construct integrated into the yeast genome through homologous recombination at the *ERG12* locus (**Figure S3C**). The confirmation primers were designed to validate correct insertion at both the 5' and 3'ends (**Figure S3B**).

The strain was characterized by growing it on three types of sugars: galactose, raffinose and dextrose (**Figure 8**). *ERG12* is an essential gene in yeast and its expression is therefore necessary for growth. The engineered strain, designed to control *ERG12* expression by galactose, should only grow in the presence of galactose. Dextrose is expected to repress the galactose-inducible promoter and prevent growth. The sugar raffinose does not repress the galactose-inducible promoter, but it does not induce it either. We expect little to no growth on Raffinose. We decided to test the engineered strain providing these three carbon sources in the medium. We noticed a rather unexpected behavior. The strain that was previously grown on galactose, could also grow on dextrose or raffinose (pGAL OFF state) without any obvious growth defect (**Figure 8, Left panel**). Therefore, we postulated that the growth profile of the culture appeared to also depend on what kind of sugar the cells had previously been grown in prior to inoculation in the new media. pGAL is a strong promoter and the *ERG12* gene may have been expressed at the highest levels in the galactose condition. The expression could have lasted well into the dextrose condition. The phenomenon is similar to the memory mechanisms observed in the *lac* operon of *E. coli [46]*. For example, a cell culture grown overnight in galactose would continue to grow well when transferred to media containing any of the dextrose and raffinose sugars as the pGAL-*ERG12* either continues to stay in ON-state or the protein is stable across generations. To test this behavior, we proposed to pre-culture the strain in dextrose or raffinose that should allow the expression of *ERG12* to be low enough such that no growth is observed in subsequent growth in these sugars. Indeed, we show that pGAL-*ERG12* yeast cells pre-cultured in raffinose and dextrose showed the kind of growth pattern we were expecting, i.e., in the presence of galactose, yeast would grow and in the presence of dextrose or raffinose,we observed no growth (**Figure 8, Middle & Right panel**) .

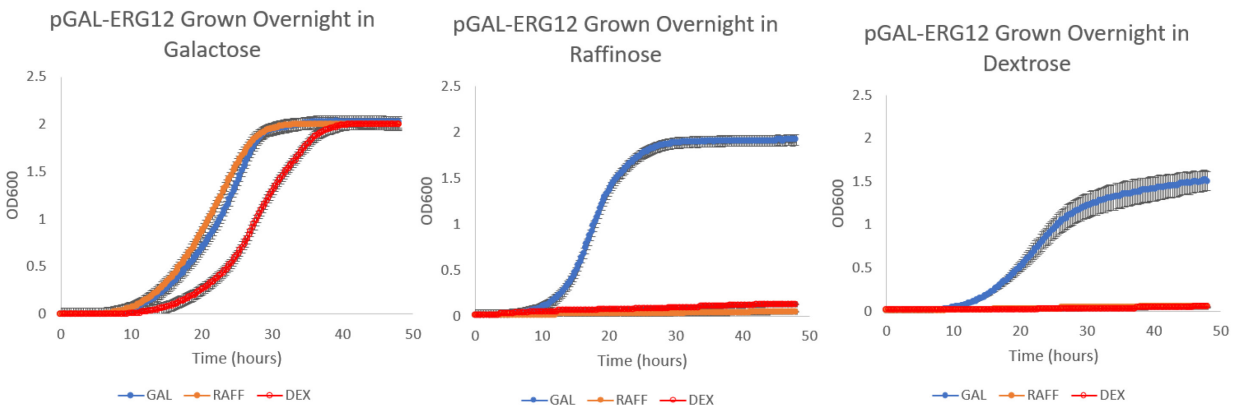

Figure 8. Growth Profile for pGAL-*ERG12* strain dependents on the carbon source in a pre-culture. The pGAL-*ERG12* cells grown in galactose medium overnight (pGAL ON state), showed no obvious growth defect when grown in the presence of dextrose or raffinose (Left panel, pGAL OFF state). However, pGAL-*ERG12* yeast cells pre-cultured in raffinose or dextrose showed no growth when subsequently grown in the presence of Raffinose or Dextrose, while showing normal growth in the presence of galactose (Middle and Right panels respectively).

2.2 Creation of a Tetracycline Operator Plasmid to Control Human Gene Expression

For tunable control of the human MVK gene expression, we wanted to construct a Tet-Off system which allows for repression of the gene in the presence of doxycycline (**Figure 5**). Such a system had already been designed in a plasmid PCM188[40]. This plasmid is commercially available through ATCC. However, PCM188 did not behave well in our lab and often showed aberrant size when tested using restriction enzyme digestion. We decided to make our own constructs. Two important components that could be amplified from PCM188 to make our own plasmid were the tetracycline operator (tetO) and the tetracycline transactivator (tTA). We developed a strategy to clone the tetO into a new Gateway compatible plasmid and integrate tTA into the previously constructed pGAL-*ERG12* strain.

The goal was to construct a plasmid that would be compatible with LR Gateway cloning [47], could be used as a shuttle vector to be expression in both *E.coli* and *S. cerevisiae*, contained the tetO, a tADH1 terminator upstream of the operator and a minimal CYC1 promoter for expression of any heterologous gene (in my case, human *MVK*). For this, we used a plasmid pAG416GPD-ccdB from the Susan Lindquist Yeast Gateway Kit (**Figure 9A**) ([47]). This plasmid serves as a destination vector and is LR Gateway compatible and contains all the other features we needed except that it did not have tetO-pCYC promoter.

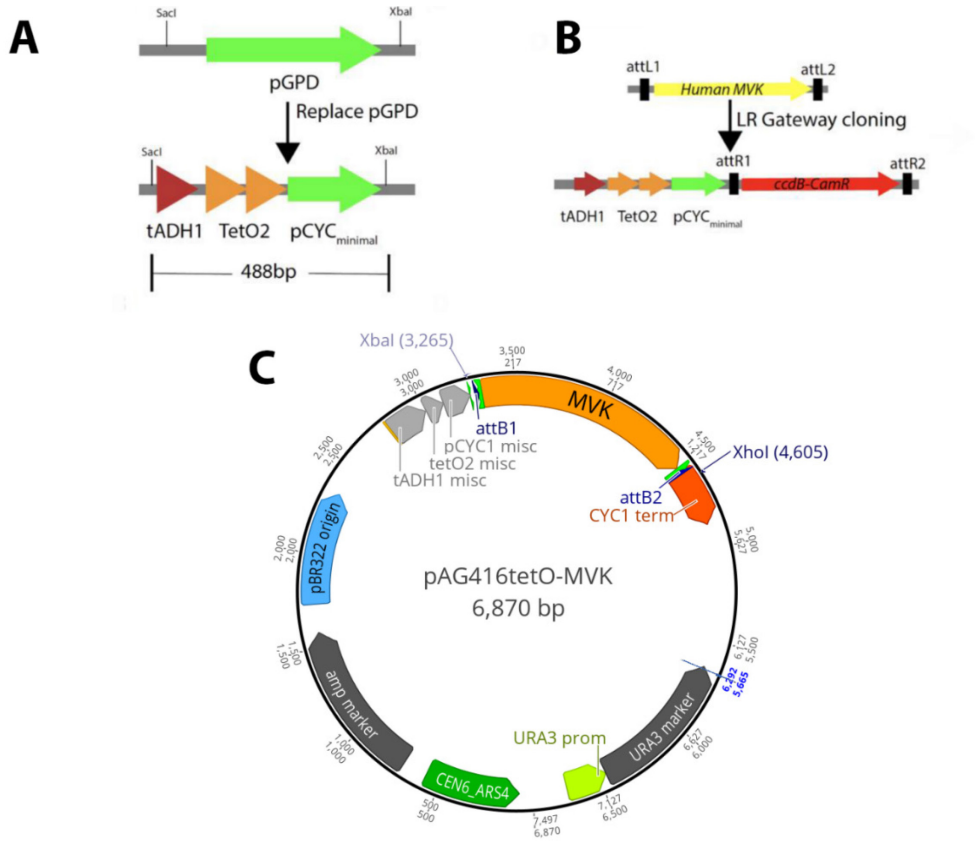

Figure 9. Design & construction of human *MVK* yeast expression vector under the control of Tet operator. (**A**) pGPD region was removed from the plasmid pAG416GPD-ccdB using flanking restriction enzyme sites (SacI & XbaI). The DNA fragment harboring tADH1-TetO2-pCYC$_{minimal}$ was PCR amplified with SacI/XbaI sites on 5' and 3' ends respectively. The Restriction enzyme digestion followed by ligation resulted in a TetO2-pCYC promoter containing yeast destination vector. (**B**) Human *MVK* gene flanked by attL1 & attL2 was used to replace the toxic ccdB-CamR cassette to generate a (**C**) yeast expression vector with human MVK under the control of Tet regulation.

We replaced the GDP promoter in pAG416GPD-ccdB by tADH1-tetO-pCYC1 from PCM188 using SacI/XbaI restriction enzymes (**Figure 9)**. To do this, we designed primers that contained SacI and XbaI restriction sites which were complementary to sites flanking the GDP promoter of pAG416GPD-ccdB **(Figure 9A)**. A pair of primers were used to amplify tADH1-tetO-pCYC1 from PCM188. Restriction enzyme cloning was used to create the new plasmid pAG416tetO-ccdB. Sequencing data confirmed successful cloning. This plasmid could now be used to clone in variants of MVK through LR Gateway cloning (**Figure 9B & C**).

2.3 Engineering a Strain Containing a Tetracycline Transactivator Expression Unit

The other important component for the tetracycline-regulated system is the tetracycline transactivator (tTA). We decided to genomically integrate tTA in yeast. The Golden Gate cloning method was used to amplify tTA from PCM188 flanked by BsmBI sites. Next, a strong constitutive yeast promoter pTDH3 and terminator tADH1 were assembled with the tTA. The promoter was taken from pYTK009 and the terminator from pYTK053 of Deuber's Golden Gate Kit[45]. After amplification of tTA, it was assembled into the entry vector pYTK001 (**Figure 10A**), verified by restriction enzyme digestion and Sanger sequencing. Next, we assembled a transcription unit using five parts, pTDH3, tTA, tADH1, along with ConS and ConE into pYTK095 (**Figure 10B**). The correct assembly of the tTA transcription unit (**Figure 10C**) was confirmed by restriction enzyme digestion and after the correct clone was confirmed, the parts were moved to the yeast integration vector pYTK097 (**Figure 10D**). This vector contains 5' and 3' homology regions for the mutant yeast *leu2* locus while also providing a functional copy of the *LEU2* gene. After integration, the strain becomes Leu+ and can be selected for on media that

lack leucine. NotI restriction sites flank the homology regions in pYTK097 and can be used to digest the plasmid to obtain our fragment of interest. The digested DNA is then added to competent cells made from the pGAL-*ERG12* strain, transformed using the EZ Yeast Transformation kit and grown on media containing galactose and G418, but lacking leucine. Several of the resulting colonies were picked and verified using colony PCR. One of the colonies (colony 3) confirmed integration and was used to test the new strain for tetracycline regulation of the human gene. This strain grows on galactose, contains G418 selection and drives the expression of human MVK by binding to the tetO on the pAG416-tetO-MVK plasmid. We used this plasmid to generate yeast expression vectors of the wild type and many variants of human MVK gene.

In **Chapter 3**, we tested if this newly engineered pGAL-Yeast gene/pTET-Human gene ON/OFF system allows the control of human gene expression and distinction between disease-causing and common variants in yeast.

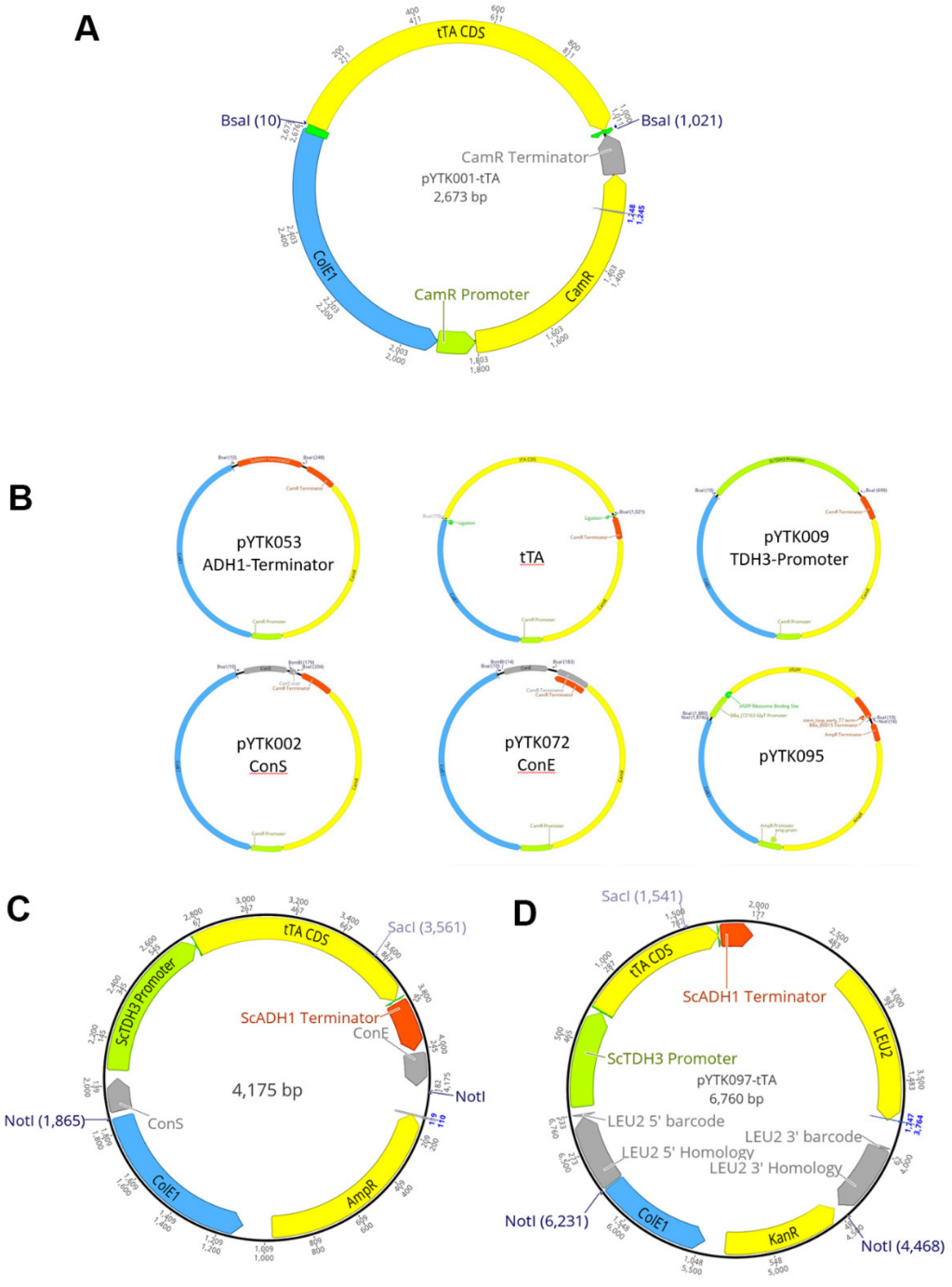

Figure 10. Cloning Tetracycline Transactivator into Entry Vector pYTK001
(A) Geneious file of plasmid map of pYTK001 vector containing Tetracycline Transactivator. **(B)** Plasmids containing parts necessary for tetracycline transactivator transcription unit (TU) **(C)** Tetracycline transactivator TU assembled into yeast integration vector pYTK097. This vector contains homology for the *leu2* locus on chromosome III. NotI digestion releases a linear DNA fragment with tetracycline transactivator and a functional *Leu2* expression unit with ends harboring homology to the native *leu2* locus.

Chapter 3: Humanized yeast model for testing human MVK variants using pGAL-Yeast/pTET-Human gene ON/OFF system

Chapter 3.1 Human MVK Variant Analysis Using Tetracycline-Regulated System

We generated the wildtype MVK and 33 of its variants into the yeast expression vector pAG416tetO-ccdB. The MVK variants were chosen from the Variant viewer database [29] such that they span the entire length of the MVK gene (**Figure 11**). The basic premise is to generate a reference set of variants for characterization in a humanized yeast setup. Majority of the variants have been found in clinical cases. Some have been analyzed in previous studies while others have never been characterized before. Wild-type MVK was obtained from the human ORFeome collection [48]. Variants were created using the Q5 Site-Directed Mutagenesis Kit from New England BioLabs (NEB). We used the wild-type MVK entry clone in PDONR221 as a background vector to introduce site-specific changes at various sites within the gene (by C. Gamache, Y. Ren and F. Zafar; **Figure 11**). Each individual variant was sequence verified before cloning into the destination vector pAG416tetO-ccdB using LR Gateway cloning.

Next, these plasmids were transformed into two yeast strains: pGAL-*ERG12* with or without tTA expression. The strain pGAL-*ERG12* allows the control of the yeast gene by galactose (**Yeast gene ON**) and dextrose (**Yeast gene OFF**) but lacks the tetracycline transactivator. Without the transactivator, the minimal pCYC promoter is incapable of strong expression, thus, the human MVK gene cannot be expressed. The strain pGAL-*ERG12* with tTA contains the transactivator and controls the yeast gene with galactose and dextrose and the presence of transactivator allows graded regulation of the human MVK gene using different concentrations of doxycycline [40]. An empty vector, pAG416tetO-ccdB, was used as a control. A total of 70 transformations were carried out using the EZ Yeast Transformation Kit (Zymogen). The transformations were done

on media containing galactose and 15µg/mL doxycycline (**Yeast gene ON and human gene OFF condition**). This condition ensured that all transformants would be able to grow due to the expression of the essential yeast gene *ERG12*. Additionally, the presence of doxycycline prevented any human gene expression to avoid any negative interactions with the yeast gene.

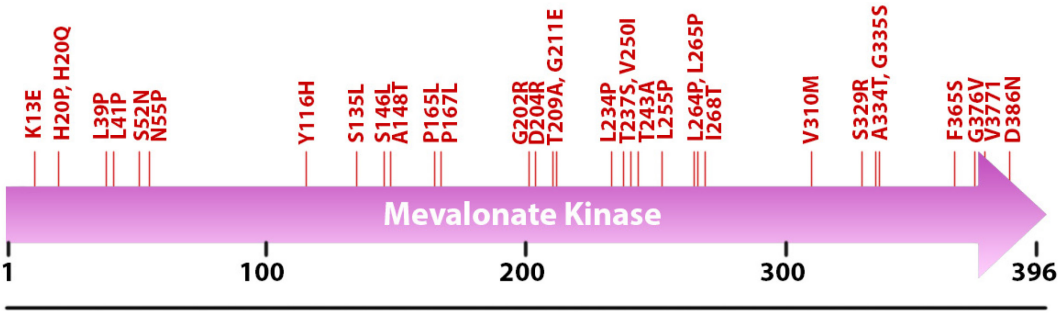

Figure 11. Mevalonate Kinase Variants Mapped on MVK Gene. MVK variants span the length of the Human Mevalonate Kinase gene. These variants were used for complementation assays in a tetracycline-regulated system.

The resulting colonies were grown overnight in dextrose and doxycycline and used for spotting assays. Spotting assays were done on media with 6 different conditions (**Table 1**).

Table 1: Growth Conditions for Spotting Assay

	Dextrose + 15µg/mL Doxycycline	Dextrose + 5µg/mL Doxycycline	Dextrose + 2µg/mL Doxycycline	Dextrose + NO Doxycycline	Galactose + 15µg/mL Doxycycline	Galactose + NO Doxycycline
Yeast Gene	OFF	OFF	OFF	OFF	ON	ON
Human Gene	OFF	ON – LOW Expression	ON – MEDIUM Expression	ON – HIGH Expression	OFF	ON – HIGH Expression

High concentrations of doxycycline such as 15µg/mL lead to complete down regulation of the human gene, while the condition without doxycycline allows high expression of the human gene. At 5µg/mL doxycycline, repression is still quite effective (**LOW expression condition**). A previous study demonstrated 1µg/mL doxycycline was enough to fully repress expression [40]. However, in our study, at 2µg/mL doxycycline, the human gene is expressed at MEDIUM levels.

An empty vector serves as a negative control when the yeast gene is turned off but no human gene is provided in its place, resulting in no yeast growth. Empty vector controls will only grow on media containing galactose, as the yeast gene is turned on in that condition and does not require a human gene to be present. Two important conditions to compare are galactose with (**Yeast gene ON, Human gene OFF**) and without doxycycline (**Yeast gene ON, Human gene ON**). In one case, the yeast gene is expressed but the human gene is not whereas the other allows for expression of both yeast and human gene. This is an important comparison to observe if there is an interaction between the human and yeast genes (considering that the two genes are orthologs with similar structure). In some cases, having both expressed could lead to a growth defect, if the human gene is dominant negative. Growth on dextrose corresponds to the yeast gene OFF, so in conditions where doxycycline concentration is high, we should see little to no growth due to the human gene also being repressed. On the other hand, growth on dextrose with no doxycycline present allows for high expression of the human gene while the yeast gene is off.

To show that expression of the human gene requires the presence of both the tTA and tetO, we did a pilot assay comparing the growth of strains with and without

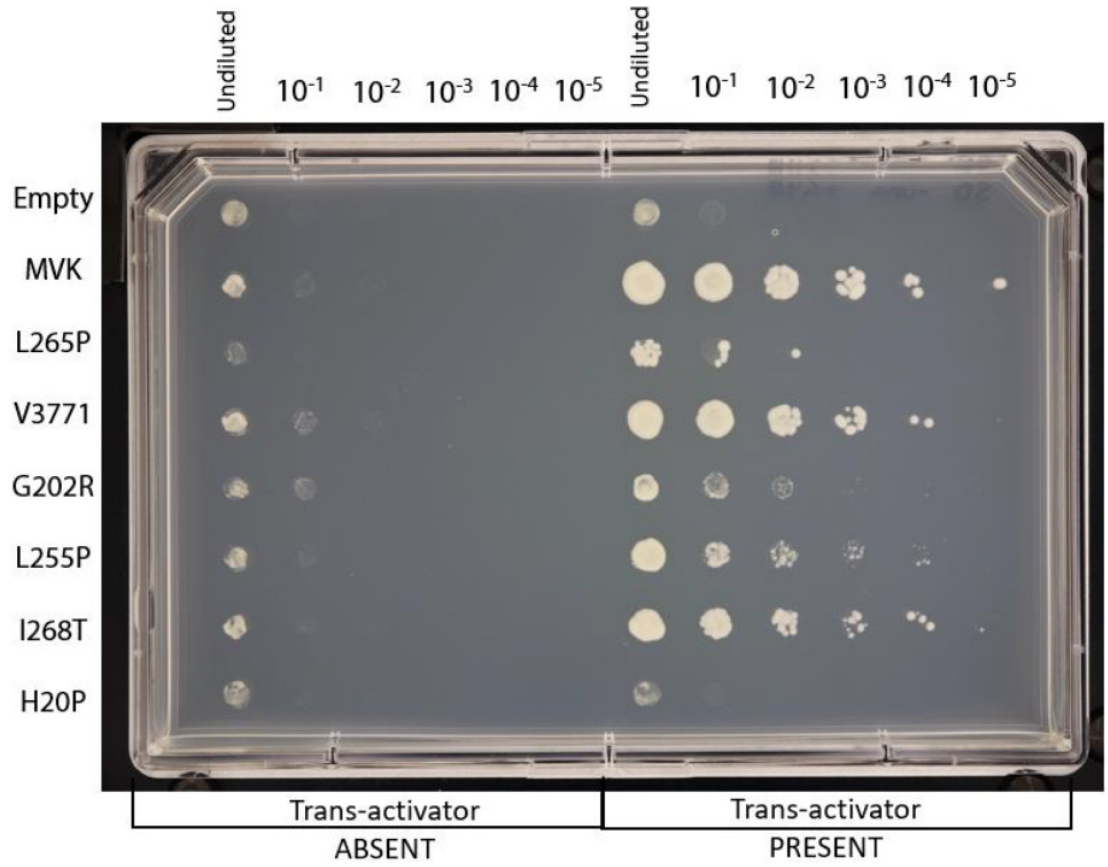

Figure 12. Complementation Assay Demonstrating the Necessity of Transactivator for Expression of the human gene *MVK*. Several variants of the *MVK* gene expression vectors were transformed into 2 yeast strains; one that contains the tetracycline transactivator (**right**) and one that does not (**left**). Spotting assay has been done on media containing dextrose, which represses the expression of the native yeast *ERG12* gene. Growth is only possible when human gene is expressed and functionally replaces the absence of yeast of *ERG12* gene by human *MVK* has been successful.

tTA in dextrose medium (**Yeast gene OFF**) in the absence of doxycycline (**Figure 12**). When the transactivator is absent, there is little to no growth in a condition when the yeast gene is turned off, regardless of which MVK variant has been transformed into the strain. This is the case for all doxycycline concentrations because without the transactivator, doxycycline does not affect gene expression. By comparison, when human gene expression clones are transformed into

the strain containing the tTA, growth varies depending on doxycycline concentration and severity of the mutation in a human gene. The empty vector is used as a negative control which shows no growth on dextrose with or without doxycycline (**Figure 12, top lane**). Wild-type MVK shows successful replaceability of the yeast gene with its human ortholog and serves as a reference to compare the functional replaceability of the variants (**Figure 12, second from top)**. Our growth assays show that MVK-L265P variant, classified as pathogenic in clinical studies, shows a severe growth defect even at HIGH expression conditions (**Figure 12**). While V377I is classified as pathogenic, however, in our assays it does not appear to show any major growth defect. This is consistent with our previous data [9] (as in **Figure 4**). Similarly, G202R is classified as pathogenic in the ClinVar database and shows a growth defect in our assay. L255P shows medium defect. HIGH expression condition shows growth but at reduced levels compared to the wild type.

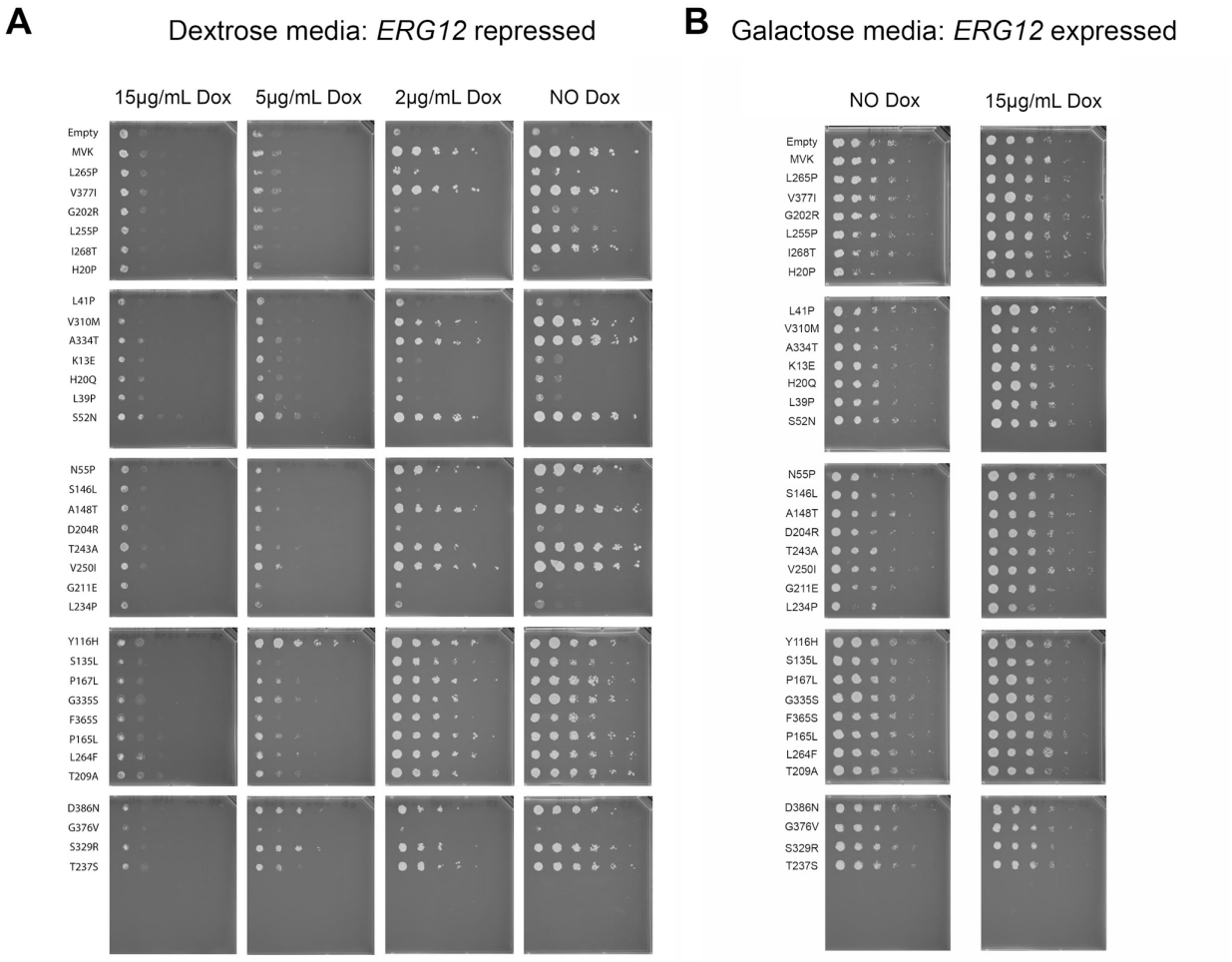

Figure 13. Complementation Assay Using Tetracycline-Regulated System. Wild type MVK and 33 of its variants transformed into the pGAL-*ERG12* + tTA strain. This strain allows expression of the essential gene *ERG12* on media containing galactose (right), while repressing the gene on media containing dextrose (left). Growth on dextrose can only occur after successful replacement of *ERG12* with *MVK*. Concentrations of doxycycline repress the human *MVK* gene to different levels.

After demonstrating the validity of our experimental setup, we next tested all 33 *MVK* variants for their ability to functionally replace the yeast gene under 4 different human gene expression conditions; No doxycycline (**Human gene ON-HIGH expression**), 2µg/mL doxycycline (**Human gene ON-MEDIUM expression**), 5µg/mL doxycycline (**Human gene ON-LOW expression)** and 15µg/mL doxycycline (**Human gene OFF**) (**Figure 13**). We show

the data for two conditions of yeast gene expression (**Yeast gene OFF - Left panel** & **Yeast gene ON - right panel**). Broadly, we see three different phenotypes. First, several human gene variants show no growth rescue even at HIGH expression conditions (such as H20P, D204R etc.) (**Figure 13A**). We refer to these human gene mutants as loss-of-function variants. Second class refers to human gene variants that show a growth rescue similar to the wild type *MVK* (such as A148T, S52N etc). These variants essentially behave similar to the wild type human gene in the conditions tested. Third class of human variants show expression dependent phenotype (such as L255P, I268T etc). These human gene variants show no or reduced rescue at low or medium expression conditions. Whereas at high expression conditions, these variants tend to replace the yeast gene function at a comparable rate to the wild type *MVK*. We refer to these human gene mutants as "**expression-sensitive**" variants. The variant database InFevers classifies L41P as likely pathogenic, it appears to have a severe growth defect in this assay. V310M showed little to no growth defect, it is clinically pathogenic. A334T is classified as likely pathogenic, it did not exhibit a growth defect. K13E is a binding site mutant, and as such we would expect a mutation in this residue to lead to a severe growth defect. This is in line with what we observe in the assay. K13E shows little to no growth on dextrose. H20Q showed little to no growth, much like the other mutant at this position H20P. L39P also exhibits a severe growth defect. S52N is considered a benign mutation, as expected it shows similar levels of growth to wild-type MVK. N55P is a binding site mutant, but it seems to be growing well. S146L is an active site mutant and shows a severe growth defect even in conditions that allow for high levels of expression. A148T is categorized as pathogenic, but does exhibit a growth defect. D204R is an active site mutant, and as expected, it shows a severe growth defect. We could not find any information on the mutant T243A, but it shows a similar growth phenotype as the wild-type. V250I shows no

growth defect but is classified as likely pathogenic. G211E shows a severe growth defect and is classified as pathogenic. L234P shows a severe growth defect and is considered pathogenic. F365S appears to have a slight growth defect and is classified as pathogenic. Interestingly, Y116H seems to show some growth even at high level of 5microg/mL doxycycline despite being classified as pathogenic. P167L, S135L, P165L, L264F and T204A seem to show comparable growth to wild-type MVK and are all classified as likely pathogenic. No information on G335S was available. D286N, which is interpreted as benign on ClinVar shows little to no growth defect. G376V shows a severe growth defect and is considered likely pathogenic. S329R and T237S show comparable growth to wild-type in the absence of doxycycline. But S329R shows a growth defect at 2µg/mL doxycycline. T237S is classified as pathogenic and S329R is likely pathogenic on ClinVar. These results have been summarized as a heatmap (**Figure 14**).

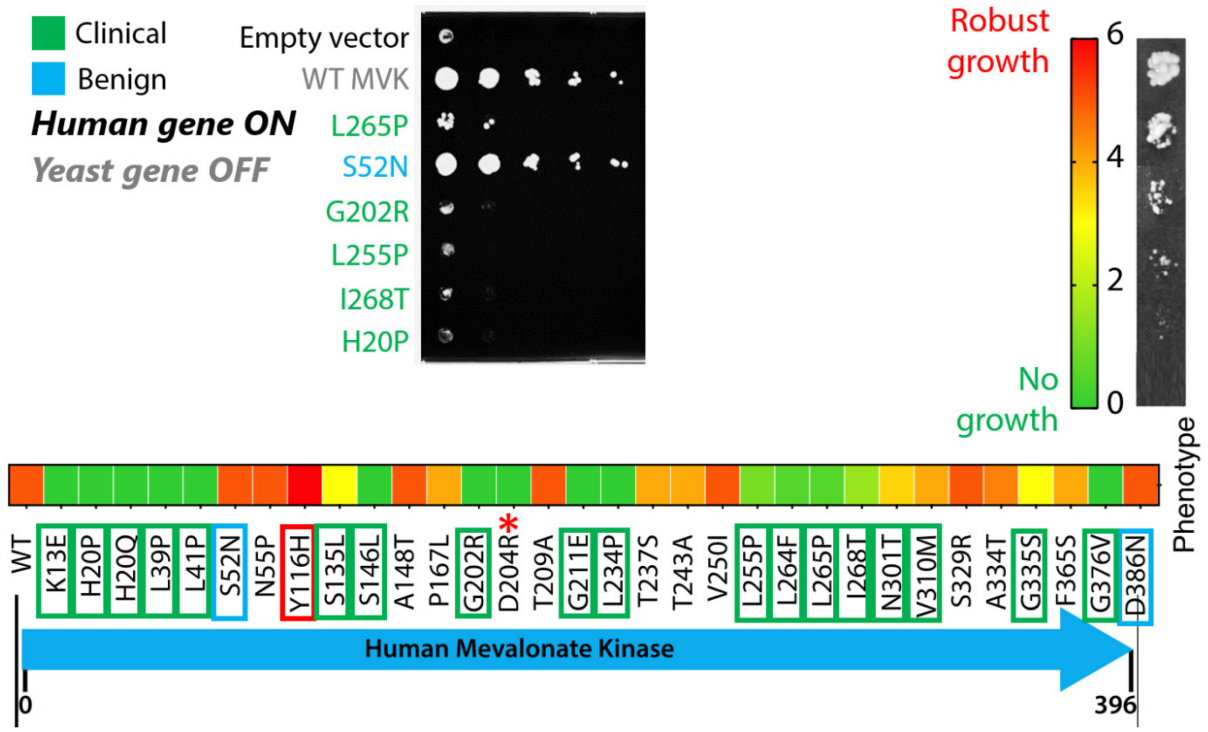

Figure 14. Heat Map Summarizing Growth Phenotype of Mutants
A heat map representing robust growth in red and no growth in green. Benign mutants marked with blue rectangle. Disease-associated mutants marked with green rectangle. Red rectangle marks a hyperactive mutant.

The protein structure for MVK can provide insight into the impact of these mutations. By mapping certain mutations onto the structure, we were able to hypothesize the reasons for mutant severity (**Figure 15**). The mutant K13E interacts with the active-site mutant D204A, these mutations likely have an impact on activity. Similarly, G202R is very close to the active site at position 204 and will likely impact the activity. In other cases, mutants such as L255P are located on the interface of two interacting subunits. Mutations at these sites will impact the interaction.

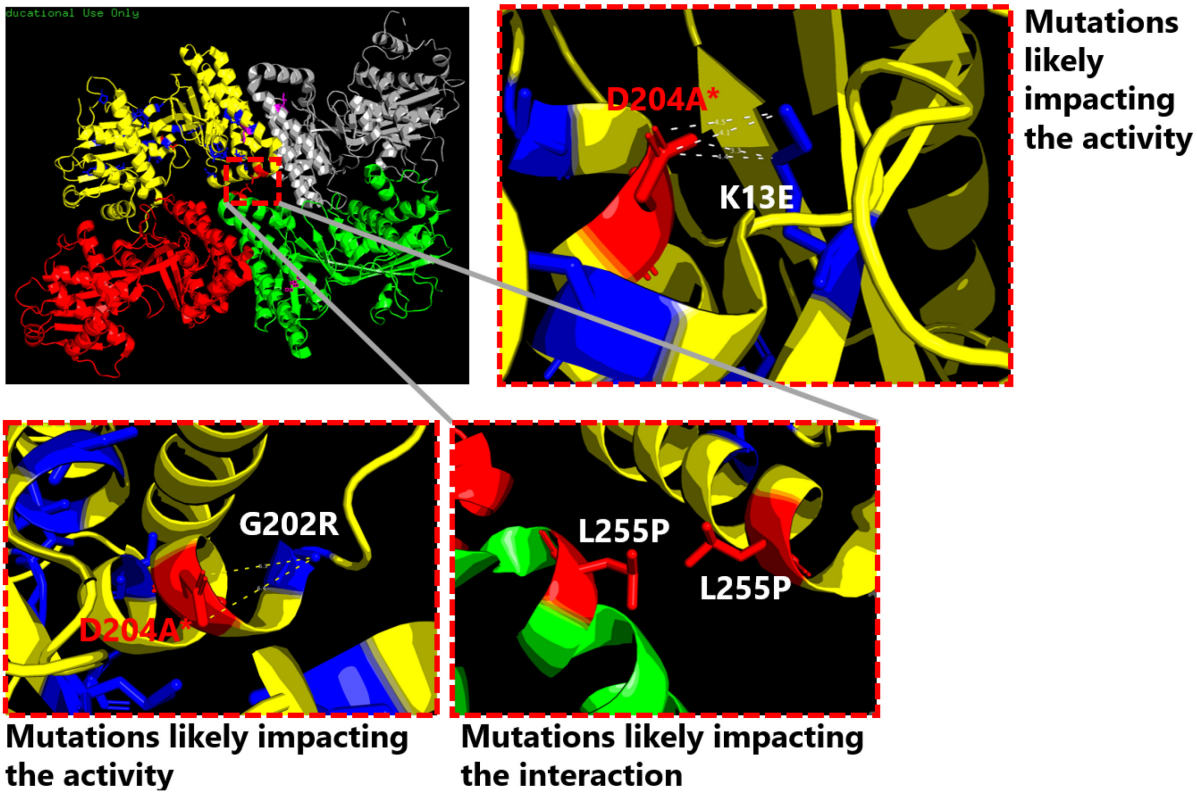

Figure 15. Disease-Causing Variants in the Context of Protein Structure
Protein structure of MVK visualized using Pymol software (PDB number 2R3V)[16]. Mutants K13E, D204R, G202R and L255P mapped on the protein structure to provide information on likely impact of mutation at those sites.

Chapter 4: Fluorescence as a Readout for Protein Activity and Stability

4.1 Generation of the wild type and mutant human MVK gene and Clones in yeast expression vectors to create MVK-eGFP Fusion Protein

All the humanized yeast assays performed thus far rely on growth as a readout of human gene function. Little is known about whether the severity of MVK mutants is due to an effect on protein activity or stability. To test this out, we decided to create a fusion protein of the wild-type MVK and some of its variants with enhanced Green Fluorescent Protein (eGFP). This would allow fluorescence as a measure of stability/activity. To understand this further, we also created a construct that included a self-cleaving peptide, E2A, between the MVK and eGFP [49]. This self-cleaving peptide would cut the protein so that the eGFP protein would be free of the MVK protein and thus, its stability would be independent of the human protein. The construct will also help to compare the MVK localizations across variants. In order to create this protein, we made use of a vector from the Susan Lindquist Yeast Gateway collection [47]. The plasmid pAG416GAL-eGFP-ccdB is compatible with LR gateway cloning, contains a GAL-inducible promoter and a ccdB cassette that can be replaced with any gene of interest to create a fusion protein (**Figure 16A**). An important step is to first remove the stop codon of the MVK gene so that the Open Reading Frame continues into the eGFP, creating a fusion protein. To remove the stop codon, we simply designed a reverse primer for the MVK gene without the Stop codon. These primers also contained attB sites that would allow for cloning into a BP cloning compatible vector pDONR221 generating an entry clone. We created two sets of primers, one that included E2A in the reverse primer and one that did not (**Figure 16B**). This method was used to create entry vectors by cloning a PCR product with attB sites into pDONR221 (**Figure 16C & D**).

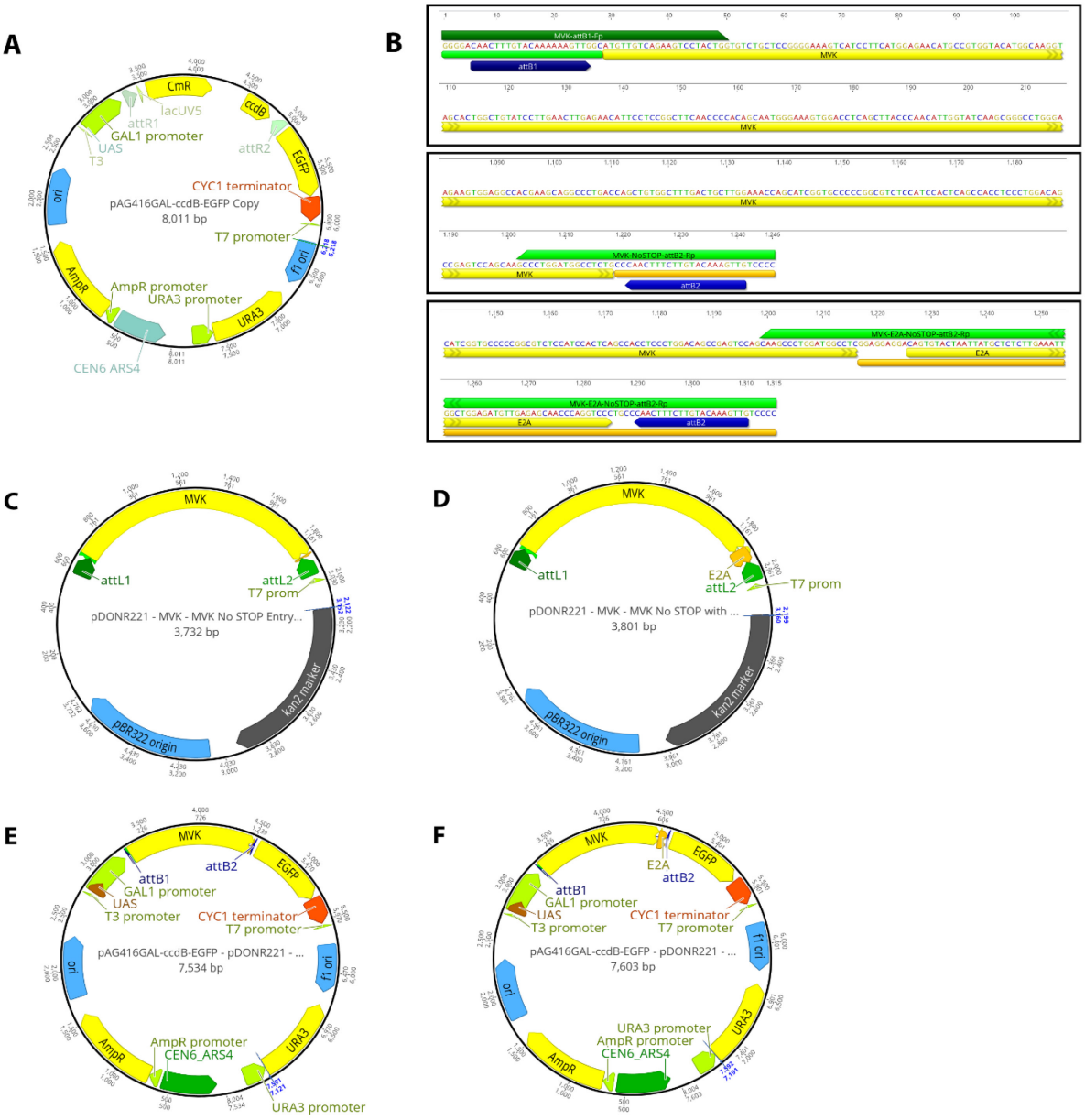

Figure 16. Human Mevalonate Kinase Gene and its Variants Fused to a Green Fluorescent Protein

(A) Plasmid pAG416GAL-ccdB-EGFP from the Susan Lindquist Yeast Gateway Collection is Gateway compatible and used to create a fusion protein tagged with enhanced green fluorescent protein (EGFP)[47]. **(B)** Primer design for addition of attB sites, removal of stop codon and addition of E2A self-cleaving peptide. **(C)** Geneious plasmid map of BP Gateway cloning of MVK into pDONR221. **(D)** Geneious plasmid map of BP Gateway cloning of MVK-E2A into pDONR221. **(E)** Geneious plasmid map of LR Gateway cloning of MVK into pAG416GAL-ccdB-EGFP to create a vector containing a fusion gene MVK-EGFP. **(F)** Geneious plasmid map of LR Gateway cloning of MVK into pAG416GAL-ccdB-EGFP to create a vector containing a fusion gene MVK-E2A-EGFP.

We chose five human MVK gene variants: wild type MVK and L265P, V377I, H20P and S52N variants and generated entry clones into pDONR221 (**Figure S4**). These were chosen based on previous data on their growth profiles when used to replace the yeast ortholog *ERG12* (**Figure 20**). As we have already established, S52N is a benign variant and shows growth levels that are similar to wild-type MVK. L265P and H20P mutants show severe growth defects in yeast. V377I seems to behave differently depending on the background of the conditions it's being tested in, but is thought to be a moderate mutant. These constructs were then tested using three forms of analysis: growth assay, microscopy, and flow cytometry. The MVK gene is controlled by a galactose promoter. Altering the level of galactose/dextrose can control expression.

4.2 Testing the Functional Replaceability of human MVK-eGFP in a yeast *ERG12* Knockout Strain

Using the yeast knockout collection, I made competent cells of a Magic Marker HetKO (heterozygous knockout) diploid strain that contained one wild-type copy of the *ERG12* gene and one knockout copy (replaced by KanMX cassette). The strains harbour SGA (Synthetic Genetic Array) markers to allow for an easy selection of haploid MATa type spores from the mixture of haploid and diploid yeast cells [32]. The MVK-eGFP and MVK-E2A-eGFP plasmids were transformed into this strain and plated on SD -Ura + G418 media. The colonies resulting from these transformations were sporulated. Sporulation efficiency was checked under the light microscope. Selection for a specific mating type can be done using a special media known as Magic Marker media. Magic Marker media containing G418 is used to select for MATa haploids harboring a knockout copy of the yeast gene, whereasile Magic Marker without G418 can be

used to select for MATa cells with the wildtype and knockout copy of the yeast gene. Furthermore, Magic Marker selection can be made with either dextrose or galactose as a carbon source allowing the control of the GAL promoter. Since our plasmid contains a galactose-inducible promoter which controls expression of the MVK gene, magic marker with the expression of the human gene in the presence of galactose allows for expression of the human MVK. Magic Marker made with dextrose will result in complete down-regulation of the human gene expression. We tested the functional replaceability in 4 conditions: magic marker (dextrose) + G418 (**Yeast gene ABSENT/(Human gene OFF**), magic marker with dextrose – G418 (**Yeast gene PRESENT/(Human gene OFF**), magic marker with galactose + G418 (**Yeast gene ABSENT/(Human gene ON**) and magic marker with galactose – G418 (**Yeast gene PRESENT/(Human gene ON**) (**Figure 17**). The functional replaceability data mimicked the previous assays such as the wildtype MVK and benign S52N variant resulted in robust growth (indicated as growth of several colonies) in the galactose + G418 condition similar to the galactose - G418 condition (**Figure 17**). Whereas, in dextrose + G418, the number of colonies observed is significantly lower than dextrose - G418 condition. This is expected since this condition corresponds to the human gene being repressed and the yeast gene is absent. However, the wild-type copy of the *ERG12* yeast gene allows for growth in dextrose - G418 condition. In these assays, since we are dealing with millions of cells post segregation, the few colonies we see are likely suppressors. For the mutants H20P and L265P, the condition galactose + G418 shows very few colonies, unlike the wild-type MVK. This condition corresponds to the knockout haploid being replaced by the human gene mutants, but since these two mutants contain severe disease-causing mutations, they cannot successfully replace the function of the yeast *ERG12* gene. These results are in line with previous growth assay data on these two mutants (**Figure 20**).

However, V377I shows robust growth in this condition suggesting no distinct phenotype. Another thing to note is that there seems to be no difference in the growth assays of constructs containing E2A and those without E2A (**Figure S2**). Additionally, the assays show that the eGFP did not affect the function of the human gene as the results of the assay are similar to non-tagged versions of these human proteins.

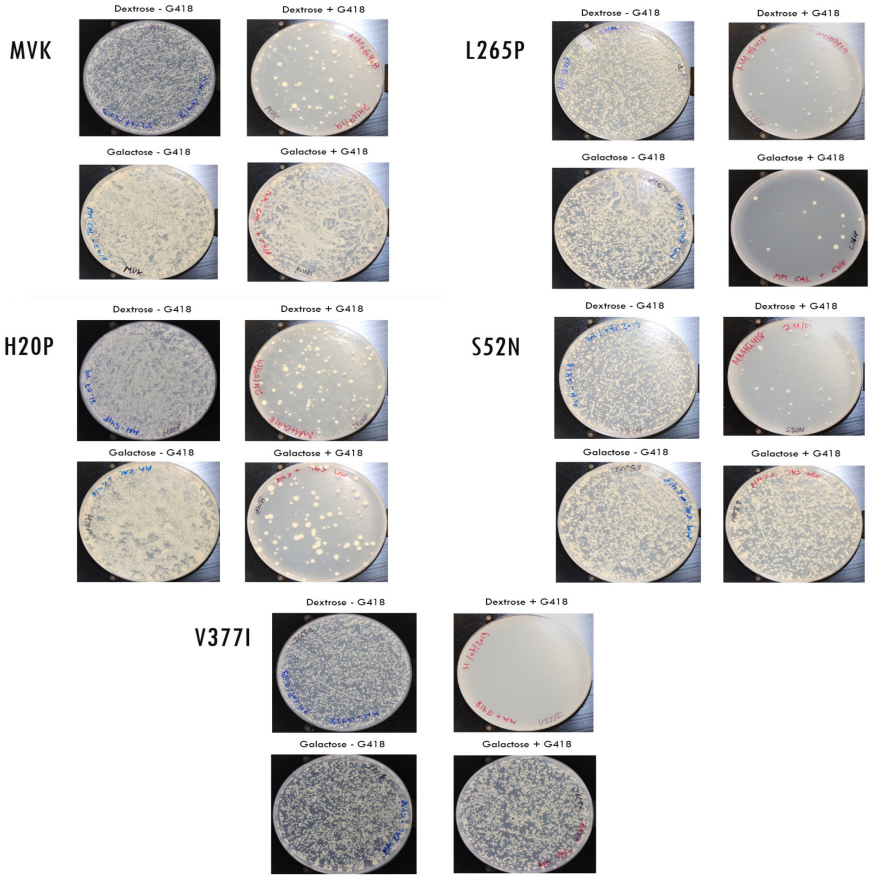

Figure 17. Growth Assay of Knockout *ERG12* Strain Transformed With MVK-EGFP Variants

Haploid strain with either a knocked out copy of *ERG12* (plates with G418) or wild type copy (plates without G418)[32]. Strains were previously diploid and transformed with plasmid containing MVK-EGFP under a galactose-inducible promoter. These plates represent haploids, following sporulation. Galactose media allows for expression of the human MVK-EGFP fusion gene and its variants. Dextrose media does not allow expression of MVK-EGFP. The condition galactose + G418 represents the knocked out yeast gene *ERG12* and expression of the human gene MVK-EGFP. In this condition, growth is rescued through successful complementation of the yeast gene by the human gene (depending on the function of the human gene).

Chapter 4.3 Microscopy and Flow Cytometry Analysis of MVK-eGFP Fusion Variants

While there is no obvious change in fluorescence between the wild type MVK and S52N variant, we saw obvious localization differences. The variants, L265P and H20P seem to show a localization of fluorescence in certain areas of the cell like foci compared to diffused cytosolic localization in the wildtype and S52N variant (**Figure 18**). For L265P in particular, these results line up with the flow cytometry data.

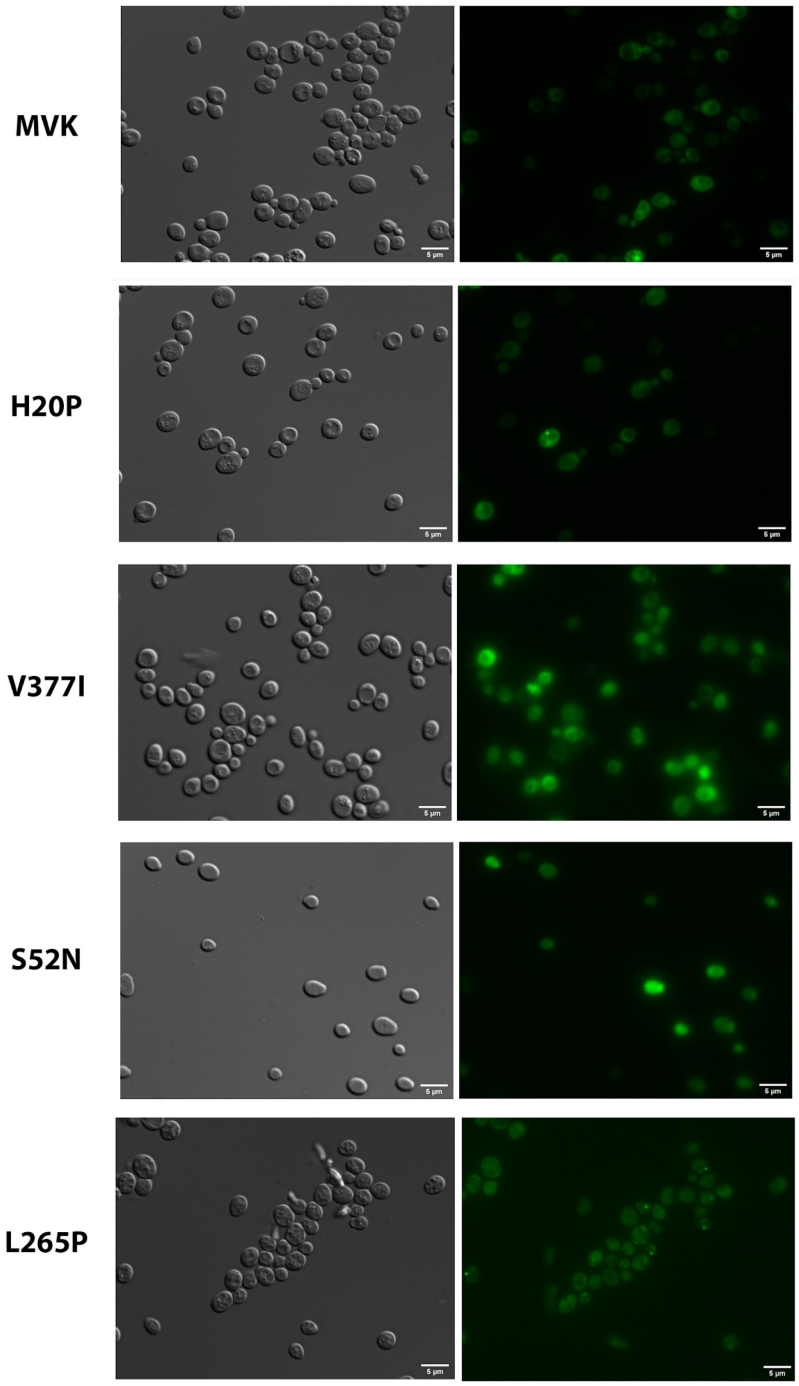

Figure 18. Microscopy of eGFP-Tagged MVK Variants

All images were taken with objective magnification 100x and ocular (camera) magnification 10x for a total magnification of 1000x. Images were taken in differential interference contrast (DITC) (left panel) and fluorescein (FITC) (right panel). A diploid strain containing one wild type copy of *ERG12* and one knockout copy was transformed with the vector pAG416-GAL-MVK-EGFP containing MVK or several of its variants. Cells from the resulting colonies were grown in 2% galactose and used for microscopy.

Another method for reading the fluorescence of the eGFP-tagged MVK variants is through flow cytometry. For the flow cytometry analysis, diploid strains transformed with pAG416GAL-MVK-eGFP had to be used. The haploid strains would pose a problem because severe mutants such as L265P and H20P would not be able to grow in the condition where there was no wild-type *ERG12* gene present. As a control, the wild-type MVK gene cloned into the vector pAG415GAL-ccdB was used. This control lacked an eGFP tag and would not link fluorescence levels to MVK expression. Instead, the fluorescence displayed by the control would represent background GFP fluorescence present in our cells. It could therefore be used as a point of comparison for the other samples.

Colonies containing each of the 5 MVK variants were inoculated into media containing different concentrations of galactose. The three concentrations of galactose used were 2%, 0.01%, and 0.005% (**Figure 19**). For the conditions with 0.01% and 0.005% galactose, 2% raffinose was added as a carbon-source. There was a clear distinction between the fluorescence level of MVK-eGFP and MVK without a tag when cells were grown in 2% galactose, with MVK-eGFP showing higher levels of fluorescence (**Figure 19**). In the high galactose condition, most of the eGFP-tagged MVK variants displayed a similar fluorescence pattern to wild-type MVK (except in the case of L265P that shows lower fluorescence) irrespective of their ability to rescue the yeast growth defect (**Figure 19, left panel**). However, at 0.01% galactose condition (medium level expression), the mutant L265P exhibited a significant drop in fluorescence compared to the other samples suggesting an unstable protein. Comparatively, a similarly defective H20P variant shows fluorescence levels similar to the wild type MVK albeit with foci formation as observed in the microscopy images (**Figure 19, middle panel**). At 0.01% galactose (lowest expression condition), this difference in fluorescence is even more apparent with L265P

exhibiting a similar fluorescence pattern as MVK without an eGFP tag (**Figure 19, medium panel**). At 0.005% galactose, expression of our tagged genes is so low that there is no distinction between the fluorescence of samples tagged with eGFP and the untagged control (**Figure 19, right panel**).

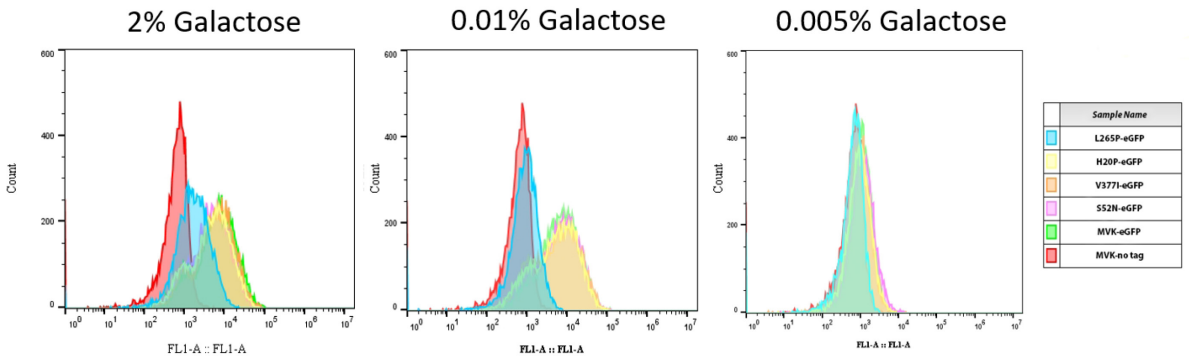

Figure 19. Flow Cytometry of eGFP-tagged MVK Variants
A diploid strain containing one wild type copy of *ERG12* and one knockout copy was transformed with the vector pAG416-GAL-MVK-EGFP containing MVK or several of its variants. Cells from the resulting colonies were grown under 3 conditions: 2% galactose (high expression of MVK), 0.01% galactose with 2% raffinose (intermediate expression of MVK) and 0.005% galactose with 2% raffinose (low expression of MVK). Samples were analyzed using the BD Accuri Flow Cytometer from Concordia's Genome Foundry. Graphs were made using FlowJo software.

Chapter 5: Materials & Methods

LR Gateway Cloning

Plasmids compatible with this form of cloning can be found in the Susan Lindquist Yeast Gateway Kit[47]. LR Cloning is characterized by “L” and “R” recombination sites that can be used to swap the ccdB cassette with our gene of interest in a plasmid. The ccdB cassette consists of chloramphenicol resistance marker and a ccdB gene which is toxic to *E. coli*. An enzyme called LR clonase II is used to catalyze the reaction. Proteinase K is an enzyme that inactivates LR clonase II, and should be used before transforming the DNA into *E. coli*. These plasmids are shuttle vectors that can function in both *E. coli* and yeast.

BP cloning follows the same logic, although the recombination sites are shorter and referred to as “B” and “P” sites. BP cloning is often used to generate entry vectors that will subsequently be used for LR cloning into a destination vector.

Yeast Gateway Kit vectors used in this study:

pAG416GAL-eGFP-ccdB

pAG416GPD-ccdB

Golden Gate Cloning

A cloning method that applies the basic concepts of engineering to biology. Golden Gate cloning relies on “parts” that can be assembled together by corresponding overhangs. For example, a promoter is considered one part while a terminator would be another. Rounds of cloning are done with BsaI and BsmbI. When a golden gate compatible plasmid is cut with BsmbI and assembled into a new vector, that cut site is transformed to a BsaI site that can be used for subsequent cloning. Selection for this type of cloning is often on ampicillin or chloramphenicol, with a superfolderGFP that can be swapped out for a gene of interest. Therefore, obtaining white

colonies from a cloning is often an indication of success. Whereas green colonies indicate a failure to replace the sfGFP with the gene of interest. Golden Gate vectors were obtained from MoClo-YTK Collection from the Deuber lab[45].

Golden gate plasmids used in this study:

pYTK001, pYTK030, pYTK077, pYTK095, pYTK097, pYTK009, pYTK053,

Restriction enzyme cloning

A destination vector and insert are cut with the same restriction enzyme, creating complementary overhangs. The overhangs are annealed together using T4 DNA ligase.

Yeast Transformations

Yeast transformations were done using the Frozen-EZ Yeast Transformation II kit from Zymo Research. EZI solution is used to create competent cells. EZ II solution is used for storage of competent cells in -80°C freezer. EZ III solution is used to transform yeast by adding 500 µL of EZ III to 50 µL of competent cells along with the DNA to be transformed. Tubes are vortexed and incubated at 30°C for 45 minutes. The resulting cell culture is centrifuged, resuspended in 50 µL sterile water and plated on selective media.

E. coli Transformations

DNA is added to 50 µL of *E. coli* competent cells and chilled on ice for 30 minutes. Cells are heat shocked at 42°C for 30 seconds, then chilled on ice for 5 minutes. 950 µL LB media is added to the cell culture and incubated for 1 hour at 37°C. The resulting cell culture is centrifuged, resuspended in 50 µL sterile water and plated on selective media.

Primer & Construct Design

All primers were designed using Geneious Prime software. Geneious was also used to edit DNA, create plasmid maps, digest plasmids, and simulate cloning.

Site-Directed Mutagenesis

Done using the Q3 Site-Directed Mutagenesis Kit by New England BioLabs (NEB).

Plasmid Mini-Prep

Using the QIAprep Spin Miniprep Kit by Qiagen. Overnight culture is centrifuged. P1 buffer is used to resuspend cells. The lysis buffer P2 is added and tubes are inverted to mix. The neutralization buffer N3 is added and the sample is centrifuged. Supernatant is added to spin columns with silica membrane for binding DNA. PE buffer containing ethanol is used as a wash buffer. Elution buffer is used to elute the DNA into Eppendorf tubes.

Sporulation

The diploid knockout strain *MATa/a ura3D0 leu2D0 his3D1 lys2D0/LYS+ met15D0/MET15+ can1delta::LEU2+-MFA1pr-HIS3/CAN1+ erg12::kanMX/ERG12+* from the Synthetic Genetic Array (SGA) Collection from the Boone lab was used for sporulation[32]. This strain has one copy of the *ERG12* gene replaced with a KanMX cassette and one wild-type copy of *ERG12*. By inoculating this strain into minimal media, the yeast undergo sporulation under starvation conditions, resulting in tetrads containing two MATa and two MATα haploids. Haploids can be selected for on Magic Marker media. Magic Marker media without G418 will select for MATa haploids with the wild-type copy of the gene. Magic Marker with G418 will select for MATα haploids consisting of the deleted copy of the gene.

Media

YPD: yeast extract, peptone, dextrose

YPG: yeast extract, peptone, galactose

YPR: yeast extract, peptone, raffinose

SD -ura: yeast nitrogen base without amino acids or ammonium sulfate, glutamate, yeast supplements without uracil (ura dropout), agar, 40% dextrose (final dextrose concentration 2%)

Galactose -ura: yeast nitrogen base without amino acids or ammonium sulfate, glutamate, yeast

supplements without uracil (ura dropout), agar, 40% galactose (final galactose concentration 2%)

Magic Marker media: 0.17% yeast nitrogen base without amino acids and ammonium sulphate, 0.1% glutamate, 0.2% dropout media (Sc-ura-his-leu-arg), 2% agar, 2% dextrose, distilled water

Sporulation media: potassium acetate 10%, zinc acetate 0.5%, autoclaved MilliQ

G418/Geneticin (200µg/ml) from Gico

Doxycycline (50µg/ml)

Graphs

Growth curve graphs in figure 9 were generated using Microsoft Excel.

Flow Cytometry

The BD Accuri C6 flow cytometer from Concordia's Genome Foundry was used to analyse EGFP-tagged samples

Microscopy

Microscopy images taken using the Leica DM 6000 microscope from Concordia's Center for Microscopy and Cellular Imaging. All images were taken at objective magnification 100x with camera magnification of 10x. Total magnification 1000x. Images were taken in Differential Interference Contrast (DIC) and Fluorescein (FITC).

Chapter 6: Conclusion

In the next few decades, one of the major challenges in human biology would be attributing heritable changes in the genome to the phenotype. The difficulty is partly due to the lack of sample size in humans. Other factors contributing to the complexity of genome-phenotype association are pleiotropy, complex genetic interactions and context-specific variations of phenotypes or what geneticists call “epistasis”. Solving this problem may lead to rationally & directly identify disease genes, perhaps even for highly polygenic ones. Recent advances in understanding the conservation of genes and their functions across different model organisms, including yeast, have revealed that despite millions of years of separate evolution, not only are genes conserved, their interaction partners are also preserved among different species[8,9,35]. Traits and diseases often reflect this shared heritage. The remarkable extent to which genes are functionally equivalent between humans and yeast demonstrates the power of distant model organisms for studying human processes[50,51]. Systematic studies have discovered extensive genetic polymorphism in these genes, including many rare variants that cause or predispose to diseases. Exploring how this variation contributes to cellular function and overall human health remains a challenge and has not followed the rate at which variants are identified. Although humans and yeast differ dramatically with respect to cell and tissue organization (i.e., unicellular vs. multicellular body plans), diet, metabolism, motility, and environment. Nonetheless, several studies have discovered that many conserved protein-coding DNAs of human genes could substitute for their yeast equivalents and sustain yeast growth. The assays, thus, uncovered which conserved processes are still interchangeable, creating direct tests of human gene function for the swappable subset in a tractable system. The humanized yeast could now act as test tubes to study human genetic variation and discover drugs or suppressor

mutations that alleviate the defects in these genes, providing a comprehensive understanding of the effect of human genetic variation on function (Figure 1). Yeast genes that are both essential and humanizable, there is a significant overlap between Mendelian diseases and rare diseases. One such gene is human MVK which functionally replaces the orthologous yeast gene, ERG12. I used the functional replaceability in yeast as a readout of human genetic variation in the MVK gene. However, it is not clear what genetic background and expression levels are needed to optimally test the effect of variation. Therefore, I established a pipeline for controlling the human and the yeast gene using regulatable promoters (pTET for human and pGAL for yeast) (Chapter 2). The assays performed in different expression control in the presence or absence of the yeast gene are illustrated in this thesis. As explained in Chapters 3 & 4, the humanized yeast could identify normal variants from disease-causing variants (mild or severe) with a relatively easy output of growth as a phenotype.

However, while most humanization assays matched the clinical phenotypes, there were a few that did not show any observable phenotypes in the conditions tested. One reason for such discrepancy could be the different genetic interactions in human vs yeast and expression conditions. Previously, in our laboratory, a former Biology 490 student from our lab Yuxiang Ren performed complementation assays that repressed the expression of ERG12 by doxycycline (Figure 24). This was done using a S.cerevisiae strain from the Tetracycline-Promoter Hughes Collection from Open Biosystems (GE Dharmacon)[52]. This strain was then transformed with MVK variants inside a pAG416-GPD-ccdB vector, allowing the MVK variants to be constitutively expressed[47]. By comparing the results of this experiment to mine, we can see that most of the results align well, but there are a few exceptions. The mutant with the most

evident difference in growth is L255P which appeared to have a severe growth defect in my experiment but grew comparably to wild-type in Ren's experiment. The mutant V377I, which is the most common mutant found in clinical studies, grew to levels comparable to wild-type in pTET assays but had a slight growth defect in the case of experiments performed earlier (Y.Ren). I would classify the mutants G335S and F365S as having a slight growth defect, whereas previous assays do not show any growth defect for these two mutants. Aside from L255P, some of these differences can be attributed to different yeast genetic backgrounds and distinct human gene expression levels. However, the choice of the promoter to control the human gene expression does seem to play a role in whether a mutant will complement. This can be more clearly observed in previous assays when we switched from using a constitutive GPD promoter to an inducible GAL promoter.

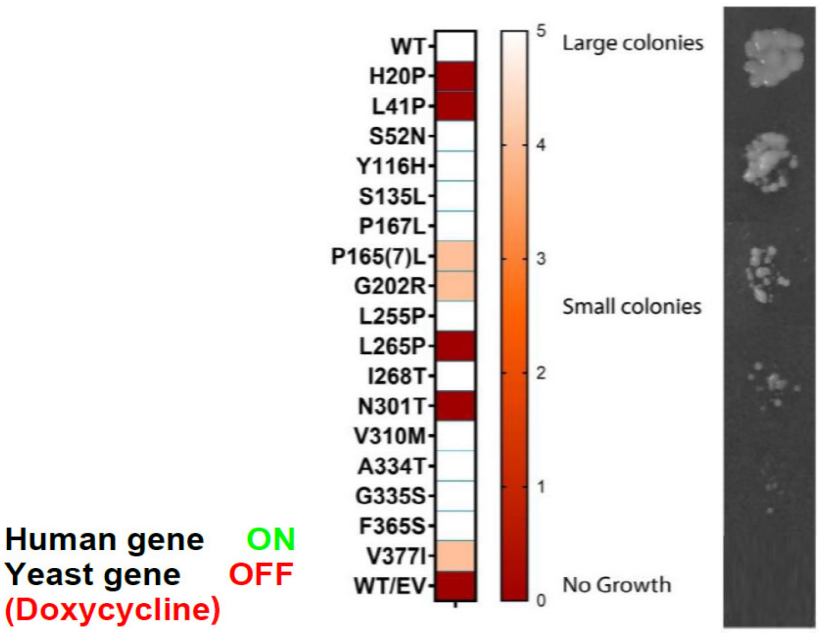

Figure 20. Complementation Assay Using Galactose-Regulated System
MVK Variants in vector pAG416-GPD-ccdB which allows control of expression of human MVK gene by the GPD constitutive promoter transformed in a yeast strain with pTET-*ERG12*. Vector containing variants have been cloned into a strain which has the yeast *ERG12* gene controlled by a tetracycline-repressible promoter[52]. This work was completed by BIOL 490 student YuXiang Ren.

Another likely rationale for lack of the yeast growth phenotype for some of the clinical cases could be the human MVK alleles present in humans as diploids. Most clinical cases of mevalonate kinase deficiency are caused by compound heterozygotes with two different alleles of MVK present as paternal and maternal copies.This study only looked at the effect of human genetic variation as a single allele. To study the effect of compound heterozygotes, previously our laboratory tested the two mutant alleles mimicking the compound heterozygote genotype in humans (**Figure 21**). These mutants were cloned into a vector pAG415-GAL-ccdB which allowed for galactose-inducible control[47]. These GAL mutants were transformed along with GDP mutants into the same tetracycline-repressible *ERG12* strain used in the above assay[52]. By varying the concentration of galactose in the media, she was able to control expression levels of one set of MVK alleles, while the other set was constitutively expressed. From these results, we can see some difference from the single mutant experiment. For example, the mutant V310M did not show a growth defect in either of our single mutant assays, but when combined with other mutants, we observed severe growth defects. For the mutant V377I, we see a range of growth phenotypes when combined with other mutants. V377I-L265P confers a severe growth defect, whereas L265P-F365S shows no growth defect.

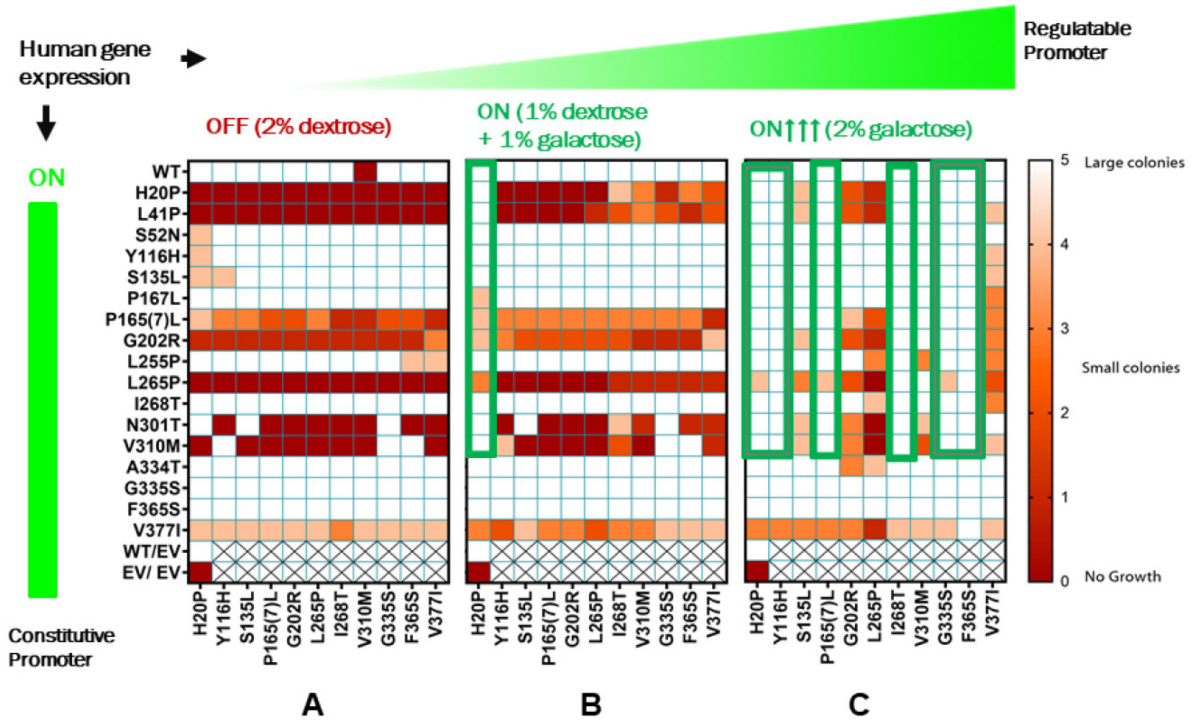

Figure 21. Complementation Assay of Double-Mutants Using Galactose-Regulated and Constitutively Expressed System

MVK variants cloned into pAG416-GPD-ccdB (y-axis) and pAG415-GAL-ccdB (x-axis) and transformed into tetracycline-repressible strain[47,52]. **(A)** 2% dextrose condition does not allow expression of human MVK variants under GAL promoter. **(B)** 1% dextrose, 2% galactose allows some expression of human MVK under GAL promoter. **(C)** 2% galactose overexpression condition where MVK variants under GAL promoter are highly expressed. Variants controlled by the GPD promoter are constitutively expressed under all three conditions. This assay allows us to observe how phenotype varies when two mutants interact.

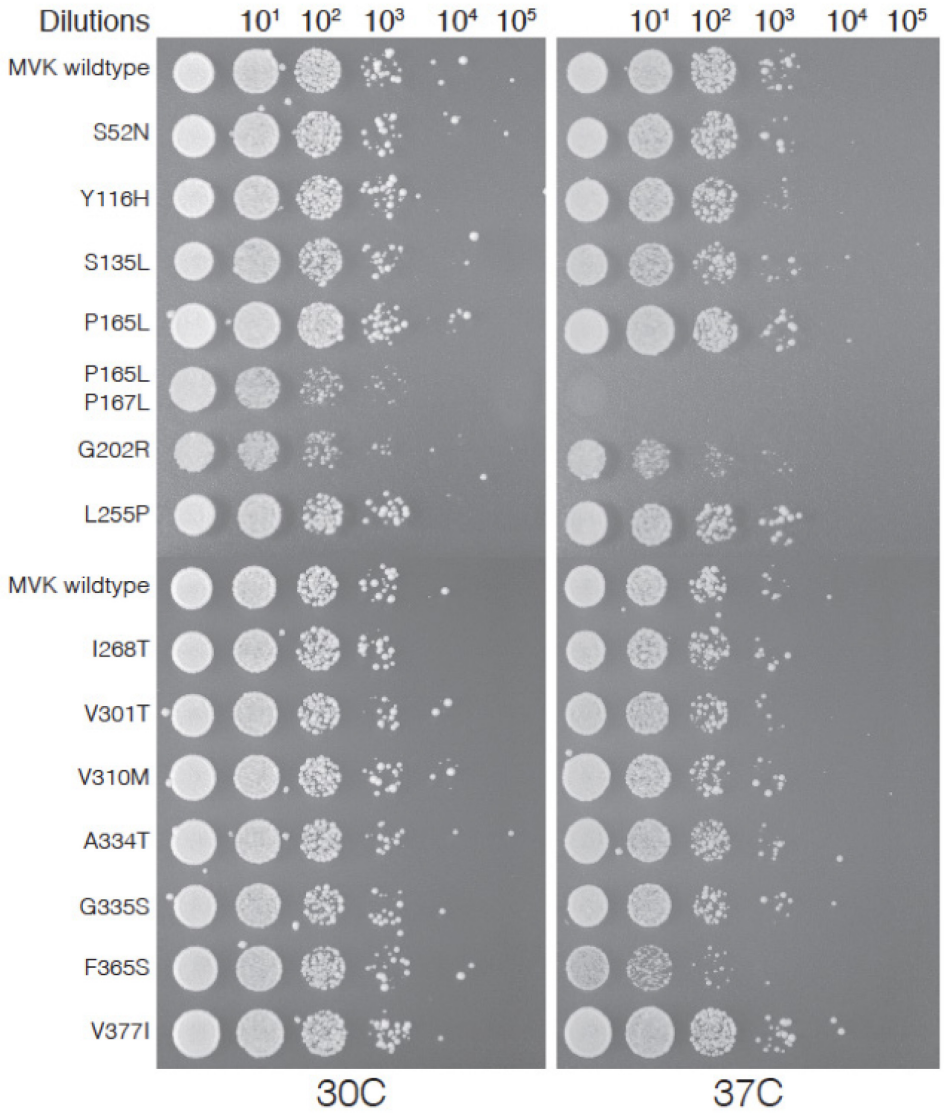

Figure 22. Complementation Assay Grown in Temperature-Sensitive Conditions

Cells were grown in permissive condition at 30°C and restrictive condition at 37°C. The double mutant P165L/P167L appears to be temperature sensitive at 37°C. G202R and F365S also have a more severe growth defect at the higher temperature. This work was completed by Dr. Aashiq Kachroo.

Another important set of experiments performed previously by Dr Kachroo were to look at the effect of temperature on our mutants (**Figure 22**). By growing the yeast cells in permissive conditions (30°C as an optimal growth condition) and comparing it to growth in restrictive

conditions (37°C), we can identify temperature-sensitive mutants. The double mutant P165L/P167L exhibits a severe growth defect when grown at 37°C. G202R and F365S also appear to show a more significant growth defect at the higher temperature. These results are exciting because one of the defining symptoms of mevalonate kinase deficiency is periods of high fever[23].

Overall, our combined data for humanized yeast MVK variant screening has established a reference set of conditions to distinguish disease-causing human gene variants from common polymorphism using simple yeast growth assays as a readout. In this project, I have successfully built a humanized yeast system that allows for the control of gene expression of both yeast and human orthologs. The control of the yeast gene acts as an ON/OFF switch, while control of the human gene is somewhat linear and depends on doxycycline concentration. The yeast ortholog is controlled by a galactose-induced promoter. There are several benefits to working with a system like this in yeast, such as their ability to be genetically modified, short generation time and easy storage.

Our approach now serves as a benchmark to study other human genes in humanized yeast setup. While mutations cause MKD in the MVK gene, several distinct diseases are caused by mutations in other genes of the cholesterol biosynthesis pathway. It would be possible to modify this system to study those other genes or to perhaps study a completely different pathway as well.

By comparing the results of this complementation assay to clinical data, we show that most severe and expression-dependent mutants can be identified using this approach. The results also matched up with most of the data from previous assays under different modes of expression

control. By analyzing MVK and 33 of its variants, we have created a reference library to test the efficiency of this system to be used in the follow-up experiment of deep mutational scanning of MVK gene that account for all possible variants of MVK.

Chapter 7: Bibliography

1. Hess JF, Kohl TA, Kotrová M, Rönsch K, Paprotka T, Mohr V, et al. Library preparation for next generation sequencing: A review of automation strategies. Biotechnol Adv. 2020;41: 107537.

2. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes JC, et al. Nucleotide sequence of bacteriophage φX174 DNA. Nature. 1977;265: 687–695.

3. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409: 860–921.

4. Garrido-Cardenas JA, Garcia-Maroto F, Alvarez-Bermejo JA, Manzano-Agugliaro F. DNA Sequencing Sensors: An Overview. Sensors . 2017;17. doi:10.3390/s17030588

5. Domchek SM, Bradbury A, Garber JE, Offit K, Robson ME. Multiplex genetic testing for cancer susceptibility: out on the high wire without a net? J Clin Oncol. 2013;31: 1267–1270.

6. Weile J, Roth FP. Multiplexed assays of variant effects contribute to a growing genotype-phenotype atlas. Hum Genet. 2018;137: 665–678.

7. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17: 405–424.

8. Sun S, Yang F, Tan G, Costanzo M, Oughtred R, Hirschman J, et al. An extended set of yeast-based functional assays accurately identifies human disease mutations. Genome Research. 2016. pp. 670–680. doi:10.1101/gr.192526.115

9. Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science. 2015;348: 921–925.

10. Kataoka T, Powers S, Cameron S, Fasano O, Goldfarb M, Broach J, et al. Functional homology of mammalian and yeast RAS genes. Cell. 1985;40: 19–26.

11. Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H. The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci U S A. 2004;101: 15386–15391.

12. Altenhoff AM, Glover NM, Dessimoz C. Inferring Orthology and Paralogy. Methods Mol Biol. 2019;1910: 149–175.

13. Luo J, Yang H, Song B-L. Mechanisms and regulation of cholesterol homeostasis. Nat Rev

Mol Cell Biol. 2020;21: 225–245.

14. Herman GE. Disorders of cholesterol biosynthesis: prototypic metabolic malformation syndromes. Hum Mol Genet. 2003;12 Spec No 1: R75–88.

15. Mevalonate kinase. In: Uniprot [Internet]. [cited 14 Feb 2021]. Available: https://www.uniprot.org/uniprot/Q03426

16. Fu Z, Voynova NE, Herdendorf TJ, Miziorko HM, Kim J-JP. Biochemical and structural basis for feedback inhibition of mevalonate kinase and isoprenoid metabolism. Biochemistry. 2008;47: 3715–3724.

17. van der Burgh R, Ter Haar NM, Boes ML, Frenkel J. Mevalonate kinase deficiency, a metabolic autoinflammatory disease. Clin Immunol. 2013;147: 197–206.

18. Favier LA, Schulert GS. Mevalonate kinase deficiency: current perspectives. Appl Clin Genet. 2016;9: 101–110.

19. van der Burgh R, Pervolaraki K, Turkenburg M, Waterham HR, Frenkel J, Boes M. Unprenylated RhoA contributes to IL-1β hypersecretion in mevalonate kinase deficiency model through stimulation of Rac1 activity. J Biol Chem. 2014;289: 27757–27765.

20. Zhang S. Natural history of mevalonate kinase deficiency: a literature review. Pediatr Rheumatol Online J. 2016;14: 30.

21. Drenth JP, Haagsma CJ, van der Meer JW. Hyperimmunoglobulinemia D and periodic fever syndrome. The clinical spectrum in a series of 50 patients. International Hyper-IgD Study Group. Medicine . 1994;73: 133–144.

22. van der Hilst JCH, Bodar EJ, Barron KS, Frenkel J, Drenth JPH, van der Meer JWM, et al. Long-term follow-up, clinical features, and quality of life in a series of 103 patients with hyperimmunoglobulinemia D syndrome. Medicine . 2008;87: 301–310.

23. Bader-Meunier B, Florkin B, Sibilia J, Acquaviva C, Hachulla E, Grateau G, et al. Mevalonate kinase deficiency: a survey of 50 patients. Pediatrics. 2011;128: e152–9.

24. Ter Haar N, Lachmann H, Woo P, Simon A, Meini A, Dolezalova P, et al. PReS-FINAL-2335: Preliminary analysis of 85 patients with mevalonate kinase deficiency from the eurofever registry. Pediatr Rheumatol Online J. 2013;11: P325.

25. Haas D, Hoffmann GF. Mevalonate kinase deficiencies: from mevalonic aciduria to hyperimmunoglobulinemia D syndrome. Orphanet J Rare Dis. 2006;1: 13.

26. Vuch J, Marcuzzi A, Bianco AM, Tommasini A, Zanin V, Crovella S. Evolutionary hypothesis of the Mevalonate Kinase Deficiency. Med Hypotheses. 2013;80: 67–69.

27. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat Rev Cancer. 2007;7: 937–948.

28. Govindaraj GM, Jain A, Peethambaran G, Bhoyar RC, Vellarikkal SK, Ganapati A, et al. Spectrum of clinical features and genetic variants in mevalonate kinase (MVK) gene of South Indian families suffering from Hyperimmunoglobulin D Syndrome. PLoS One. 2020;15: e0237999.

29. Variant Viewer. In: National Center for Biotechnology Information (NCBI) - Variant Viewer Database - MVK [Internet]. [cited 20 Mar 2021]. Available: https://www.ncbi.nlm.nih.gov/variation/view/?q=MVK

30. Botstein D, Fink GR. Yeast: an experimental organism for 21st Century biology. Genetics. 2011;189: 695–704.

31. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, et al. Life with 6000 genes. Science. 1996;274: 546, 563–7.

32. Tong AHY, Boone C. Synthetic Genetic Array Analysis in Saccharomyces cerevisiae. Yeast Protocols. 2005. pp. 171–192. doi:10.1385/1-59259-958-3:171

33. Duina AA, Miller ME, Keeney JB. Budding yeast for budding geneticists: a primer on the Saccharomyces cerevisiae model system. Genetics. 2014;197: 33–48.

34. Cuperus JT, Lo RS, Shumaker L, Proctor J, Fields S. A tetO Toolkit To Alter Expression of Genes in Saccharomyces cerevisiae. ACS Synth Biol. 2015;4: 842–852.

35. Kachroo AH, Laurent JM, Akhmetov A, Szilagyi-Jones M, McWhite CD, Zhao A, et al. Systematic bacterialization of yeast genes identifies a near-universally swappable pathway. Elife. 2017;6. doi:10.7554/eLife.25093

36. Houten SM, Koster J, Romeijn G-J, Frenkel J, Di Rocco M, Caruso U, et al. Organization of the mevalonate kinase (MVK) gene and identification of novel mutations causing mevalonic aciduria and hyperimmunoglobulinaemia D and periodic fever syndrome. Eur J Hum Genet. 2001;9: 253–259.

37. Samkari A, Borzutzky A, Fermo E, Treaba DO, Dedeoglu F, Altura RA. A novel missense mutation in MVK associated with MK deficiency and dyserythropoietic anemia. Pediatrics. 2010;125: e964–8.

38. ClinVar. In: National Center for Biotechnology Information (NCBI) - ClinVar Database - MVK [gene] [Internet]. [cited 20 Mar 2021]. Available: https://www.ncbi.nlm.nih.gov/clinvar/?term=mvk%5Bgene%5D

39. Kachroo AH, Laurent JM, Yellman CM, Meyer AG, Wilke CO, Marcotte EM. Systematic humanization of yeast genes reveals conserved functions and genetic modularity.

Supplementary Materials. Science. 2015.

40. Garí E, Piedrafita L, Aldea M, Herrero E. A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae. Yeast. 1997;13: 837–848.

41. Lane S, Xu H, Oh EJ, Kim H, Lesmana A, Jeong D, et al. Glucose repression can be alleviated by reducing glucose phosphorylation rate in Saccharomyces cerevisiae. Sci Rep. 2018;8: 2613.

42. Sprengel R, Hasan MT. Tetracycline-controlled genetic switches. Handb Exp Pharmacol. 2007; 49–72.

43. Flick JS, Johnston M. Two systems of glucose repression of the GAL1 promoter in Saccharomyces cerevisiae. Mol Cell Biol. 1990;10: 4757–4769.

44. Lenstra TL, Coulon A, Chow CC, Larson DR. Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription. Mol Cell. 2015;60: 597–610.

45. Lee ME, DeLoache WC, Cervantes B, Dueber JE. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. ACS Synth Biol. 2015;4: 975–986.

46. Lambert G, Kussell E. Memory and fitness optimization of bacteria under fluctuating environments. PLoS Genet. 2014;10: e1004556.

47. Alberti S, Gitler AD, Lindquist S. A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast. 2007;24: 913–919.

48. Lamesch P, Li N, Milstein S, Fan C, Hao T, Szabo G, et al. hORFeome v3.1: A resource of human open reading frames representing over 10,000 human genes. Genomics. 2007. pp. 307–315. doi:10.1016/j.ygeno.2006.11.012

49. Souza-Moreira TM, Navarrete C, Chen X, Zanelli CF, Valentini SR, Furlan M, et al. Screening of 2A peptides for polycistronic gene expression in yeast. FEMS Yeast Res. 2018;18. doi:10.1093/femsyr/foy036

50. Koch EN, Costanzo M, Bellay J, Deshpande R, Chatfield-Reed K, Chua G, et al. Conserved rules govern genetic interaction degree across species. Genome Biol. 2012;13: R57.

51. Srivas R, Shen JP, Yang CC, Sun SM, Li J, Gross AM, et al. A Network of Conserved Synthetic Lethal Interactions for Exploration of Precision Cancer Therapy. Mol Cell. 2016;63: 514–525.

52. Mnaimneh S, Davierwala AP, Haynes J, Moffat J, Peng W-T, Zhang W, et al. Exploration of essential gene functions via titratable promoter alleles. Cell. 2004;118: 31–44.

Supplementary Materials

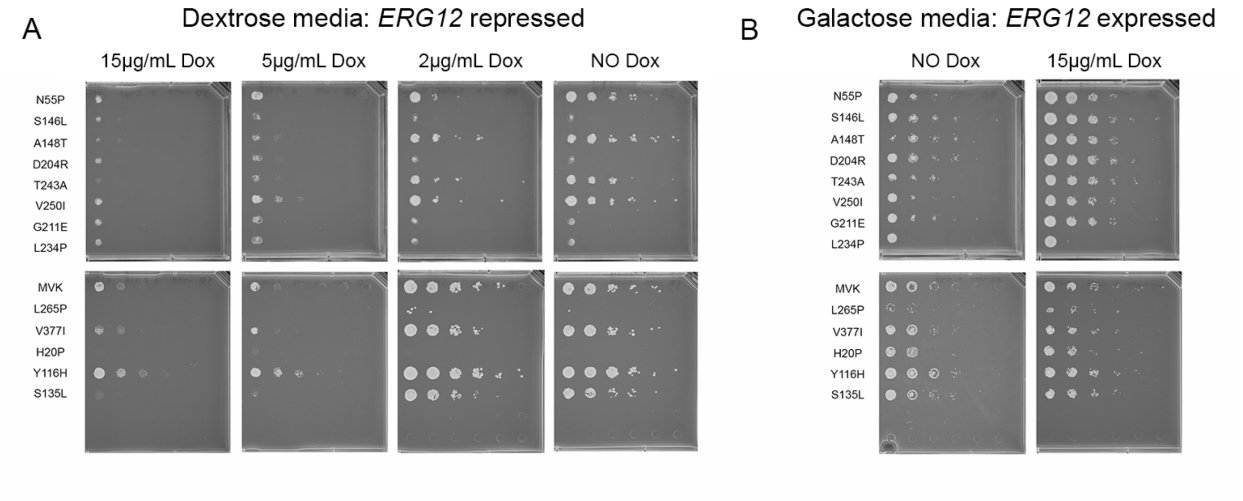

Figure S1. Complementation Assay Using Tetracycline-Regulated System (Precultured in Dextrose without Doxycycline)

Wild type MVK and several of its variants transformed into the pGAL-*ERG12* + tTA strain. This strain allows the expression of the essential gene *ERG12* on media containing galactose (right) while repressing the gene on media containing dextrose (left). Growth on dextrose can only occur after the successful replacement of *ERG12* with *MVK*. Concentrations of doxycycline repress the human *MVK* gene to different levels. These results are distinct from Figure 16 because the samples were precultured in dextrose without doxycycline, while the previous assay was precultured in dextrose with doxycycline.

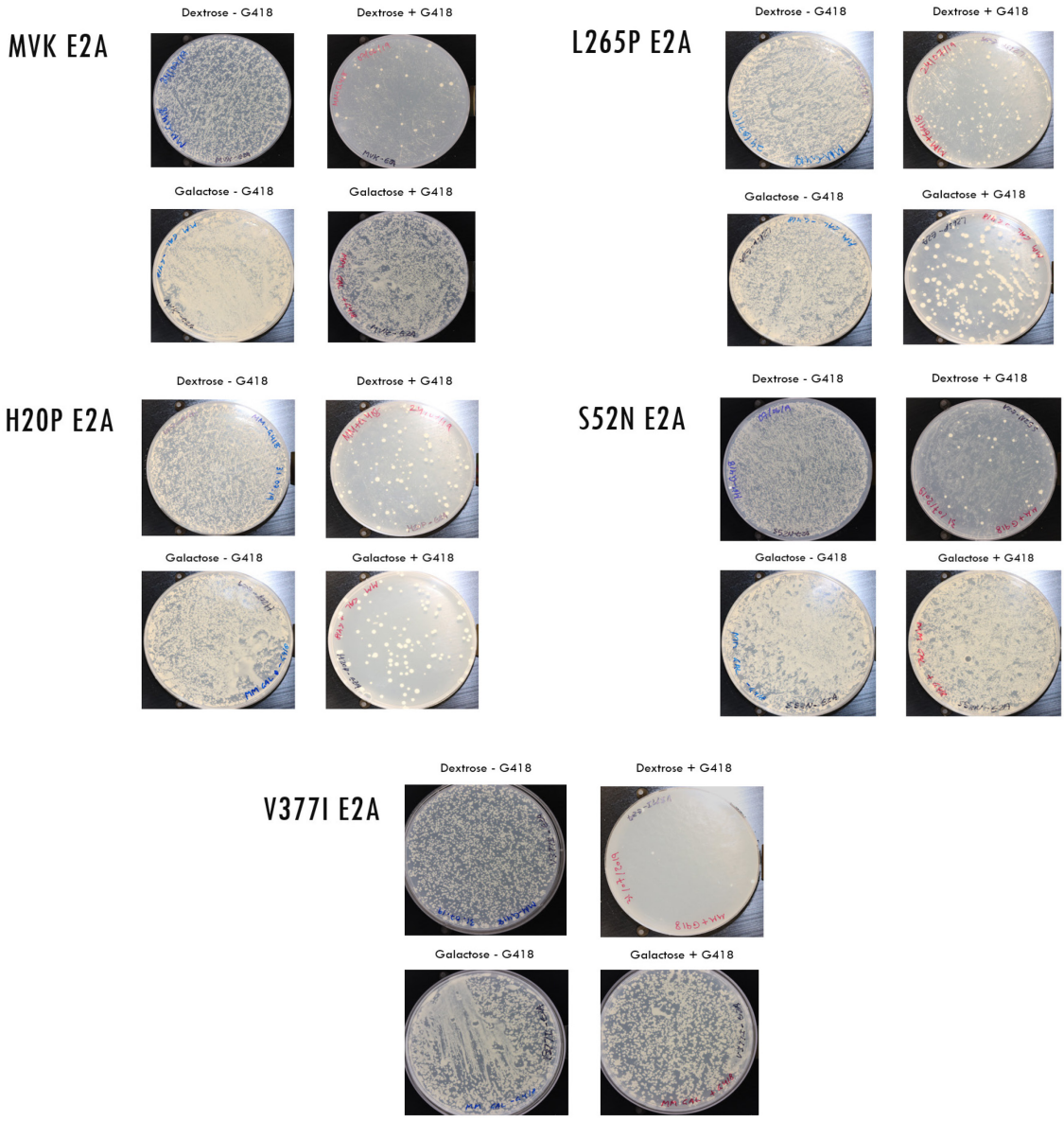

Figure S2. Growth Assay of Knockout *ERG12* Strain Transformed With MVK-EGFP Variants

Haploid strain with either a knocked out copy of *ERG12* (plates with G418) or wild type copy (plates without G418)[32]. Strains were previously diploid and transformed with plasmid containing MVK-E2A-EGFP under a galactose-inducible promoter. These plates represent haploids, following sporulation. Galactose media allows for expression of the human MVK-E2A-EGFP fusion gene and its variants. Dextrose media does not allow expression of MVK-E2A-EGFP. The condition galactose + G418 represents the knocked out yeast gene *ERG12* and expression of the human gene MVK. In this condition, growth is rescued through successful complementation of the yeast gene by the human gene.

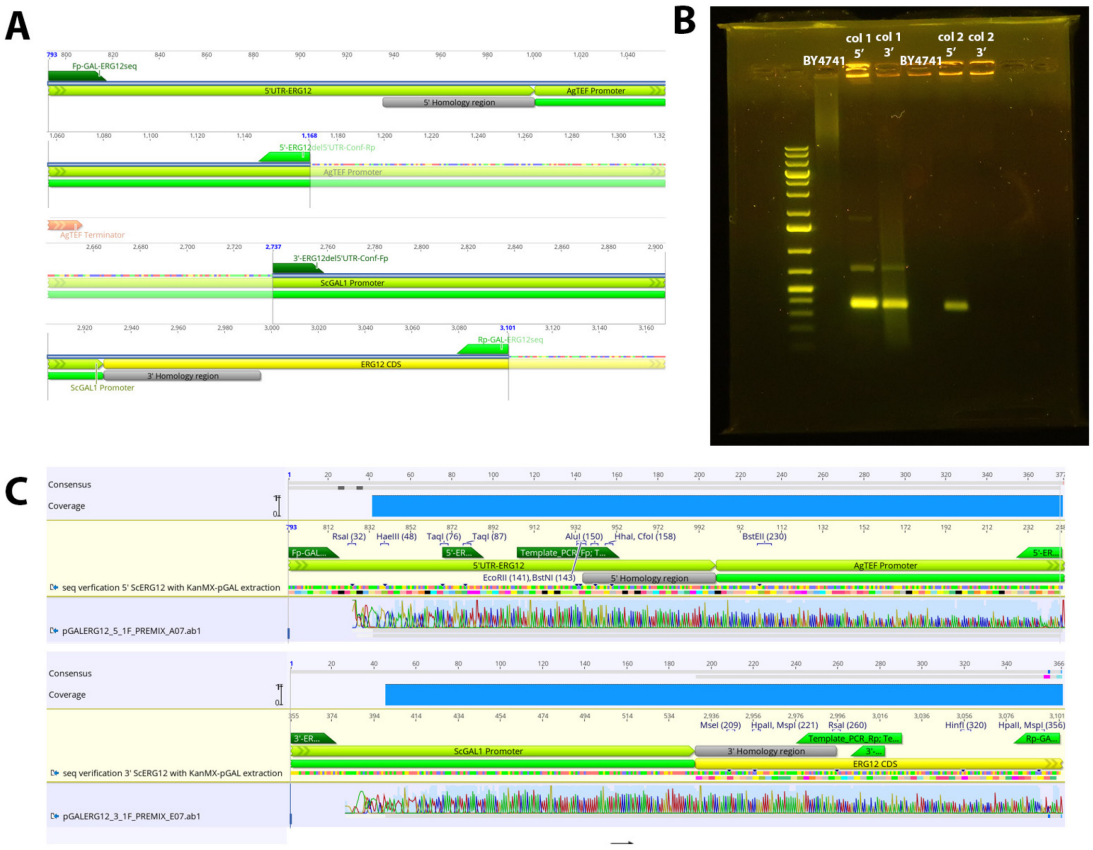

Figure S3. Confirming the insertion of pGAL at *ERG12* locus by PCR and Sanger Sequencing. **(A)** The Geneious software files show the illustration of the *ERG12* locus. The primers used for adding additional sequence homology to facilitate integration via homologous recombination are indicated. **(B)** The sequencing primers that bind outside the region of homology used for insertion were used for confirmation PCR that yield an amplification only in the case of an engineered locus. The PCR was sequence verified. **(C)** The sequencing reads were aligned with the reference map of the engineered locus (Top, 5' end & Bottom, 3' end) confirming the replacement of the native promoter by pGAL.

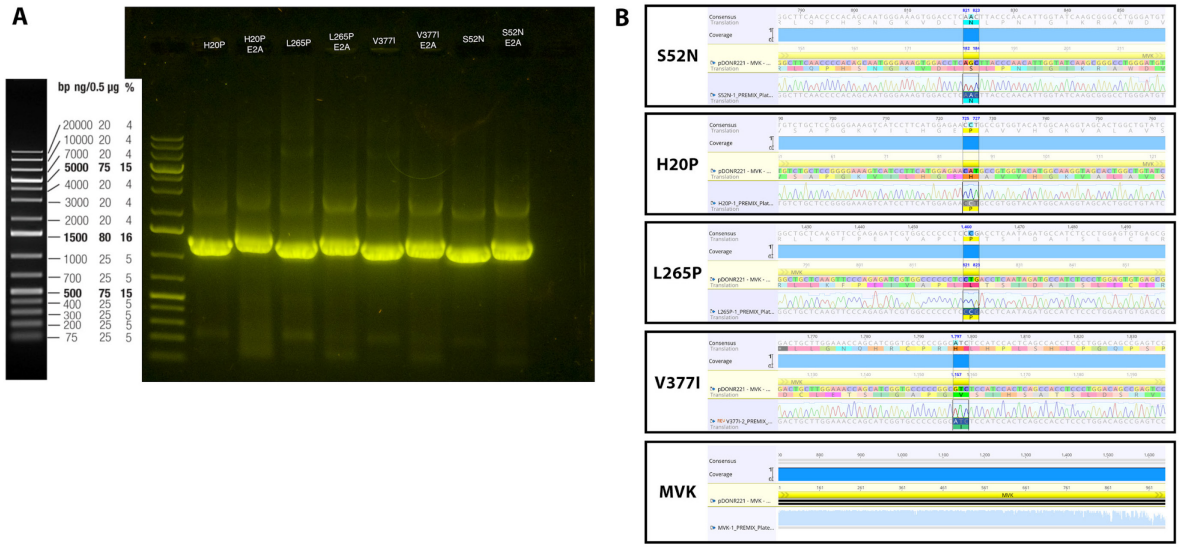

Figure S4: MVK and Several Variants Cloned into pDONR221 Vector and Sequence Verified

(A) Gel electrophoresis image of MVK variants amplified with and without E2A. Expected band size with E2A is 1315 bp. Expected band size without E2A is 1255 bp. These bands also include the attB1 and attB2 sites. **(B)** Geneious files showing sequencing results for MVK and several variants cloned into pDONR221.