Characterization of *ORF19.7608* (*PPP1*), a Biofilm-induced Gene Encoding a Protein of Unknown Function in *Candida albicans*

Nmerichukwu Iwuchukwu

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

Montréal, Québec, Canada

February 2024

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Nmerichukwu Iwuchukwu

Entitled: Characterization of *ORF19.7608*(*PPP1*), a Biofilm-induced Gene Encoding a Protein of Unknown Function in *Candida albicans*

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:

_____Chair

Dr. A. Piekny

_____Examiner

Dr. D. Kwan

_____Examiner

Dr. I. Benoit-Gelber

_____ Supervisor

Dr. M. Whiteway

Approved by _____

Dr. Robert Weladji (Graduate Program Director)

_____2024 _____

Dr. Pascale Sicotte (Dean of the Faculty of Arts and Science)

ABSTRACT

Characterization of *ORF19.7608*(*PPP1*), a Biofilm-induced Gene Encoding a Protein of Unknown Function in *Candida albicans*

Nmerichukwu Iwuchukwu

Biofilms are a major source of pathogenicity in *Candida albicans* and are connected to about half the deaths due to systemic candidiasis. Biofilm formation is controlled by a transcriptional network and involves over a thousand genes including a small gene we have named *PPP1* (Punctate Pattern Protein 1). This is a biofilm-upregulated gene encoding an unexamined protein found only in *Candida albicans* and its sister species *Candida dubliniensis*. In this study, we characterize this protein by analysing its effect on biofilm formation and its cellular localization. We focus on fluorescent microscopy to identify its subcellular localization and links to other cellular processes. First, we tagged Ppp1 with the green fluorescent protein (GFP) and imaged the protein expression pattern. This expression pattern was used to identify subcellular compartments with similar patterns and direct knockouts and tags of proteins defining these domains. Next, we screened the mutant and complement strain through stress and invasive assays and confocal microscopy to identify a distinct phenotype. Disruptions and tagging of Sur7 defining the eisosome and Arc35 defining the actin cytoskeleton suggested that Ppp1 was not a component of either the eisosome or actin cytoskeleton. However, tagging Sed5, a Golgi defining protein, identified colocalization between the puncta. Since there is no clear distinction between Golgi compartment proteins and secretory proteins, we tagged and analyzed Sap2 and compared data with results for Ppp1. We also tagged Orf19.6274 and Orf19.4654, proteins with similar characteristics to Ppp1 and compared their patterns to Ppp1 puncta. These experiments support a model that Ppp1 is localized in the Golgi.

Acknowledgements

I would like to sincerely appreciate everyone that has helped this dream become a reality. It is because of you that I can do this.

My heartfelt gratitude to Dr. Malcolm Whiteway my supervisor for his immense support and dedication to this project. His guidance, suggestion and directions are the reason I have this thesis. I would like to appreciate Dr. David Kwan and Dr. Isabelle Benoit-Gelber my thesis committee for their support and suggestions throughout this project.

I would like thank Dr. Chris Law for his guidance in the imaging experiments. He oversaw the microscopy aspect of this project, suggesting imaging techniques and analysis tools. This project would have not been successful without the input of the Whiteway lab members, present and former. I would like to especially thank: Dr. Anna Carolina Borges Pereira da Costa who laid the foundation for this project by identifying the gene, constructing some strains and doing some biochemical assays; Dr. Raha Parvizi Omran who guided me throughout the project, helping me troubleshoot and find answers and Dr. Manjari Shrivastava who helped me get started on the project and continued to support me.

I would like to appreciate the Mitchell lab most notably Min-Ju Kim and Dr. Aaron Mitchell for their help with the confocal microscopy. I thank Dr. Bernard Hube for gifting us the *sap2Δ/Δ* strain.

I want to thank my support system, my dearest family, and friends for encouraging, uplifting me and reminding me of who I am. Thank you for understanding the craziness that is my life and helping me make the best of my free time.

Lastly, I want to thank NSERC and SynBioApps Concordia for the fellowship that covered my living expenses while I focused on doing science.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Gene of Interest	GOI
Fluorescent Protein	FP
Clustered regularly interspaced short palindromic repeats	CRISPR
Polymerase Chain Reaction	PCR
Wildtype	WT
Differential Inference Contrast	DIC
Green Fluorescent Protein	GFP
Red Fluorescent Protein	RFP
Candida Genome Database	CGD
Albicans specific gene	ASG
Amino acid	AA
Reactive oxygen species	ROS
5- flucytosine	5-FC
Minimum Inhibitory Concentrations	MICs
Mitogen-activated protein kinase	MAPK
heat shock proteins	HSPs
secretory aspartyl proteinases	SAPs
Endoplasmic reticulum	ER
Glycosyl phosphatidylinositol	GPI
Regions of Interest	ROIs
Gastrointestinal tract	GIT
Manders' Correlation Coefficient	MCC

Chapter 1: Introduction

Living organisms are grouped into three domains, Eukarya, Bacteria and Archae. Eukarya is a domain classification for organisms possessing a nucleus [1]. This taxon precedes that of the kingdom, the foundation of most classifications in recent times. Domain Eukarya can be classified into four kingdoms: Plantae, Animalia, Fungi, and Protista, each with distinct features.

Fungi is a kingdom with organisms that can have characteristics similar to both plants and animals. For example, most fungi are sessile – as are plants - and possess root-like mycelia which branch out to hyphae resembling plant branches. At the same time, these organisms are heterotrophs like animals and metabolize their food by secreting digestive enzymes to degrade substrates for easy absorption [3]. Fungi also possess other distinct characteristics which distinguish them from other kingdoms and classify them together. A major one is the chitin-glucan complex cell wall, one of the major features targeted by pharmaceutical compounds used for anti-fungal treatments [4]. Chitin, an abundant polysaccharide, also forms the exoskeleton of arthropods while glucans are found in plants as starch, amylose, and amylopectin [4].

The ecological, biological, and industrial importance of fungi cannot be overemphasized. In soil, fungal mycorrhiza embedded in the roots of plants help in moisture retention, nutrient acquisition, and pathogen resistance, thereby improving crop yields [5]. The industrial applications of fungi includes Saccharomyces yeast involved in the making of bread and fermented beverages such as beer and wine, as well as various fungi implicated in the production of different cheeses. Fungi are part of the human microbiome and help build host immunity [6], but they can also be pathogens and cause superficial or invasive infections.

Fungi can be classified into 8 phyla each with distinct features and relationships [7]. The ascomycota phylum is home to *Saccharomyces cerevisiae*, the baker's, brewer's and powerful experimental model yeast, as well as *Candida albicans,* the most common fungal pathogen; both belong to the family Saccharomycetaceae [8]. Organisms in the ascomycota phylum can be either saprophytic or parasitic and can undergo sexual or asexual reproduction. Some of them are edible mushrooms, some form symbiotic relationships with other organisms, some cause diseases in plants and others are lethal pathogens [8]. Candida

species are pathogenic fungi that cause different types of candidiasis sometimes leading to sepsis (candidemia). *Candida albicans* is a human commensal in this genus with the highest prevalence of infection of humans among the *Candida* spp. [9].

1.1. *Candida albicans*, a part of normal human flora, can be an opportunistic pathogen

Humans are able to live a healthy life and fight infections in part with the help of microorganisms located in specific areas of the body. These organisms are called the normal flora, and they help build the immune system and fight off pathogens entering the body, thus preventing infections [6]. The body's microbiota or flora is found in almost all parts of the body with the exception of internal organs and fluids [6]. This flora is made up of bacteria, such as *Staphylococci, Streptococci, Lactobacilli* and even fungi such as *Candida* [6,10]. These microorganisms are found in the mouth, nose, skin, gastrointestinal tract (GIT), and the urogenital tract.

Candida as a part of the normal flora is found in the mouth, skin, GIT, and vagina [11]. When the body's pH is altered, the microbiota composition is changed or the immune system is compromised, it can become a pathogen. It is an opportunistic pathogen in the sense that it will only cause infection when it takes advantage of the imbalance in the host to colonize it by overgrowing and suppressing the flora, the responses of the flora and even the body's immunity. Because it is part of the microbiota of humans, *Candida albicans* morphology and biological processes are well adapted to the human body, which help it in evading immunological responses. It causes both superficial and invasive infections in all age groups from diaper rash in babies to athlete's foot, vaginal candidiasis, oral thrush, and systemic candidiasis affecting the heart, bone, liver, and spleen [14]. *Candida* spp. like *albicans, tropicalis*, and *auris* cause numerous infections, with *albicans* accounting for about half of those [12].

Candida albicans is a pathogenic fungus and a leading cause of nosocomial infections, with a mortality rate as high as 50%. [13]. Most of the nosocomial infections are as a result of implanted devices such as catheters and prosthesis. These implants provide a conducive environment for *Candida* biofilms to grow as the fluids on them are rich in glucose and proteins. Candida cells attach to these devices as biofilms with the aid of adhesion enzymes,

then invade the patient and cause disease. The combination of *Candida albicans* as part of the human microbiome and being able to switch to a lethal pathogen makes it a top priority in biomedical research. Current efforts have produced 4 classes of antifungals each targeting different cellular processes in the pathogen. Echinocandins inhibit 1-3, β-D-glucan affecting the cell wall synthesis; polyenes inhibit ergosterol synthesis needed for cell membrane formation; azoles inhibit cytochrome P450-dependent 14α-lanosterol demethylase inhibiting ergosterol biosynthesis and pyrimidine analogs disrupt DNA and RNA synthesis [15,16,17].

1.2 Virulence factors help *Candida* become pathogenic

Candida albicans can change interaction with its host from commensal to pathogenic through a variety of environmental factors, cellular processes, genetic responses, and induction pathways. A major virulence factor is dimorphism, the phenotypic switch from yeast to hyphae. This morphological change allows cells to adhere to and invade various tissues and to cause infection as the hyphal form is responsible for disseminated candidiasis. [12, 18]. It can also switch from the white yeast form to the opaque oval form for mating (regulated by the Mitogen-activated protein kinase (MAPK) pathway). Other virulence factors include biofilm formation, hydrolytic enzymes, genome plasticity, adherence and invasion enzymes, metabolic plasticity, heat shock proteins (HSP), and antifungal resistance [12].

Hydrolytic enzymes like hemolysins, phospholipases, candidalysin and proteinases are involved in tissue damage. Hemolysins help *Candida* acquire iron by breaking down hemoglobin in the host's blood [19]. Phospholipases aid virulence by cleaving the ester bond of the cell membrane leading to cell lysis [20]. Proteinases cleave proteins with a major class being the secretory aspartyl proteinases (SAPs). Candidalysin (Ece1) a toxin lyses the host epithelial lining resulting in cell damage and mucosal infections [21].

HSPs are proteins that give *Candida* metabolic plasticity, cushioning the effect of stress factors like heat, osmotic stress, oxidative stress, nutrient depletion, and pH changes [22]. They act as chaperones preventing protein denaturation and are upregulated in the yeast form of *Candida* [12]. *Candida albicans* has genes like Pma1 that help it adapt to alkalinity by releasing ammonia [23, 24]. Genome plasticity helps *Candida* to survive in different

environments and conditions. The *Candida* genome is regulated in such a way that specific genes are induced under certain conditions as an adaptation to the host or to mitigate the effect of lytic factors. This adaptation helps it evade host immunity, stress factors and antifungals.

1.2.1 Biofilm formation in *Candida albicans*

A biofilm is a community of microorganisms attached to a surface and enclosed in an extracellular matrix [25, 26]. These microorganisms form colonies of cells that can resist the host's immune response as well as antimicrobial therapy. Adherence of biofilms to surfaces like catheters and prostheses make them a major cause of nosocomial infections. Oftentimes, biofilms are a network of different bacteria and fungi that colonize the host. Many microorganisms can make biofilms and *Candida albicans* is among that group. *Candida albicans* biofilms affect patients from different categories: immunocompromised individuals like HIV and cancer patients, elderly patients with dentures, and even young patients with prostheses and catheters. Biofilm cells are resistant to antifungals and host immunity so pose a huge problem in healthcare. In the United States, more than 50% of catheters become infected with *Candida* biofilms and billions are spent yearly to treat these infections [27].

Candida albicans biofilm formation is regulated by environmental factors like osmotic stress, oxidative stress, and heat shock, and genetic factors like secondary messengers, transcription factors, and regulatory RNAs [28, 29]. Other factors like drug efflux pumps, quorum sensing, host immunity, and extracellular matrix also influence its formation. A transcriptional network controls biofilm formation and each phase in the developmental cycle is regulated by a specific set of genes [30, 31]. Biofilm formation takes between 24 – 48 hours and occurs in four stages as discussed below [32, 30, 33]:

1. Cell Adherence: Within few hours of inoculation or infection, Candida cells attach to host tissues or abiotic surfaces forming the basal layer, with the help of adhesion enzymes like the Als family [32].
2. Cell Proliferation and Hyphae Formation: After cells have attached properly to host tissues or implant surfaces, they begin to replicate, increasing their numbers and branch into hyphae so

they can pass through tissues and disseminate infection. This is also called the initiation step because of the hyphae formation [32, 33].

3. Production of Extracellular Matrix (ECM): In this stage, the biofilms are mature having grown for up to 24 hours. This phase is marked by the formation of an extra polymeric substance called the ECM which acts as a capsule for the cells [30,32,33].
4. Accumulation of ECM and Cell Dispersion: This is the final stage in biofilm formation where more ECM is produced and some mature biofilm cells dislodge and are dispersed to other tissues or surfaces to initiate another biofilm [32,33].

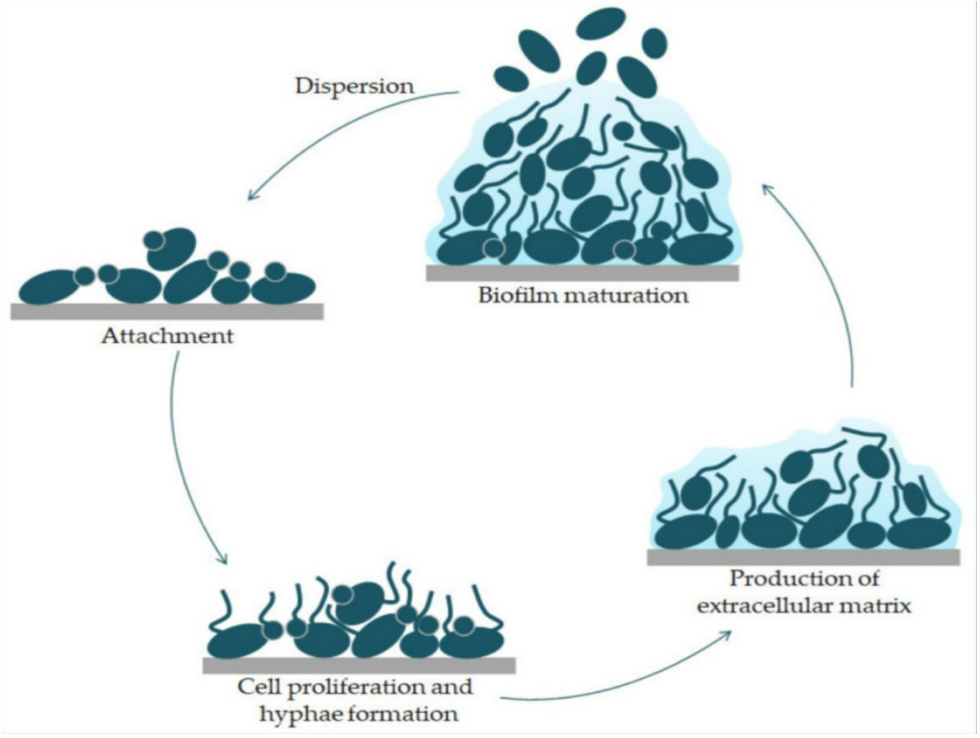


Figure 1: Developmental cycle of biofilm formation in *Candida albicans.* Life cycle begins with yeast cells attaching to biotic or abiotic surfaces followed by cell elongation to hyphae and proliferation. In the third stage, cells have matured and formed the ECM and the last stage sees the accumulation of ECM and dispersal of mature cells [32].

1.3 Genetic regulation of *Candida albicans* biofilms

Candida albicans has a regulatory network of genes differentially expressed at different stages of biofilm formation. In the adherence stage, there is an abundance of adhesins such as the Als proteins Als2 and Als3 [30]. During hyphae formation, Efg1, Tec1, Suv3, Nup85, Mds3 and Kem1 are involved [34]. In the third stage where the ECM is produced, negative regulators such as Zap1 and Adh1 repress formation of β 1,3 glucan and biofilm formation [30, 35]. Sometimes, expression of these genes and transcription factors are identified in more than one stage of biofilm formation.

A useful strategy to understanding biological processes is through extensive study of the genetic networks that drive them [31]. In biofilm formation, investigating the regulatory network of biofilm formation involves first identifying the key transcription factors (TFs) and subsequently the genes regulated by them. This approach identifies hundreds of genes that could otherwise have gone unnoticed [31], and studying these genes has possible clinical importance as antifungals targeting the products of these genes can be found, improving antifungal therapy. A study by Nobile et al. 2012 investigated the transcriptional network regulating *Candida albicans* biofilms with the aim of mapping these TFs controlling genes involved in biofilm formation [31]. More than 1000 genes were identified, many of which have already been extensively studied. However, several such regulated genes are uncharacterized; one of the uncharacterized genes designated *ORF19.7608*, encodes a protein, found in only *Candida albicans* and *Candida dubliniensis,* that has an N-terminal signal peptide.

1.4 Signal peptides, their role in protein transport, and their distribution in protein families

Signal peptides are short, typically hydrophobic amino acid sequences in nascent proteins associated with the secretory pathway, including proteins that are localized in the secretory organelles like the endoplasmic reticulum (ER) and Golgi, proteins that are secreted, and proteins localized in other subcellular compartments [36]. They are N-terminal sequences that help in the transport and secretion of proteins. As the name implies, signal peptides are

recognition signals for the signal recognition particle (SRP) that binds to the ribosome-nascent chain complex (RNC) [37]. This interaction targets the ribosome-nascent chain-SRP complex to the ER, where protein translocation occurs [37]. GTP dependence ensures protein transport to the correct subcellular location by coordinating the presence of a signal sequence on the ribosome with the availability of a translocon in the membrane [38]. After the proteins are processed by the ER, they are transported to the Golgi for further processing, packaging, and transport to their subcellular locations.

Signal peptides are found in the Secretory Aspartyl Proteinases (SAP), a class of secreted proteins associated with virulence in *Candida albicans*. SAPs have been linked to proteolysis [39], cell adhesion [40] and disseminated Candida infection [41]. They are also found in cell wall proteins like Kre1 involved in the synthesis of cell wall β 1,6-glucan [42] and Cht3 a chitinase also involved in cell separation [43].

Signal peptides are also found in proteins of the secretory pathway. In the Golgi, Kex1 and Kex2, proteins with N-signal peptides are needed for cleaving α-factor to form a mature protein and Arf1 is an ADP-ribosylation factor needed for the hydrolysis of GTP [44]. Some ER proteins like yeast α-glucosidase I (Cwh41), a type II integral membrane N-glycoprotein needed for protein glycosylation [45]; Kar2, a heat shock protein in the Hsp70 family with an ER retention signal at the carboxyl terminus [46]; Mpd1, a member of the protein disulfide isomerase (PDI) family involved in formation of disulphide bonds for protein folding [47] and Dpm1 (dolichol phosphate mannose synthase) needed for GPI anchoring, N-glycosylation, and O-mannosylation [48] also have an N-signal sequence. In the plasma membrane, proteins like Gas1, a Glycosyl phosphatidylinositol (GPI) anchored cell surface protein [49] and Ste2, a receptor for α-factor mating pheromone, MF-α also have a signal peptide.

1.5 The Golgi and its role in biofilm formation

The Golgi is part of the group of organelles that make up the endomembrane system and is an element of the membrane trafficking system consisting of the Golgi, ER, endosomes, lysosomes, and the plasma membrane [50]. It lies at the heart of the secretory pathway and is

involved in cellular processes such as protein glycosylation, sorting into vesicles, and trafficking of newly synthesized membrane and secretory proteins to cellular destinations [50]. Its location within the membrane trafficking system helps facilitate transport of proteins to target sites and post translational modification.

This organelle has a clearly defined structure comprising stacks of flattened membrane discs called cisternae, layered on each other to form what is termed the Golgi stack. Golgi cisternae mature from the early/cis to the late/trans Golgi network cisternae [51]. This maturation process allows cargo proteins to transverse the Golgi by remaining within the maturing cisternae after which they are transported to their target membrane by the vesicles [51]. Each stage of maturation is characterized by resident proteins, some of which represent characterized markers defining each phase. The early cisternae are defined by proteins like Sed5, Vrg4 and the late cisternae by proteins such as Sec7, Sys1, and Sft2 [51, 52].

Protein trafficking through the Golgi begins with the transport of synthesized proteins from the ER into the cis-Golgi. These proteins then pass through the Golgi and are either secreted or retained. Retention of proteins in the Golgi is through iterative rounds of anterograde and retrograde transport. This is influenced by protein-protein interaction, affinity of proteins for the Golgi lipid bilayer and binding affinity for the COPI and COPII vesicle coat complexes [56]. The transport of material in the membrane trafficking system involves the packaging of proteins into transport carriers called vesicles [50]. This process occurs through vesicular intermediates that bud from a donor compartment and fuse with an acceptor compartment [53]. The budding of vesicles and cargo protein selection are mediated by protein coats [54] while vesicle targeting, and fusion are dependent on SNARE proteins [53]. The final step in this protein transport is the fusion of the vesicle to the acceptor membrane, mediated by the SNARE proteins. In this step, a vesicle carrying "v-SNARE" binds to the complementary "t-SNARE" on the target membrane releasing the cargo [55].

The Golgi does not merely serve as a tunnel for protein transport to target membranes, it is also involved in post-translational modifications such as glycosylation. Glycosylation helps in proper folding of protein which makes them functional and aids their proper localization. The glycosylation process adds sugar residues to the proteins and is critical for the synthesis of the glycoproteins that make up the ECM of Candida biofilms and are thus a major source

of antifungal and immunological resistance. This process is regulated by glycotransferases which attach sugar molecules on the nitrogen atom of the side chain of asparagine in N-glycosylation and the oxygen atom of the side chain of serine and threonine in O-glycosylation. Other post translational modifications in the Golgi include N-acetylation, phosphorylation, proteolysis, methylation, lipidation etc.

1.6 Predicting protein localization based on signal sequences is difficult

A signal peptide is a characteristic of newly synthesized proteins representing both secreted proteins and integral proteins of various subcellular compartments, including the secretory pathway. It is an N-terminal short hydrophobic region at the start of a protein sequence. Protein retention in the secretory pathway and sorting to different organelles depends on various factors like amino acid composition, signal motifs, transmembrane domains, cytoplasmic tails etc. [56]. It can also depend on protein interactions with transport vesicles, and integral membrane proteins [56]. Even within a family of proteins, sorting and retention mechanisms differ [56]. Protein retention within the secretory pathway is a dynamic process with each protein having its own mechanism.

The secretory pathway and its proteins were initially characterized decades ago. Newly identified proteins such as Orf19.7608 can find their characterization under the aforementioned retention mechanisms. On the other hand, secretory proteins like Sap2 have also been extensively studied but nothing is known about the protein expression pattern and if this is consistent with a group of proteins. To understand the protein localization and biological function of proteins, visualization techniques can be employed with the goal of establishing subcellular localization.

1.7 Fluorescent microscopy: a visualization technique to characterize Orf19.7608

Identifying the organelle a protein is localized in is one step in understanding the role of a protein in a cellular process. Microscopy has been in existence for centuries, advancing from simple light microscopes to electron microscopes, confocal microscopes, and fluorescent

microscopes [57]. Fluorescent microscopy is an imaging technique that can be used for characterizing organellular compartmentation. It is used to study the subcellular resolution of proteins, thus enhancing our understanding of cell physiology [58]. Fluorescent microscopy involves various techniques like protein localization, ion transport etc. to localize proteins and understand physiological processes [58]. Wide-field fluorescence microscopy (WFFM), also called epifluorescence microscopy relies on the absorption and emission of light from fluorophores like GFP to form a contrast with the background, illuminating the tagged protein or organelle which is then captured. It elicits fluorescence from the sample using a light source, a microscope, and excitation and emission filters. This results in emitted light, of longer wavelength than the excitation which is collected by the objective lens and observed through the microscope eyepieces or by a camera followed by computer digitization [58].

During image capturing, light passes through the object and this spreads, giving the captured image a blurrier resolution, this forms background. Background constitutes an analysis problem because it can alter the minimum intensity values for the probe signals thereby significantly affecting colocalization values. Not being able to differentiate between low intensity signals and background can result in wrong colocalization coefficient values. Deconvolution is an imaging technique that recreates the image to be most similar to its parent object using the Point Spread Function (PSF) of the optical system. This procedure increases contrast, improves resolution, reduces noise to a minimum, eliminates blur and suppresses background to very low levels allowing for analysis of low intensity signals [59]. This is especially useful for images with low signal to noise ratio (fluorophore signal intensity is low), images with objects of near resolution size [59] and for image analysis involving pixel intensity or comparing signal overlap. Deconvolution is available in software like AutoQuant.

Fluorescent microscopy can be used to capture the distribution pattern of proteins. This pattern can then be compared with markers of subcellular compartments with similar patterns. To determine if the two proteins are localized in the same organelle, colocalization experiments are conducted. Colocalization between a novel protein and a compartment marker can help understand the function of a protein [60]. Colocalization depends on the

premises that repeated co-occurrence of two probes throughout a cell is an indicator of colocalization of the two proteins in the same domain [60]. This method of functional characterization has been used frequently [61, 62].

Colocalization can be visualized through merged images and further quantified through different statistical analysis parameters. The color of the image produced from merging the pixels from the two channels can shows if the two proteins colocalize in the same subcellular compartment such that a secondary color to the channels depicts a perfect colocalization while individual colors of the channels in the merged image show no colocalization. The distribution of pixels in both channels can be represented by a scatterplot called a cytofluorogram.

Colocalization can also be quantified using the Pearson's Correlation Coefficient (PCC) or Mander's Colocalization Coefficient (MCC) [60]. PCC measures the linear relationship between the pixel intensities of the two tags [60]. This is ideal for experiments where we know the biology of the signals in the two channels have a linear relationship. MCC on the other hand measures the fractional overlap of the intensity of one signal with the other [60] i.e. fractional overlap of A with B and B with A, ideal for experiments that involve compartmentalization of signals.

1.8 Thesis objective

Candida albicans is a fungal pathogen of medical importance. Biofilm formation, a virulence factor, is a major cause of infections in hospitalized patients, immunocompromised individuals, and even healthy people. This study aims to understand the role of an uncharacterized gene upregulated during biofilm formation *ORF19.7608.* **The main objective is to identify the subcellular location of this protein encoded by this gene, investigate its biological function and evaluate its role in biofilm formation**. To do this, we made a null mutant of *ORF19.7608*, a complemented strain of the null mutant by reintegrating *ORF19.7608* at its native promoter, and a tagged strain where the *ORF19.7608* coding sequence was fused to GFP. We investigated the subcellular location of the fusion protein in strains deleted for components of the eisosome (*SUR7)* and the actin cytoskeleton

(*ARC35*). We also made double tagged strains in *ORF19.7608* for proteins localized to the Golgi (*SED5*), *SUR7* and *ARC35.* This was used to identify the protein expression pattern of Orf19.7608 and infer its subcellular localization. We compared the Orf19.7608-GFP pattern to that of the secreted protein Sap2 to verify Orf19.7608 is localized in an organelle and not passing the secretory pathway. We evaluated the effect of antifungals and stress factors on biofilm formation and visualized biofilm formation in the mutant and complement strain using confocal microscopy. We also made tags of similar proteins (small, N-signal peptide proteins upregulated in biofilm formation) with Orf19.7608 and compared their expression pattern.

Chapter 2: Materials and Methods

2.1. Bioinformatic analysis of *ORF19.7608*

We used database searches to investigate the structure and role of this uncharacterized gene found highly expressed in biofilms. To begin, an ortholog synteny of *ORF19.7608* established by the Butler lab was retrieved through the Candida Genome Database (CGD). We put the amino acid sequence through NCBI blast to check across all species for proteins with sequence similarity. We also checked protein databases like Uniprot to see if biological or functional annotations have been made or other unique characteristics identified. Lastly, we looked at the predicted 3D structure using AlphaFold.

2.2 CRISPR constructs and strains:

To investigate the role of Orf19.7608 in biofilm formation, a mutant strain, a complemented strain and a strain with a GFP tagged version of Orf19.7608 were obtained from Dr. Anna Costa (unpublished data). Other strains were constructed using the transient CRISPR-Cas9 system protocol [64]. We constructed strains N120, N130, N131, N140, N150, N160, N170 (see table 1). Cas9 was PCR amplified from pV1093. Benchling was used to design the sgRNAs, the repair DNAs and to visualize the designed constructs. The sgRNAs were chosen using the CRISPR feature which gives a pool of sgRNAs for a selected sequence showing on and off target scores, an indicator of guide efficiency, and selecting the ones with the highest on target scores closest to my target.

CRISPR was used to knock out *ARC35* and to introduce fusion constructs to selected encoded proteins. For the disruption, two guide RNAs were chosen from the genomic DNA region and repair DNA was constructed using upstream and downstream primers suitable to direct the selectable marker *ARG4* to *ARC35* (80 nucleotides from *ARC35* upstream or downstream and 20 nucleotides from *ARG4*).

In the tags, the region of interest (ROI) for fusion was downstream of the gene of interest (GOI). Guide RNAs were chosen from 200bp downstream these GOIs. This allowed integration of the fluorescent protein (FP) without disruption of protein expression of the

GOI. In preparing the repair DNA for the tags, a selectable marker was linked to the FP by PCR to enable identification of transformants. To do this, the repair DNA is constructed in three rounds of PCR similar to the formation of the sgRNA cassette with the first round involving the amplification of *URA3* from pFA-GFP-CaUra3. This forms a single joint template with Scarlet from CIp10-mScarlet-I IDT in the second round of PCR. The third round forms the cassette, introducing the homology arms complementary to downstream the GOI.

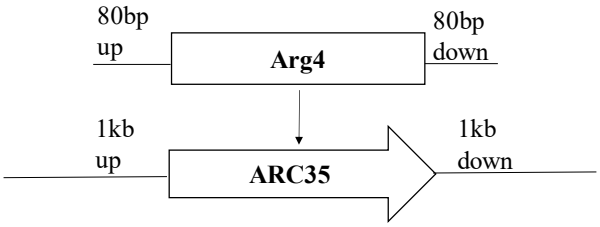


Figure 2: Schematic representation of gene deletion of *ARC35* in Orf19.7608-GFP

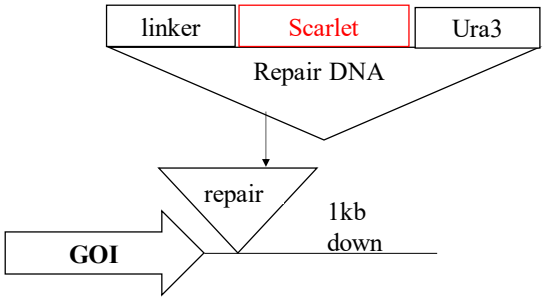


Figure 3: Schematic representation of gene tagging in Orf19.7608-GFP and SN148

Table 1: Strains used in this study

Strains	Parent	Description	Source
N101	SN148 a/α	*orf19.7608*Δ::*ARG4*/*orf19.7608*Δ::*ARG4*	Anna Costa (unpublished)

N100	N101	*ORF19.7608*/*ORF19.7608*; *his1*/*his1*; leu2/leu2; *arg4/arg4*; *ura3*::imm434/*ura3*::imm434; *iro1*::imm434/*iro1*::imm434 + pV1093-gRNA *ARG4*	Anna Costa (unpublished)
N110	SN148a/α	*ORF19.7608-GFP-*::*HIS1*/*ORF19.7608-GFP::HIS1*; *leu2/leu2; arg4/arg4; ura3*::imm434/*ura3*::imm434; *iro1*::imm434/*iro1*::imm434	Anna Costa (unpublished)
N121	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1; sur7Δ/LEU2::sur7Δ/LEU2; arg4/arg4; ura3*:: imm434*/ura3*::imm434*; iro1*::imm434*/iro1*::imm434	Anna Costa (unpublished)
N120	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1*; *SUR7-Scarlet::URA3/ SUR7-Scarlet::URA3; arg4/arg4, leu2/leu2*	This study
N130	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1 ARC35-Scarlet::URA3/ ARC35-Scarlet::URA3; arg4/arg4, leu2/leu2*	This study
N131	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1 arc35Δ/ARG4::arc35Δ/ARG4; leu2/leu2; ura3* imm434*/ura3*::imm434*; iro1*::imm434*/iro1*::imm434	This study
N140	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1; SED5-Scarlet::URA3/ SED5-Scarlet::URA3; arg4/arg4, leu2/leu2*	This study
N150	SN148a/α	*SAP2-Scarlet::URA3/ SAP2-Scarlet::URA3; arg4/arg4, leu2/leu2; his1/his1*	This study
N151	CAI4	*Δ/sap2::hisG/Δsap2::hisG::URA3::hisG*	Hube et al. 1997
N160	N110	*ORF19.7608-GFP-::HIS1/ORF19.7608-GFP::HIS1; PBR1-Scarlet::URA3/ PBR1-Scarlet::URA3; arg4/arg4, leu2/leu2*	This study
N170	SN148a/α	*ORF19.4654-Scarlet::URA3/ ORF19.4654-Scarlet::URA3; arg4/arg4, leu2/leu2; his1/his1*	This study

2.3 Yeast transformation

PCR products containing *CaCAS9*, sgRNAs and repair DNA were transformed into Candida cells using an improved version of the lithium acetate protocol [65]. First, cells were grown overnight at 30^0C in YPD with uridine (10 g yeast extract, 20 g peptone, 20 g glucose, 50mg uridine, 1L water), shaking at 220rpm. They were then made transformation competent by growing to log phase (OD_{600}=0.8), washing 10ml of culture with nuclease free water, and incubating in 600ul 1x LiAc for 20mins to improve the cell's ability to absorb foreign DNA. In an Eppendorf tube, 100µl of the competent Candida cells, 10µl of single stranded salmon sperm DNA (boiled for 10mins at 98ºC), 1µg CaCas9, 1µg sgRNA and 3-6µg repair DNA was added. 600µl 1xLiAc/40%PEG was also added to the mix, and this was incubated for 3hours at 30ºC. Next, the transformation mix was put in a water bath at 44ºC for 15mins and then on ice for 1min. The mix was then centrifuged for 10mins at 6000 rpm and the pellet washed with sterile nuclease free water. The pellet was then resuspended in 900µl YPD with uridine and grown for 2hours to allow cell recovery. Another centrifugation was done to remove the media, cells were washed and resuspended in nuclease free water, and then plated on selection plates suitable to detect to the marker used in constructing the strain. Transformants were grown on selection plates at 30^0C for 3-4 days. Colony PCR was then used to screen transformants and identify successful constructs.

2.4 DNA sequencing

The *SAP2-SCARLET-URA3* fusion construct was amplified from transformants using forward primer from *SAP2* 1kb upstream and reverse primer from *SAP2* 1kb downstream to capture the entire fusion construct. The PCR product was purified to an $OD_{260/280}$ ratio of 1.7/1.8. 15µl of 30ng/µl of this amplicon was sent to plasmidsaurus for sequencing to ensure proper fusion junctions and identify any mutations that could impact translation. This sequencing data was compared to the genomic sequences of *SAP2* in SN148 and *SCARLET* in CIp10-mScarlet- I IDT.

2.5 Yeast carbon base- bovine serum albumin (YCB-BSA) assay

To confirm that tagging Sap2 did not affect its function, colonies of SN148, Sap2-Scarlet and *sap2Δ/Δ* were grown on YCB-BSA plate (23.4 g yeast carbon base, 4 g BSA per liter, 20g agar, pH 4.0) [66] at 30°C for 24 hours. Zones of clearance were observed for each strain and images of the plates were captured using Epson Perfection v500 photo scanner.

2.6 Cellular stress assay

The stress response phenotype of mutants of *ORF19.7608* was analyzed by growing cells in YPD with uridine under different stress conditions. Single colonies of the WT, mutant and complemented strains were grown overnight in YPD with uridine at 220rpm, 30^0C. Cells were washed twice in 1x PBS and a 1:10 dilution made to take the OD_{600} reading. Calculations were made for OD_{600}= 1 and serial dilutions to 10^3. 3µl of samples were spotted on plates and grown at 30^0C and 37^0C for two days (except caspofungin plates where growth was for 5 days). Strains were subjected to osmotic stress of $CaCl_2$ (400mM), NaCl (500nM), glycerol (100nM) and oxidative stress of hydrogen peroxide (3mM). Antifungal resistance to the different classes of antifungals: flucanozole (1ug/ml), 5-flucytosine (0.5ug/ml), amphotericin B (0.25ug/ml), and caspofungin (0.75ug/ml) was also determined. Growth at different temperatures sought to analyze the effect of temperature on the strains. Experiments were repeated three times for each test to evaluate reproducibility. The selected concentrations of salts, glycerol, hydrogen peroxide and antifungals were determined from literature review [67,68,69,70]. Images of plates and colonies were scanned by an Epson Perfection v500 photo scanner.

2.6.2 Invasive assay

Overnight YPD cultures of samples were washed in 1x PBS and diluted 1:10 to take an OD600 reading. OD_{600}=1 was calculated and 1:10 dilutions to 10^3 were prepared. 3ul of each sample dilution is spotted on YPD agar plates and grown at 30^0C and 37^0C for 3 days. Non-adherent cells were removed by washing plates gently under running water. Colonies that remain on the agar after washing are considered invasive. Images of plates before and after washing were taken using the Epson Perfection v500 photo scanner.

2.6.3 Crystal violet assay

Biomass accumulation in biofilms was measured by the crystal violet assay. WT, mutant and complement strain were cultured in YPD with uridine overnight at 30ºC/220 rpm. Cells were washed twice with 1x PBS and resuspended in Spider medium. The concentration of cells were adjusted to OD_{600}= 0.5 and 200µl of the inoculum transferred to a 96-well plate with a non-treated-surface for initial incubation at 37ºC for 1.5 h in static conditions to allow cell adherence. Next, non-adherent cells were washed off with 1x PBS and 200 µl of fresh Spider medium added to each well. The plates were incubated for 24 h at 37ºC/75 rpm in the shaking incubator. The biofilms were washed once with 1x PBS, air dried for 45 mins and stained with 0.4% aqueous crystal violet solution for 45 min at room temperature (RT). Next, the biofilms are washed twice with autoclaved distilled water and then de-stained with 95% ethanol for 45 mins. The de-stained solution was diluted at 1:10 in 95% ethanol and the absorbance measured at OD_{595} (Tecan Infinite M200 Pro, Grodig, Austria). This was done in triplicates and repeated to minimize experimental errors.

2.7 Biofilm and planktonic growth

To visualize protein distribution of Orf19.7608 under biofilm and planktonic conditions, cells were grown vegetatively in YPD with uridine media and in *in-vitro* biofilm conditions with Spider media. For planktonic growth, cells were grown in 50ml Falcon tubes at 220rpm, 30^0C for 12-16 hours. For biofilms, they were grown in 6 well plates at 75rpm, 37^0C for 48 hours. Cells were washed twice in 1x PBS to eliminate autofluorescence before being imaged.

Imaging to determine the phase of biofilm where Orf19.7608-GFP is abundant was done using prepared slides from biofilm wells at 1.5 hours, 8 hours, 24 hours, and 48 hours time points representative of the various phases of biofilm formation [73,74]. For filament induction, cells were grown for 16hours at 220rpm/37ºC in YPD supplemented with 10% FBS, RPMI, GlcNac and Spider media.

2.8.1. Confocal microscopy

Biofilm growth and imaging:

To assay biofilm formation in a 96-well plate (Greiner 96-well plate; catalog no. 655090), strains were inoculated to an optical density at 600 nm (OD600) of 0.05 from overnight cultures into 100µl of prewarmed 0.2% glucose containing RPMI with or without 10% fetal bovine serum (FBS). First, the cells were incubated in a shaker incubator at 37°C for 90 min with mild shaking (60rpm) to allow adherence, and then each well was gently washed twice with 1x PBS. Next, 100µl of prewarmed RPMI with or without 10% FBS was added into each well, and cells were allowed to form biofilms in a shaker incubator with 60 rpm at 37°C for 24 hours. At that time, the medium was discarded from each well, and biofilms were fixed by incubation with 100µl of 4% formaldehyde in 1x PBS for 1 hour and then gently washed twice with 1x PBS. Biofilms were stained with calcofluor white (200µg/ml in PBS) overnight at room temperature (RT) with mild shaking (60rpm), and then each well was gently washed twice with 1x PBS. To clarify biofilms in the 96-well plates, we used 100% TDE (thiodiethanol), which has a refractive index of 1.521. We removed the PBS from each well and added 100µl of 50% TDE in PBS. We incubated the 96-well plate at room temperature for an hour and then removed the solution from each well. Finally, 100µl of 100% TDE solution was added to the wells and incubated at RT for an hour, when the clarified biofilm was transparent. The biofilms for strains of SN148, N100, and N101 were imaged using a Keyence BZ-X800E fluorescence microscope [71].

Biofilm image processing:

Optical sections of the biofilms were collected in several series of planes at an 0.45- or 0.85-µm step size. The stacks were concatenated and processed using Fiji [72]. The images were processed using the Background Subtract plugin. The side view projection images were obtained by re-slicing the stack and using the maximum intensity Z-projection. The scale of the stacked side view images was adjusted based on the objective used for the imaging.

2.8.2. Fluorescent microscopy

Orf19.7608-GFP grown both under biofilm conditions and vegetatively were imaged with a Leica DM6000 epifluorescent microscope at 1000x with exposures of about 2s for increased contrast between background and images. Images were taken in DIC and FITC (GFP) channels using focus drive of 2.0 to capture images from all planes and give a final 3D view or focus drive of none for a 2D image. The captured images were collected on the volocity image acquisition software. 3D images were stacked into one using Fiji [72]. In the double tagged strains, this procedure was adjusted to include the red (RFP) channel.

2.8.3. Colocalization

Images of the double tagged strains were analyzed using Fiji for co-localization between the green signal from GFP and the red signal from Scarlet. Regions of interests (ROIs) were selected from the images to minimize background and duplicated. These images were then split into individual channels and the red and green channels merged to colocalize the pixels of GFP and Scarlet by the following steps: Image > color > split channels > merge channels. JACoP plugin in Fiji was used to plot the cytofluorogram of the pixel intensities of the two channels and quantify colocalization. To quantify colocalization, we used MCC to measure the overlap of GFP in Scarlet and Scarlet in GFP. A value of 1 shows perfect colocalization, 0 no colocalization and -1 shows signals are inversely related. To do this, GFP and Scarlet images derived from “split channels” were uploaded and analyzed by JACoP, an ImageJ plugin. To ensure colocalization was not random, we employed Costes’ randomization which shuffles pixels for 1000 rounds. For thresholding of signal intensity, Costes’ automatic threshold which assigns appropriate values per image was used. These two parameters are contained in the JACoP plugin.

Chapter 3: Results

3.1 Bioinformatic analysis identifies characteristics of Orf19.7608

A study by Nobile et al. into the transcriptional network of biofilm regulation in *Candida albicans* identified more than 1000 genes upregulated in biofilm formation [27]. Many of the highly upregulated genes, like *HWP1, ECE1,* and *OPT4* have been extensively studied. We expanded the investigation into other genes, initially *ORF19.7608,* which exhibits an RNA seq log 2 score of 8.92 for biofilm-induced-expression, one of the highest upregulated genes identified [31], but which had not been further analyzed. Subsequent analysis looked at the protein expression pattern of *ORF19.6274* and *ORF19.4654*, proteins with similar characteristics to *ORF19.7608*.

Bioinformatic analysis can provide an efficient way of identifying important characteristics of unstudied genes and proteins. We consulted the Candida Genome Database (CGD), NCBI protein blast, Uniprot, and InterPro and performed a Kyte Dolittle plot to establish a basic characterization of Orf19.7608. Orf19.7608 was defined as an uncharacterized protein with no known biological function or cellular compartment [75]. The ortholog synteny in figure 4 shows that this gene has only one ortholog, Cd36_35264 in *Candida dubliniensis*; even the closely related *Candida tropicalis* lacks any ortholog and has no gene at the syntenic location. The *Candida albicans* gene *ORF19.7608* is located on chromosome R.

Searching protein databases gives insight to protein structure, function and evolution as these repositories have experimental evidence and inferred functionality on proteins. We blasted our protein sequence against the RefSeq database to identify proteins similar to Orf19.7608. As with the CGD, the only ortholog found was in *Candida dubliniensis* with a percentage identity of 78% [76]. Looking at the protein alignment in Figure 5 we can see 2 amino acid insertions after the 49th AA and another 2 insertions after the 91st AA giving *Candida dubliniensis* 135 AA as opposed to 131 in *albicans*.

We investigated GO annotations to help understand a potential role for Orf19.7608 in biofilm formation. UniProt identified the presence of a signal peptide MKTSVVLAAALTAMNTANA representing 1-19AA [36]. This was corroborated by Kyte Dolittle plot in figure 6 [77,78]. Checking the predicted 3D modeling structure with

AlphaFold as seen in figure 7 gave a structure made up of α-helices, beta sheets and an abundance of loops [79,80].

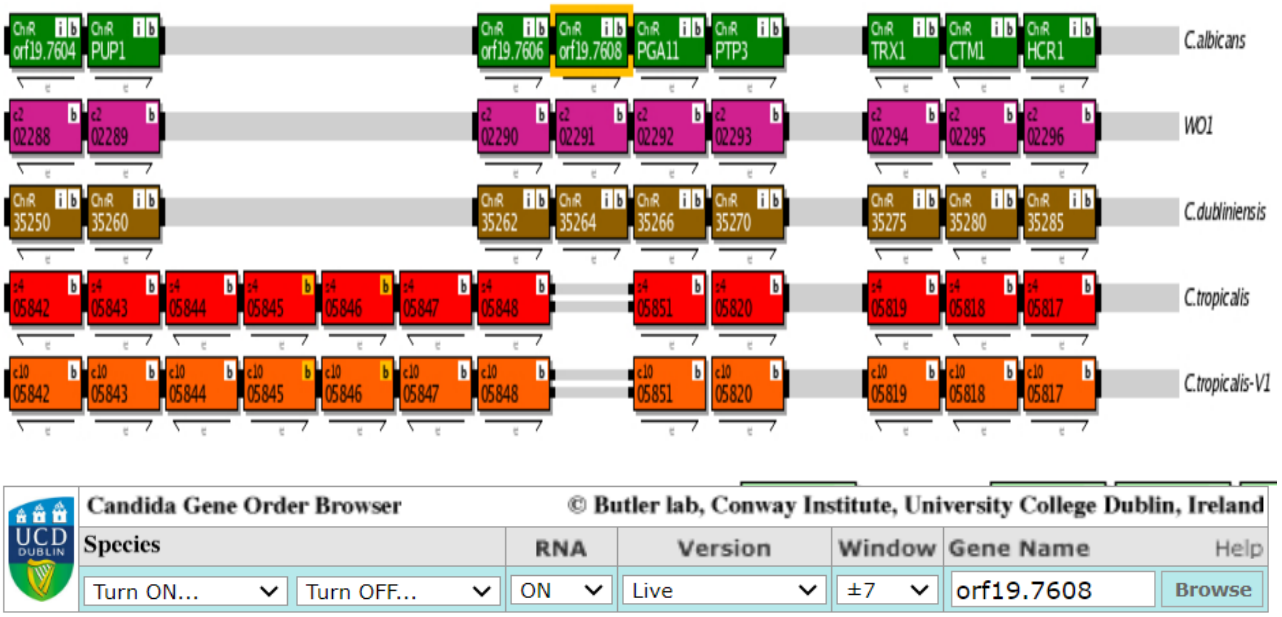


Figure 4: Ortholog synteny of *ORF19.7608* showing *ORF19.7608* in *Candida albicans* and *dubliniensis* with the locus in *tropicalis* unoccupied. Flanking genes *ORF19.7606*, *PGA11* and *PTP3* have verified orthologs. *35262* and *05848* are orthologs of *ORF19.7606*, *35266* and *05841* are orthologs of *PGA11*, *35270* and *05820* are orthologs of *PTP3* [75]. *WO1* same as *C. albicans*

conserved hypothetical protein [Candida dubliniensis CD36]

Sequence ID: XP_002422490.1 Length: 135 Number of Matches: 1

Range 1: 1 to 135 GenPept Graphics ▼ Next Match ▲ Pr

Score	Expect	Method	Identities	Positives	Gaps
171 bits(434)	2e-52	Compositional matrix adjust.	105/135(78%)	114/135(84%)	4/135(2%)

```
Query  1    MKTSVVLAAALTAMNTANAAVIPFFDGELSINYDAHSGWSFGFDKQPAQ--SGSPPATTT  58
            MKTSVVLAAALTAMNTANAAVIPFFDGELSINYDA+SGWSFGFDKQP+Q  SGS PATT
Sbjct  1    MKTSVVLAAALTAMNTANAAVIPFFDGELSINYDANSGWSFGFDKQPSQSGSGSAPATTN  60

Query  59   PGSGSGSTTTPATTADDGDDDEGSCDDSGSAPPA--NNGGSNGSSSGSGDYFSQWKQGLD  116
             G+ +GS TT A+T DDGDD+E SC DSGS PPA       + S SGSGDYFSQWKQGLD
Sbjct  61   TGNTNGSATTTASTGDDGDDEEDSCGDSGSTPPANSGGSNGSSSGSGSGDYFSQWKQGLD  120

Query  117  NLIQKGKTWFSGLFG  131
            +LIQKGK+WFSGLFG
Sbjct  121  DLIQKGKSWFSGLFG  135
```

Figure 5: NCBI protein BLAST alignment showing insertions and substitutions between the amino acid alignments of Orf19.7608 in *Candida albicans* and Cd36_35264 in *Candida dubliniensis* [36].

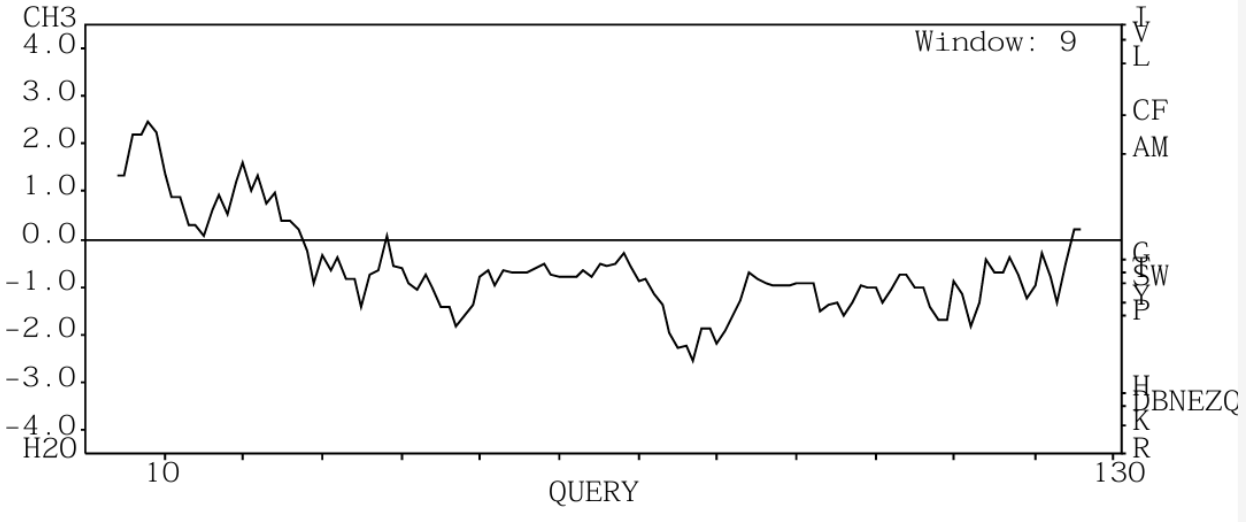
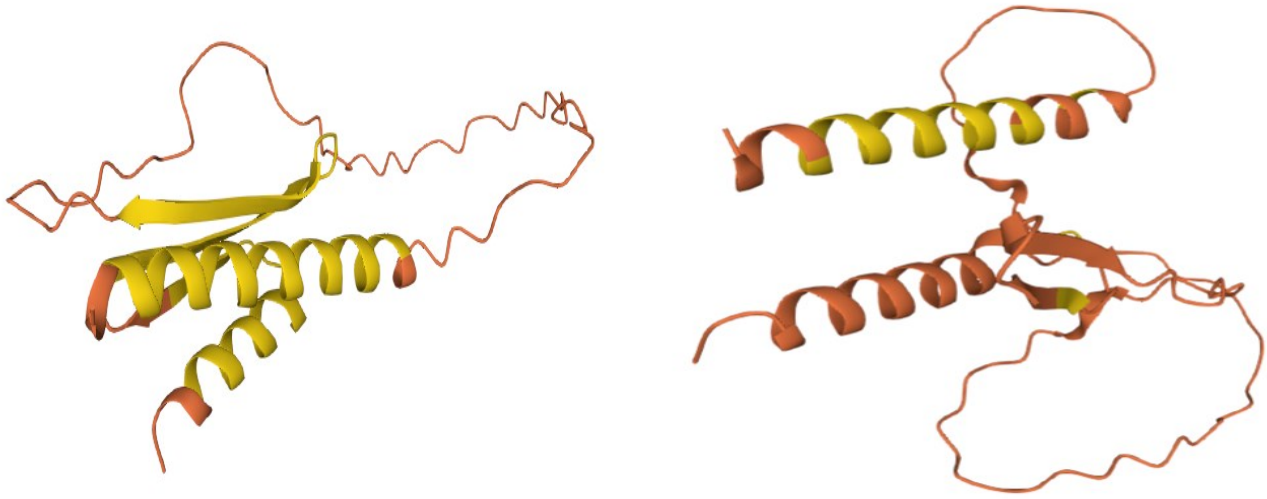


Figure 6: Kyte Dolittle plot of Orf19.7608 shows a hydrophobic N-terminus (N-signal peptide) within the first 20 amino acids [77,78]



Orf19.7608 in *Candida albicans* **Cd36_35264 in *Candida dubliniensis***

Low (70 > pLDDT > 50) for the α-helices and β- sheets

Very low (pLDDT < 50) for the loops

Figure 7: AlphaFold comparing the 3D structure of Orf19.7608 in *Candida albicans* with its ortholog Cd36_35264 in *Candida dubliniensis*. Per-residue confidence score (pLDDT) are between 0 and 100. Prediction reveals poorly defined 3D structures made up of mainly loops for both Orf19.7608 and Cd36_35264[79, 80].

3.2.1 Fluorescent microscopy identifies a pattern for Orf19.7608 under biofilm conditions

Subcellular localization through fluorescent microscopy is one way of understanding a protein's function as it visualizes the physiological processes a cell undergoes to annotate protein function from organellular compartmentalization [58]. We employed this to identify the organelle where Orf19.7608 is localized and to help suggest its possible functions. We grew the GFP tagged strain (N110) under biofilm conditions for 48 hours and planktonic conditions for 16 hours to check if growth conditions affect protein expression. As expected, N110 gave different protein expression patterns under these different conditions. In the vegetative cells, the cells were yeast-like green cells and the biofilm cells had distinct puncta

all over the cells as seen in fig 8b. Because of this protein expression pattern, we named our uncharacterized gene Punctate Pattern Protein 1 (*PPP1*).

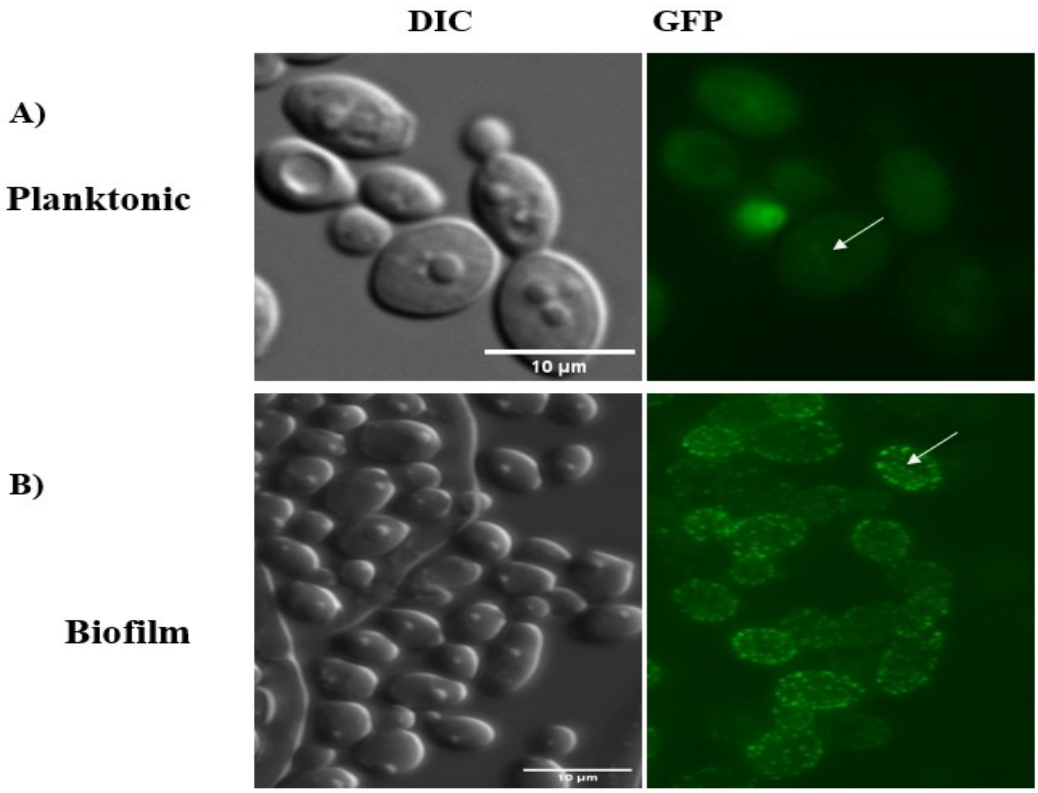


Figure 8: Fluorescent microscopy images of ORF19.7608-GFP

A) Planktonic conditions: overnight culture in YPD at 220rpm/30ºC

B) Biofilm condition: 48hrs growth in Spider at 75rpm/37ºC

Images were captured at 1000x. Scale bars represent 10µm.

3.2.2.1 Hyphae cells do not express the puncta distribution of Ppp1 even under filament inducing conditions

Hyphae cells in the biofilm do not express the punctate pattern of Ppp1. We grew N110 overnight in filament inducing conditions to see if the punctate pattern identified in the yeast cells can be visualized in the hyphae cells. Cells were grown for 16hrs at 37ºC in GlcNac,

RPMI, YPD + 10% serum and spider media. As seen in figure 9, none of the cells expressed the punctate pattern of Ppp1.

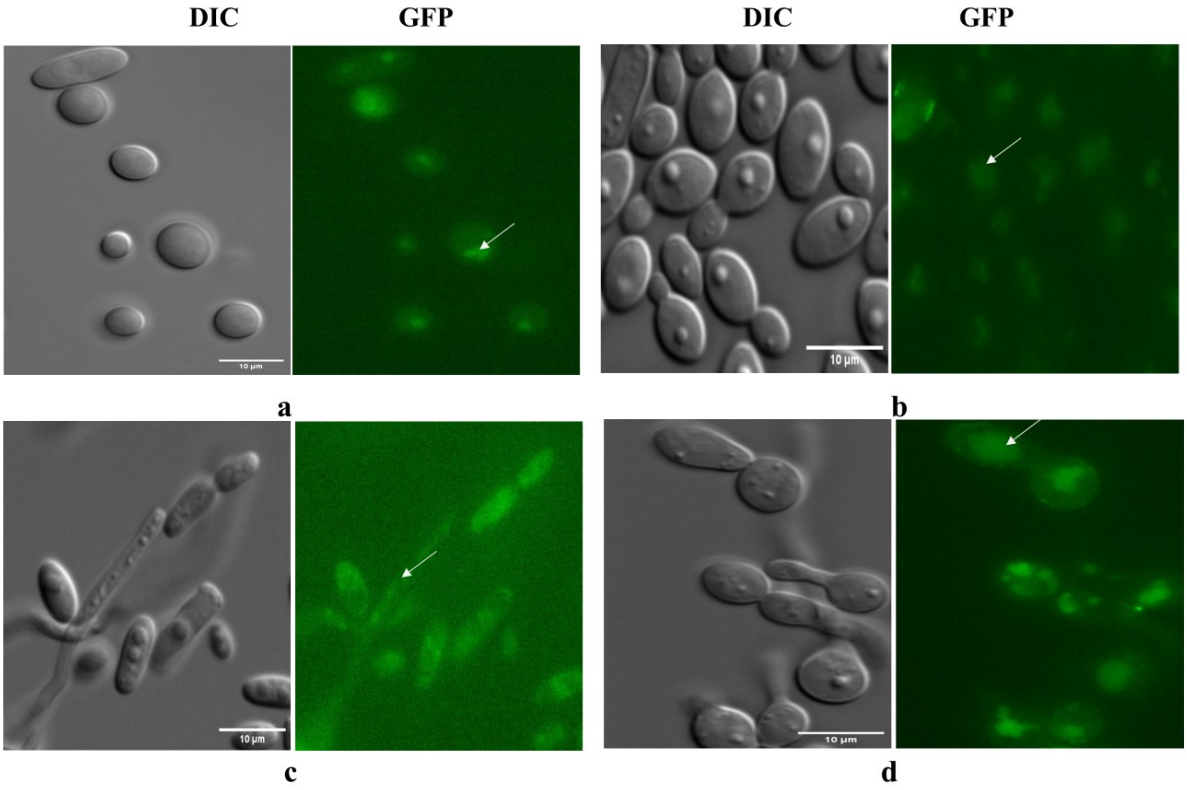


Figure 9: Fluorescent microscopy images of Ppp1-GFP grown in a) YPD +10% serum, b) Spider, c) GlcNac and d) RPMI for 16hrs at 220rpm/37ºC. Images were captured at 1000x. Scale bars represent 10µm.

3.2.2 Analysis of biofilm cells of Ppp1 shows this protein is highly expressed during ECM production

Although *PPP1* has been identified as a biofilm upregulated gene and its protein expression pattern determined, it's function in biofilm formation and in *Candida albicans* generally is still unknown. We approach this issue in the next series of experiments. First, we identify the stage of biofilm formation where Ppp1 is present by imaging N110 at different time points representative of the developmental stages of biofilm formation and extrapolating the stage of biofilm formation that Ppp1 is found in from the protein expression pattern at these time

points. To do this, N110 is grown as a biofilm for 48 hours in a 6 well plate. Images were taken at 1000x on LeicaDM6000 at time points representative of the phases of biofilm formation: 1.5 hours for cell adherence, 8 hours for the cell proliferation and hyphae formation, 24 hours for ECM production and 48 hours for ECM accumulation and dispersal of mature biofilm [73,74]. Ppp1 expression begins at the start of the biofilm cycle. However, during ECM production as seen in figure 10c and d, we see the puncta expression unique to the Ppp1 biofilm highly defined.

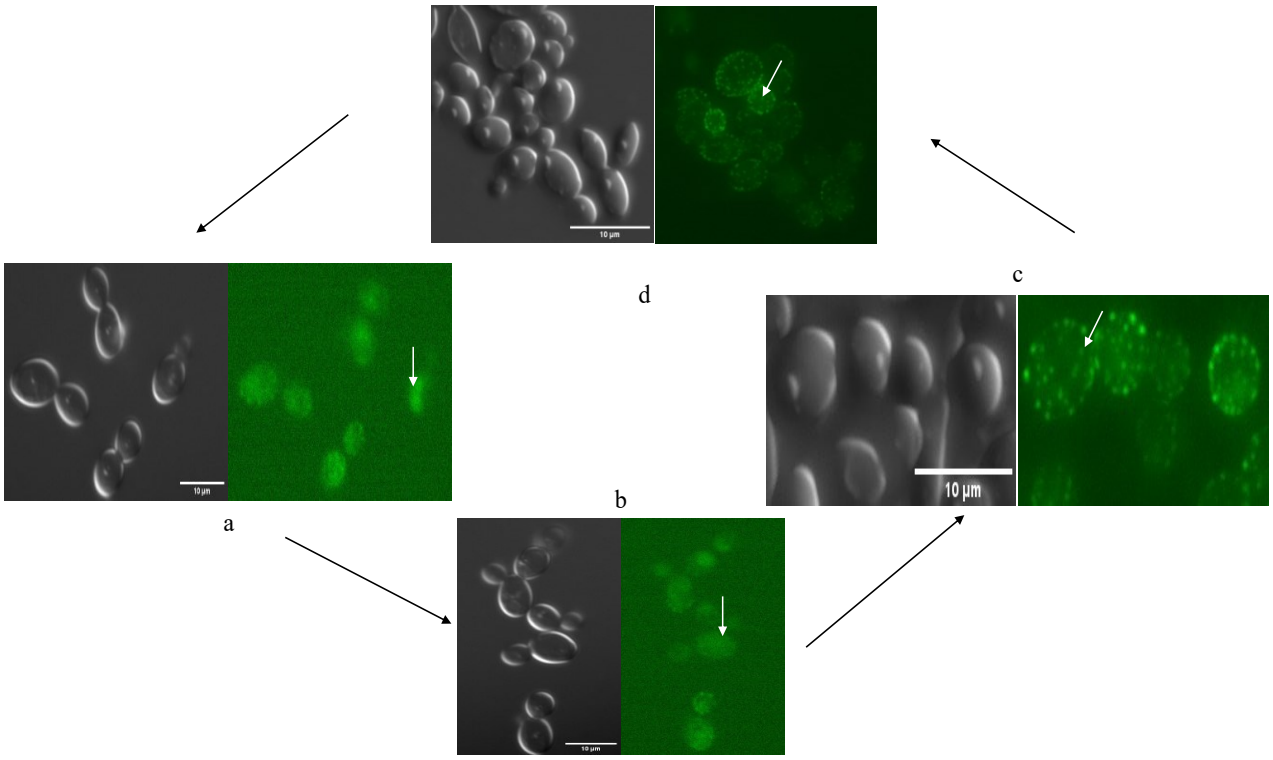


Figure 10: Ppp1 expressed at the different stages of biofilm formation

a) Cell adherence at 1.5 hours b) Cell proliferation and hyphae formation at 8 hours

c) Production of ECM at 24 hours

d) ECM accumulation and dispersal at 48 hours

Images were captured at 1000x. Scale bars represent 10µm.

3.2.3 sur7Δ/ Δ and arc35 Δ/ Δ do not show a different expression pattern for Ppp1

The next set of experiments were aimed at linking Ppp1 to a marker of a subcellular organelle with a similar protein expression pattern to Ppp1. *S. cerevisiae,* the model yeast, has most of its subcellular localizations annotated by fluorescent microscopy. It has three organelles with similar distribution patterns to that of Ppp1; the eisosome, the actin cytoskeleton and the Golgi, represented by Sur7, Arc35 and Sed5 respectively [81,82]. Sur7 is an eisosome defining protein localized in the plasma membrane involved in endocytosis [81]; Arc35 is localized in the actin cytoskeleton and is involved in hyphae formation [83] and Sed5 is a Golgi marker that is involved in protein transport [55].

Disruption of a subcellular compartment by deleting one of the domain proteins is one way of identifying the localization of a gene of interest. In this study, we separately made a homozygous mutant of *sur7* (N121) and *arc35* (N131) in N110, visualized the expression pattern of Ppp1, and compared them to the pattern in the WT N110. This was testing if deleting a component of a subcellular organelle may disrupt the organelle localization pattern visualized by microscopy. In our case, disrupting Arc35 or Sur7 did not affect the cell's protein expression pattern of Ppp1 as shown in Figure 11. The puncta look exactly like those in N110, suggesting that Ppp1 may not be localized in the eisosome or the actin cytoskeleton. However, a failed disruption of organelle patterning is not conclusive evidence that a protein is not localized in an organelle; these organelle structures are a result of many proteins, so a single disruption may not create any phenotypic difference. *SED5* is an essential gene so its deletion in the tagged Ppp1 strain could not be achieved [84].

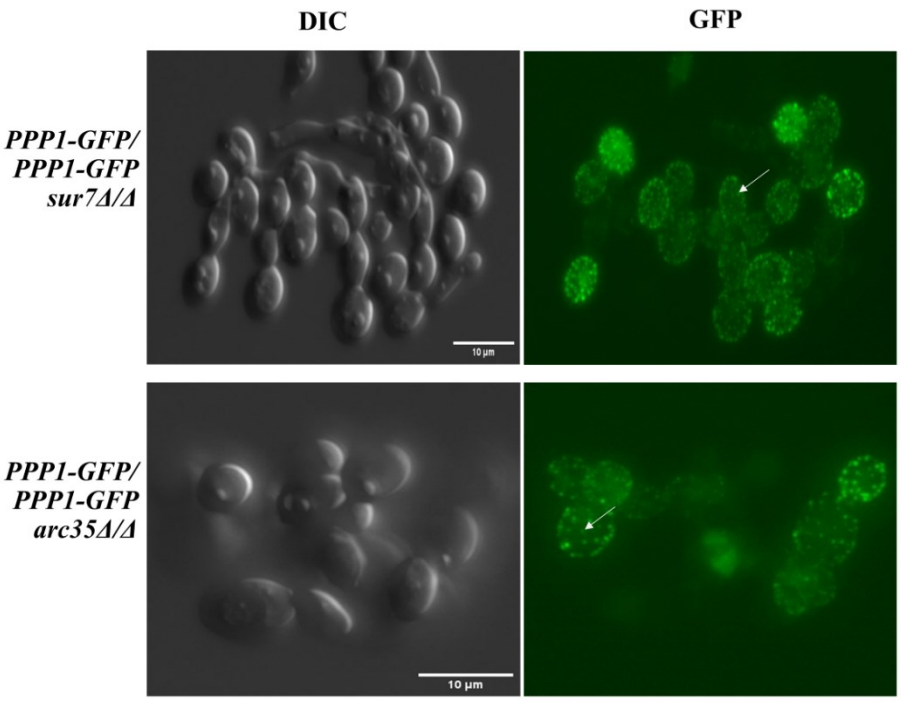


Figure 11: Fluorescent microscopy imaging of N121 (*sur7Δ/Δ PPP1-GFP*) and N131 (*arc35Δ/Δ PPP1-GFP*). *sur7Δ/Δ* and *arc35Δ/Δ* GFP signal was the same as *PPP1-GFP*. Cells were grown as biofilm and imaged DIC and GFP channels with LeicaDM6000. Images were captured at 1000x. Scale bars represent 10µm.

3.2.4 Colocalization analysis of Sur7, Arc35 and Sed5 show Ppp1 and Sed5 colocalize

Since deletion of a resident protein in an organelle may not disrupt the protein expression pattern of that subcellular compartment, we investigated co-localization to establish the relationship among the fusion proteins. Co-localization depends on the premise that repeated co-occurrence of two probes within a cell is an indicator of co-localization of the two proteins in the same compartment [60]. We tagged the organelle markers Sur7, Arc35 and Sed5 with Scarlet (an RFP) in the GFP tagged strain of Ppp1. Overlap merges the green and red signal to yield a yellow color which is easily distinguishable, or if the signals are distributed differently in the cell and do not overlap, the red and green signals remain distinct.

Strains N120, N130 and N140 were grown as biofilms and imaged with LeicaDM6000. Since two proteins have been tagged in each strain, we imaged the cells in FITC for the GFP signal and RFP for the Scarlet signal. Merging the images from the two fluorescent channels gave a quick visualization of the signal pattern and it was different for each strain as we can see in fig 11. In N120, where Scarlet is tagging Sur7, a component of the eisosome, as expected the Scarlet signals are on the cell's periphery while the GFP puncta are evenly distributed within the cell. Consequently, the merged channel shows most of GFP unmerged with RFP. In N130, where Scarlet is linked to the actin cytoskeleton marker Arc35, the signal distribution is more dispersed than for the eisosome marker but does not colocalize with the GFP signal from Ppp1. In contrast, in strain N140 where Scarlet is tagging the Golgi marker Sed5, we can see an almost perfect signal overlap between Scarlet and GFP.

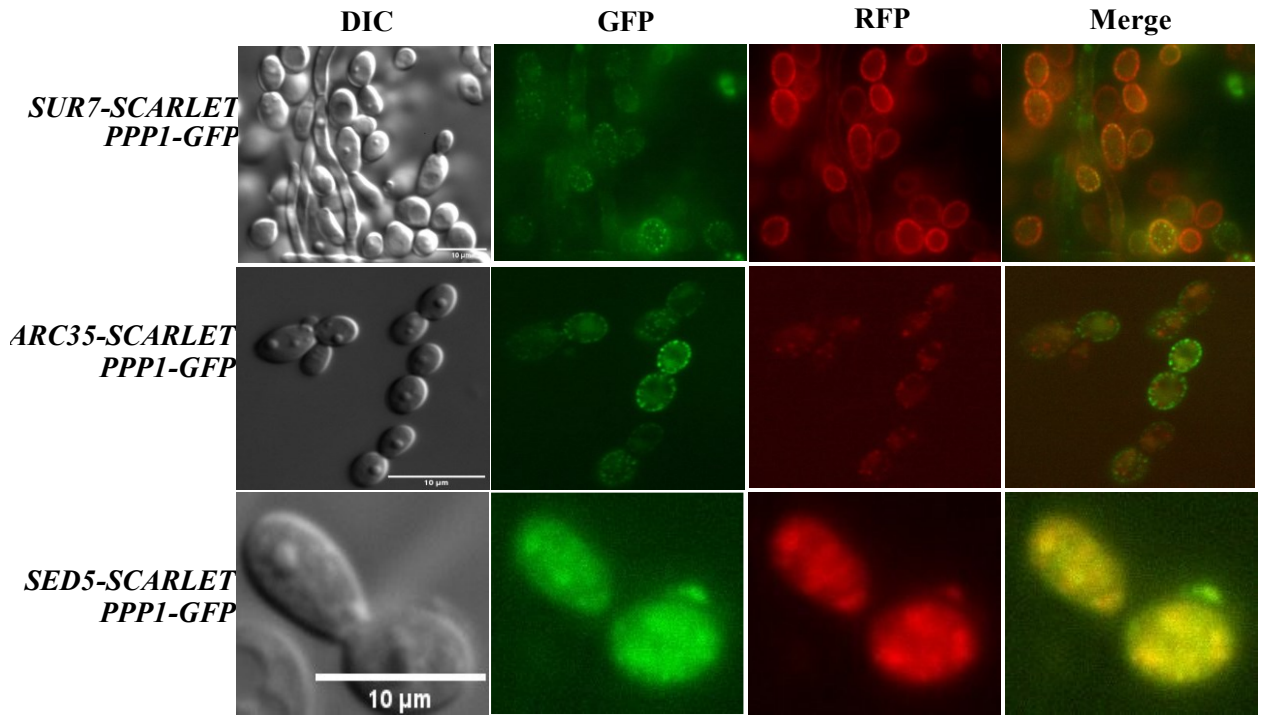


Fig 12: Colocalization of the GFP and Scarlet signal of Sur7, Arc35 and Sed5 suggesting that Ppp1 colocalizes with Sed5. Images were captured at 1000x. Scale bars represent 10µm.

Colocalization was quantified using the Mander's Colocalization Coefficient with a value of 1 showing perfect colocalization, 0 no colocalization and -1 inverse relationship. As seen in fig 13, for Sed5, we get a value of about 0.95 for GFP vs Sca and 0.9 for Sca vs GFP. Arc35 has a colocalization value of about 0.6 and Sur7 has a coefficient of 0.7 for GFP vs Sca and 0.6 for Sca vs GFP.

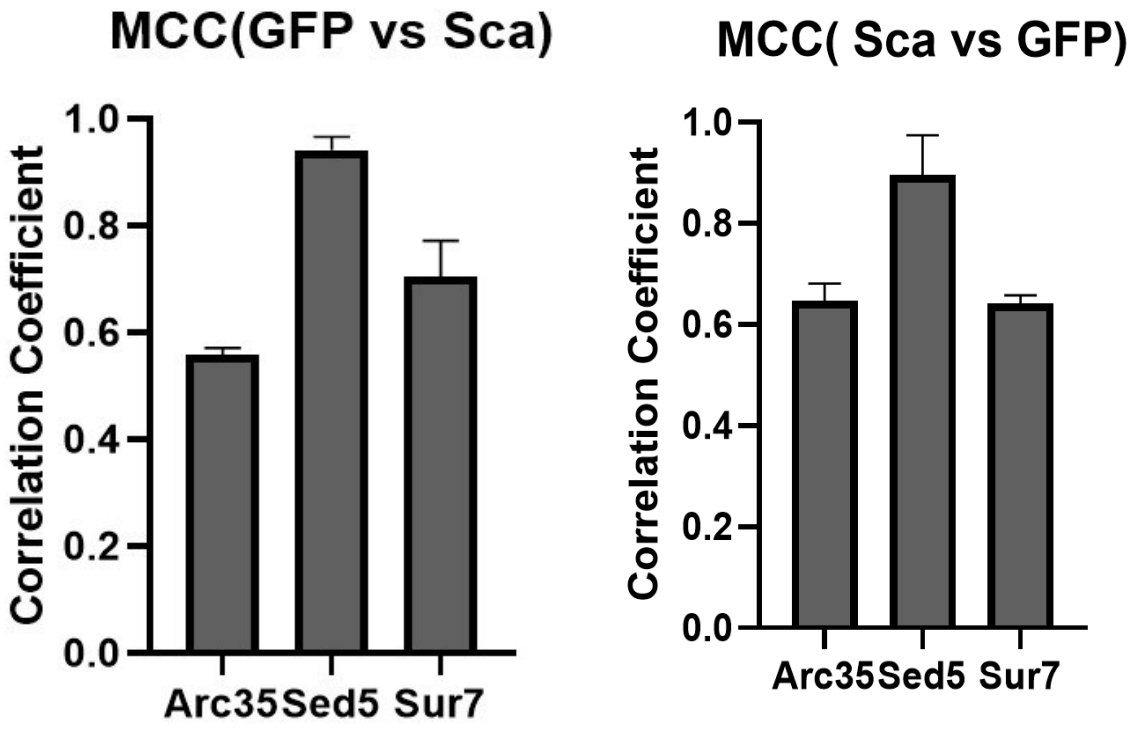


Fig 13: Mander's Colocalization Coefficient (MCC) measuring the colocalization between GFP in Ppp1 and Scarlet (Sca) in Arc35, Sed5 and Sur7. Error bars indicate SEM.

3.3 Sap2, a secretory protein has a different protein expression pattern from Ppp1

A signal peptide is a characteristic of many proteins in the secretory pathway and secretory proteins. We have examined the protein expression pattern of Ppp1 and compared it with components of the secretory pathway using imaging techniques and analysis. Our results

show that Ppp1 colocalizes with Sed5 in the Golgi. To validate that Ppp1 expression pattern is as a result of its localization in the Golgi and not just because it is passing through the secretory pathway, we also tagged a secretory aspartyl protease, Sap2 which has been characterized as a key player in the Sap family.

3.3.1 Sap2-Scarlet (N150) construct is properly tagged and encodes a functional protein

To confirm that the construct was properly inserted, we performed three colony PCR reactions which we have named S1 (FW Ext *SAP2*, RV Int *SCA* = 979bp), S2 (FW Int *URA3,* RV Ext *SAP2* = 1274bp) and P (FW Ext *SAP2*, RV Ext *SAP2* = 2849bp). In Fig 14A, colony PCR shows our strain had the correct band size for S1, S2 and P. There was no band for the WT in S1 and S2 and in P the size was between 850bp and 1000bp using the 1kb plus ruler (actual size is 877bp).

With the construct confirmed by colony PCR, we then did nanopore sequencing of our whole construct (Sap2 – Ura). This sequence was selected to check if the repair DNA insertion introduced some mutations that could affect the gene translation of Sap2. We sent an amplicon of about 3.3kb containing all of *SAP2, SCARLET* and *URA3* (*SAP2* = 1194bp, linker = 18bp, *SCARLET* = 699bp, *URA3* = 813bp, CIp10-mScarlet-IDT and pFA-GFP-Ura3 plasmids = 593bp) Sequencing data as shown in figure 14C confirms that the construct was properly inserted but had some mutations. Between Sap2 and Scarlet, there were 7 point mutations (4 in Sap2 and 3 in Scarlet). Most of the mutations in this amplicon fall between 2kb and 3.3kb in the Ura3 region. Ura3 was placed after Scarlet and only serves as a selection for this strain so mutations in this region are not important to this study. Upon sequencing of Sap2 in the WT strain (data not shown), we discovered that only 2 mutations were as a result of the strain construct, the other two were due to differences in the genomic sequence and plasmidsaurus sequence data. None of these mutations introduced a stop codon so we expected our protein to function properly.

To confirm that the Scarlet fusion does not interfere with Sap2 function, we conducted a YCB-BSA assay. SAP proteins are Secreted Aspartyl Proteinases involved in pathogenicity, and among this family of proteins, Sap2 is needed to process substrates containing proteins

as the sole nitrogen source [41]. When cells are grown in YCB-BSA, if Sap2 is non-functional, the substrate is not degraded, and no zone of clearance is seen on the YCB-BSA plates. For negative control, we included a *sap2Δ/Δ* (N151) from Hube et al. 1997. As seen in fig 14B, there was a large zone of clearance for SN148 and N150 and a minimal one for N151. The negative control confirms that our assay works, and the wild type zone of clearance from N150 confirms our fusion protein is functional and is properly secreted. Therefore, the pattern of RFP signal from strain N150 should represent that of a fluorescently tagged protein passing through the secretory pathway.

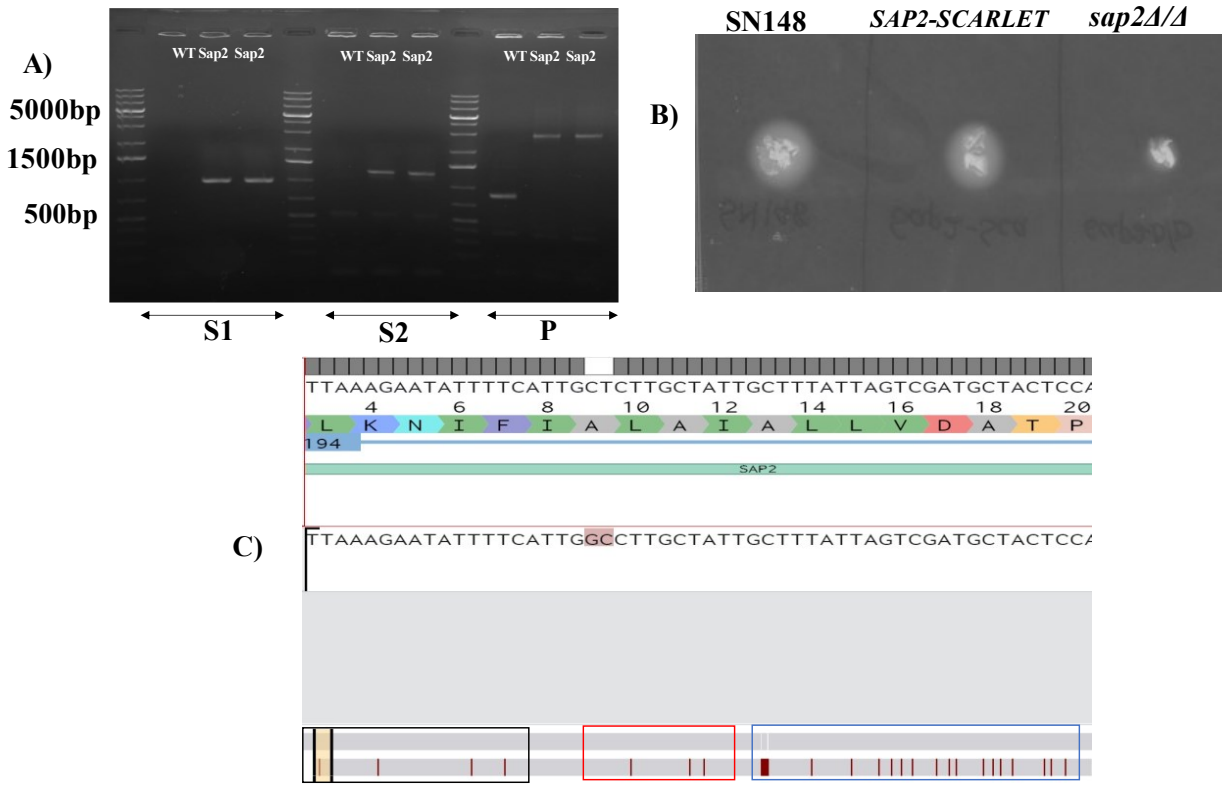


Figure 14: Diagnostic testing of Sap2-Scarlet

A) Colony PCR for Sap2-Scarlet using WT(SN148) as control. S1= FWSap2 - RVSca, S2= FWUra3 - RVSap2 and P= FWSap2- RVSap2. 1kb plus ladder measures band size

B) YCB-BSA assay for Sap2-Scarlet with SN148 as positive control and sap2Δ/Δ as negative control. Zone of clearance shows degradation of BSA.

C) Sap2-Scarlet-Ura3 sequence by plasmidsaurus. Black box enclose Sap2 mutations, red box enclose Scarlet mutations and blue box enclose Ura3 mutations.

3.3.2 Sap2 does not have a punctate distribution similar to that of Ppp1

Upon confirmation that N150 was properly constructed and is functional, we visualized the tagged protein. We grew a colony of N150 as a biofilm and prepared slides as described in the materials and methods. Comparison of the distribution of Sap2 (as seen in fig 15) with Ppp1 (as seen in fig 8) shows that Sap2 and Ppp1 do not have the same distribution pattern. Scarlet signal from Sap2 does not form puncta but is instead spread throughout the cell. This shows that even though both have N-signal peptides, the punctate expression pattern in Ppp1 is characteristic of Golgi resident proteins and not secreted proteins.

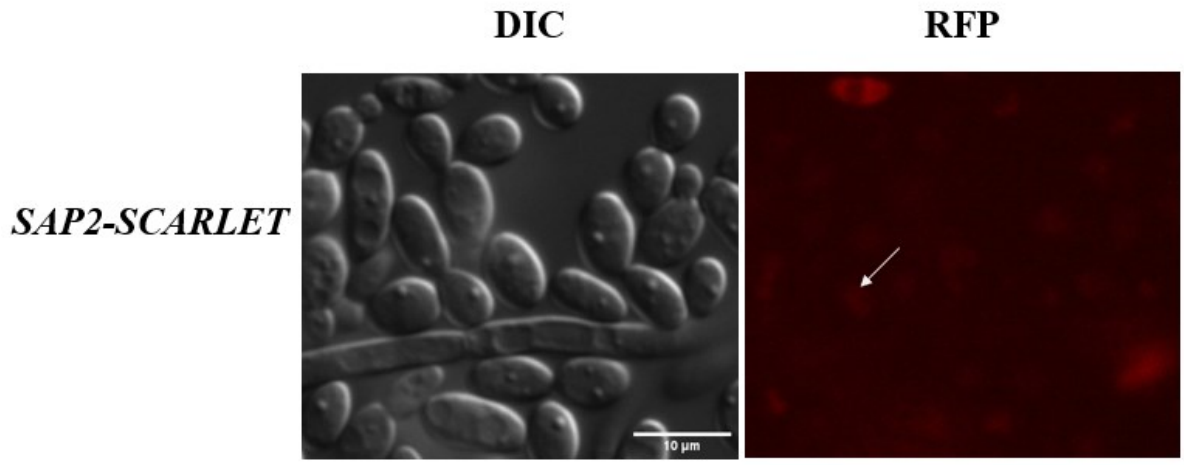


Figure 15*:* Fluorescent microscopy imaging of N150 for Sap2-Scarlet. Sap2-Scarlet was grown under same conditions as Ppp1 biofilm (48hrs growth in 6 well plate at 37^{o}C, 75rpm). Sap2 does not form puncta like Ppp1-GFP. Cells were imaged in DIC and RFP channels with LeicaDM6000. Images were captured at 1000x. Scale bars represent 10µm.

3.4.1 Cellular stress and antifungal assays do not show an altered growth for Ppp1

Candida albicans has an environmental response pathway made up of genes that respond to factors like osmotic and oxidative stress [85]. Osmotic and oxidative stress play an important role in fungal virulence with osmotic stress increasing the tonicity of the cell thereby causing water loss [28] and cell size reduction and oxidative stress causing the release of reactive

oxygen species (ROS) [86]. We tested these stress factors at 30ºC and 37ºC on strains N101 (deletion mutant) and N100 (complemented strain) to see if loss of *PPP1* is affecting response to stress inducers and/or any of the classes of antifungals. These experiments were done using overnight YPD uridine culture with a starting concentration of 10^7 cells/ml. Spot assays of YPD uridine agar supplemented with the appropriate stress-inducing factors were repeated for consistency in data and cell concentrations from 10^6 to 10^4 were analyzed. For osmotic stress, sodium chloride (NaCl), calcium chloride ($CaCl_2$) and glycerol were used, while hydrogen peroxide (H_2O_2) was used for oxidative stress. To identify distinct phenotypes in the stress assays, concentrations where cells can grow at varying rates and morphologies were optimized from literature review: NaCl (500nM), $CaCl_2$ (400mM), H_2O_2 (3mM) and glycerol (100nM) [67]. No altered growth was identified in any of the stress indicators as seen in figure 16A suggesting that *PPP1* does not have any effect on osmotolerance or oxidative stress.

One of the ways in which biofilms persist in the hospital environment is by resistance to antifungals. Due to the structure of the biofilm cells and the genes induced, they develop mechanisms to evade antifungal therapy. In this study, we also monitored the effects of the four classes of antifungals on strains deleted for *PPP1.* Testing the four classes of antifungals on our strains allows the effect of the drugs on different subcellular compartments to be observed [9]. Caspofungin, an echinocandin, inhibits 1-3 β-D-glucan affecting the cell wall synthesis; amphotericin B, a polyene, inhibits ergosterol synthesis needed for cell membrane formation; fluconazole, an azole, inhibits cytochrome P450-dependent 14α-lanosterol demethylase inhibiting ergosterol biosynthesis, and 5- flucytosine (5-FC), a pyrimidine analog, disrupts DNA and RNA synthesis [15,16,17]. Like the stress assays, Minimum Inhibitory Concentrations (MICs) were obtained from literature ensuring that phenotypic changes can be observed. Concentration of antifungals were as follows: flucanozole (1ug/ml), 5-Flucytosine (5-FC) (0.5ug/ml), amphotericin B (0.25ug/ml), and caspofungin (0.75ug/ml) [68, 69, 70]. Strains were also grown overnight, diluted to OD_{600}=1, and spotted in dilutions from 10^6 to 10^3. As with the stress assays, we did not also observe any altered growth in the mutant and complemented strain in comparison with the wildtype (see Figure 16B).

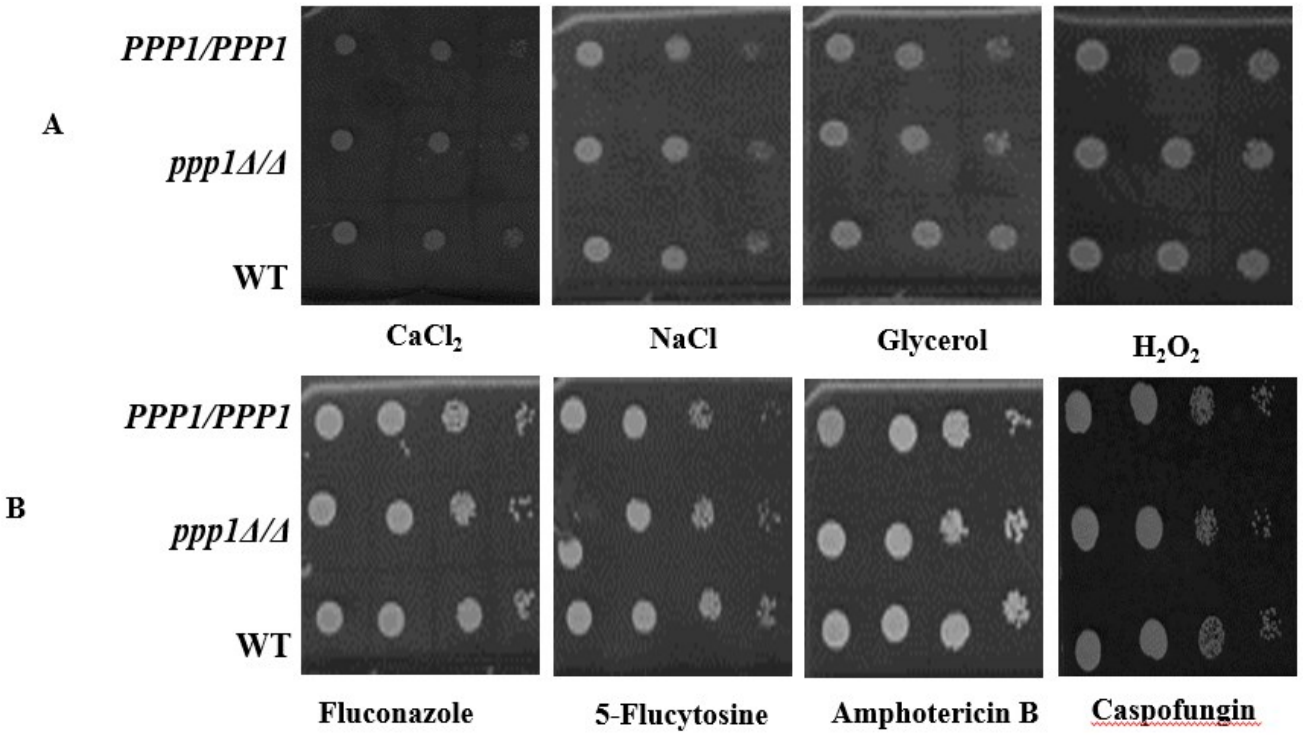


Figure 16: Genotoxic and cellular stress assay. A 1:10 serial dilution of overnight cultures grown in yeast growth conditions were spotted (highest cell concentrations being 1 x 10^6 cells/ml and lowest being 1 x 10^4 cells/ml for stress and 1 x 10^3 cells/ml for antifungals) onto YPD agar plates containing different chemicals and were incubated at 30^{o}C for 2 days except for caspofungin (5 days). Experiments were repeated three times for each sample for consistent results. Results were the same at 30^{o}C and 37^{o}C.

(A) Strains were subjected to osmotic stress of CaCl2 (400mM), NaCl (500nM), glycerol (100nM) and oxidative stress of H_2O_2 (3mM)

(B) Determination of resistance to different antifungal drugs: flucanozole (1ug/ml), 5-Flucytosine (0.5ug/ml), amphotericin B (0.25ug/ml), and caspofungin (0.75ug/ml).

3.4.2 Ppp1 does not influence invasive response

Invasiveness is an adaptation of *Candida albicans* to host tissues. Since filamentation, one of the steps in biofilm formation, aids invasiveness, we tested our mutant and complemented strains for invasiveness using a plate washing assay [67]. Cells were grown overnight in YPD uridine, diluted to OD_{600}=1 and 3ul plated on YPD and spider media and incubated at 30ºC

and 37ºC for 72hours. There were no colonies in the plates after cells were washed to remove non-adherent cells.

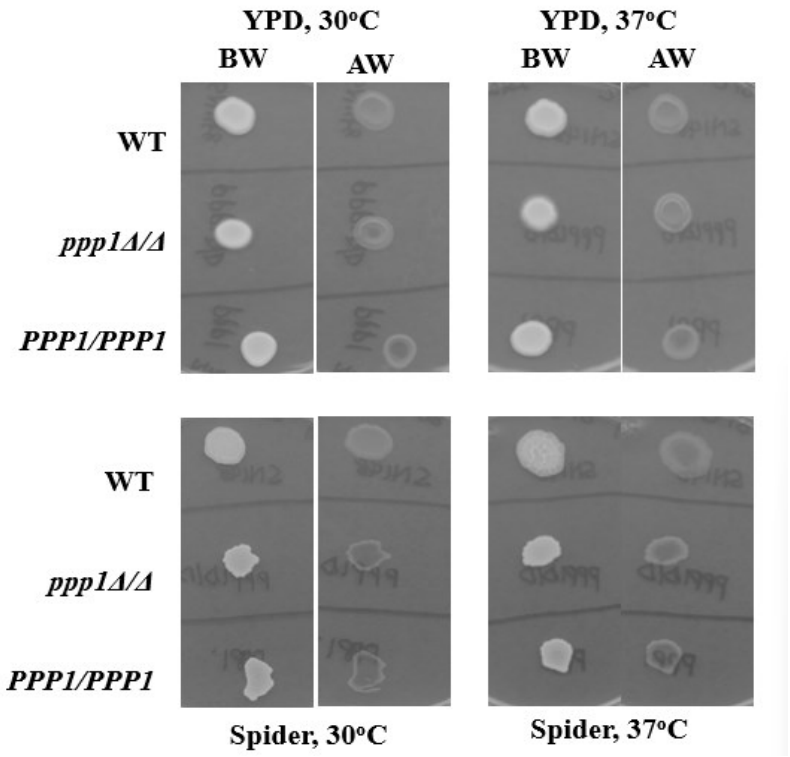


Figure 17: Invasiveness of *PPP1* in *Candida albicans*. Overnight cultures grown in YPD uridine were diluted to OD_{600} of 1.0. 3µl of samples were spotted on YPD and Spider media plates and incubated at 30˚C and 37˚C for 72 hours. Resulting colonies were washed gently by spritzing water on plates to remove the non-adherent cells, and invasiveness of the samples was observed. BW = before wash, AW= after wash

3.4.3 Crystal violet assay shows Ppp1 does not have a significant effect on biofilm formation

Since Ppp1 is highly expressed during biofilm formation, we measured biofilm biomass in strains N100 and N101. Strains were grown in different biofilm inducing media to identify potential differences in biofilm biomass based on chemical composition of these media. Crystal violet assay measures how much of the dye is taken up by the biofilm, providing a measure of the biofilm biomass. In buffered Spider medium (pH adjusted to 7.2), deletion of

PPP1 did not cause any difference in biofilm production as biomass of all strains are similar. In unbuffered Spider media biofilm production was also low as for the buffered media. The mutant produced a little more biofilm than the wildtype but less than the complement strain. A significantly greater amount of biofilm was formed when cells were grown in RPMI media. In RPMI, biofilm formation in the mutant strain was a bit lower than others and the complement produced almost the same amount of biofilm as the WT. These data show that there is no significant change in biofilm formation when *PPP1* is deleted.

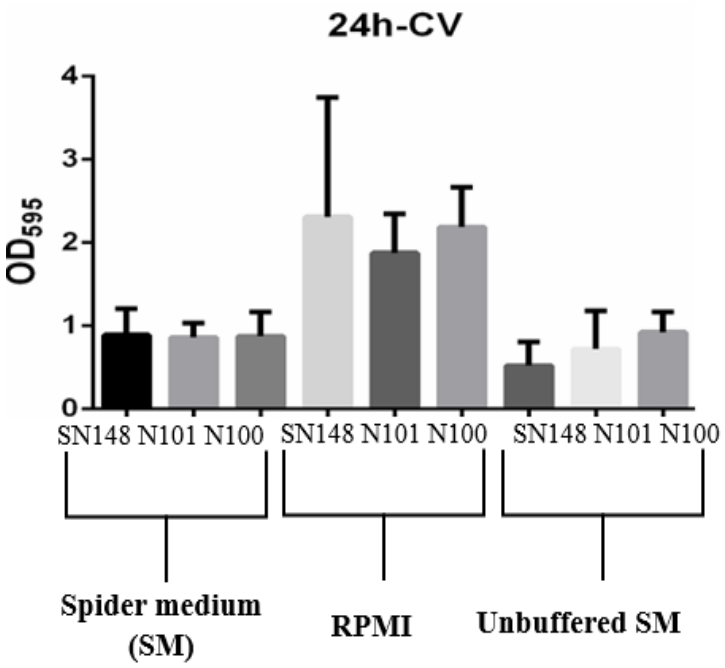


Figure 18: Crystal violet assay measuring the biofilm biomass of *ppp1Δ/Δ* (N101) and *PPP1 + pV1093* (N100) comparing it to the WT. Experiments were done in triplicates and repeated for consistence. Error bars represent standard deviation.

3.4.4 Confocal microscopy shows minor effects in the biofilm formed by the complemented strain

The mutant and complement strain were tested by the Mitchell lab who used confocal microscopy to analyze the biofilm structure. Side and apical views of the strains were imaged, and these were compared against the wildtype. As seen in figure 19, there was no

difference in biofilm formation in the mutant and complement strain compared to the WT in the side view, but the complement strain had somewhat more biofilm in the apical view compared to the mutant and wildtype. When 10% FBS was added (fig 20), there was a substantial increase in biofilm formation, but the complement strain had shorter hyphae and less biofilm compared to the mutant and WT. In general, deletion of *PPP1* does not significantly impact biofilm formation. However, reinsertion of *PPP1*at it's native locus does not exactly return the strain back to the WT, as there are subtle differences in biofilm formation.

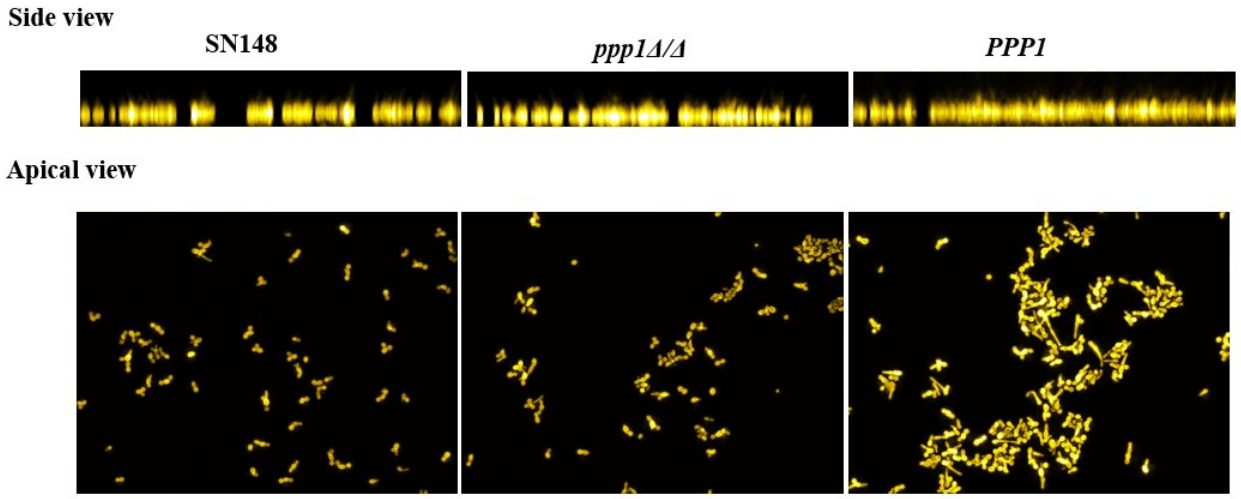


Figure 19: Side and apical view confocal microscopy images of SN148 (WT), *ppp1Δ/Δ* and *PPP1+ pV1093* (complement strain) grown in RPMI for 24 hours at 37°C. Imaging was done in duplicates.

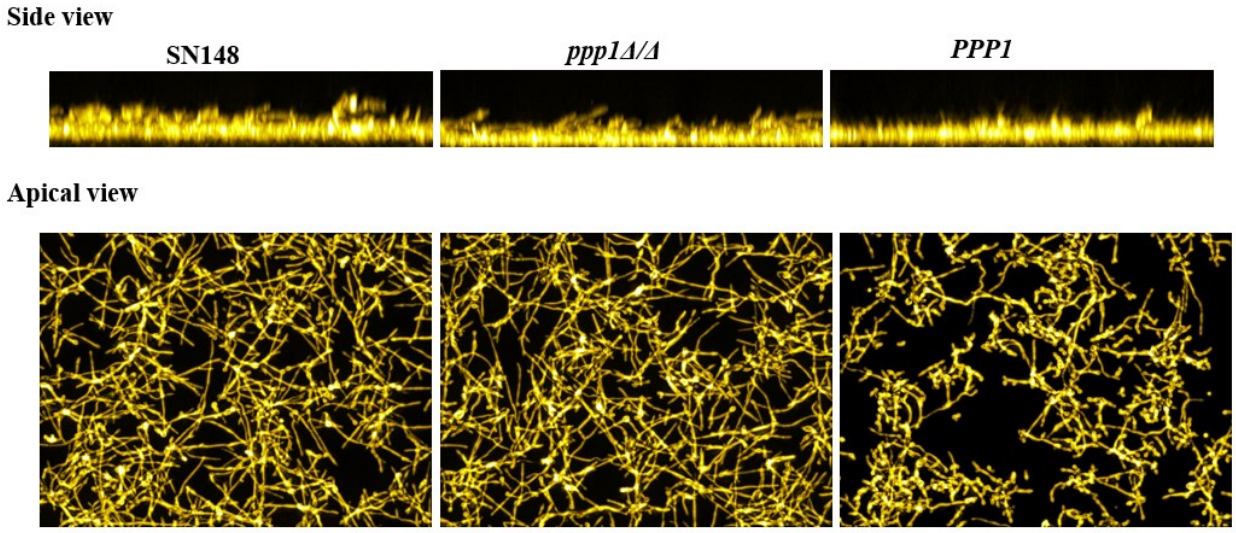


Figure 20: Side and apical view confocal microscopy images of SN148 (WT), *ppp1Δ/Δ* and *PPP1+ pV1093* (complement strain) grown in RPMI + 10% FBS for 24 hours at 37°C. Imaging was done in duplicates.

3.5 Orf19.6274 and Orf19.4654 proteins with similar characteristics to Ppp1 have a different distribution pattern to that of Ppp1

Previous experiments identified Ppp1 as being a Golgi resident protein, but do not identify an altered growth caused by its deletion. We investigated other proteins with characteristics similar to Ppp1 to see if they perhaps belong to a complex and work in tandem to influence biofilm formation. To explore this genetic network, we looked at Orf19.6274 and Orf19.4654, proteins with similar characteristics to Ppp1. Orf19.6274 (Pbr1) and Orf19.4654 are small proteins with N-signal peptides highly upregulated in *Candida albicans* biofilms. Pbr1 is a protein of 175aa with orthologs in *dubliniensis* and *tropicalis* while Orf19.4654 is a protein of 109aa with an ortholog only in *dubliniensis.* Both have no annotated subcellular location and have similar log2 differential expression in biofilms: 8.79 for Pbr1 and 7.2 for Orf19.4654. Pbr1 is located on chromosome 1 while Orf19.4654 is located on chromosome 4

We tagged these proteins with Scarlet and visualized their expression patterns. As seen in fig 21, neither of these proteins have the same pattern as Ppp1, making Ppp1's uniqueness all the more interesting. Pbr1 and Orf19.4654 have different localization patterns under biofilm

conditions with Orf19.4654 having a nuclear like localization similar to Smr1[82] and Pbr1 having a localization similar to Sap2 found in the extracellular matrix.

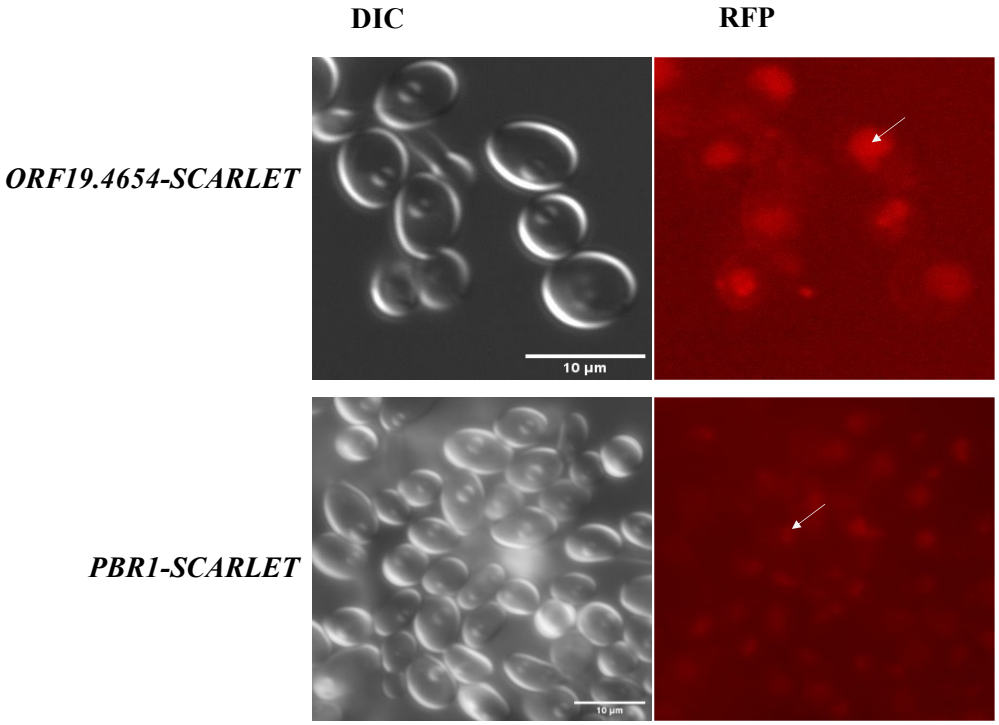


Figure 21: Fluorescent microscopy images of PBR1-Scarlet and Orf19.4654-Scarlet. Cells were grown as biofilm and imaged DIC and RFP channels with LeicaDM6000. Images were captured at 1000x. Scale bars represent 10µm.

Chapter 4: Discussion

Candida spp. are the third most frequent cause of nosocomial infections and have a major role in biofilms [87]. Biofilms are a virulence factor of medical importance due to their prevalence in nosocomial infections especially among patients with implanted devices [33]. They are also the leading cause of infection in immunocompromised patients and can affect healthy individuals. Due to the nature of biofilm cells, they are resistant to most antifungals. They can cause systemic candidiasis leading to 30-50% mortality rate in infected patients [33]. Decades of studies have gone into understanding how *Candida albicans* biofilms are formed and defining the network of genes and transcription factors (TFs) involved in this process. Nobile et al. 2012 identified more than 1000 genes regulated by 6 key TFs (Bcr1, Tec1, Efg1, Ndt80, Rob1, and Brg1) involved in biofilm formation [31]. Some of the genes identified like *HWP1, ALS1* have been extensively studied. *ORF19.7608* was one of the most highly upregulated genes and due to its uncharacterized nature, we decided to study it in-depth to annotate its function and investigate the localization of the encoded protein.

Orf19.7608 is a small protein of 131aa with an N-terminal signal peptide between aa 1 and 19. Protein sequence alignments of Orf19.7608 reveals an ortholog in only its sister species *Candida dubliniensis* with a percentage identity of 78%. Orf19.7608 belongs to the class of genes called Albicans Specific Genes (ASGs) [88]. This family of 65 genes are found in only *Candida albicans* and *dubliniensis* [88]. As seen in figure 4, looking at the synteny of Orf19.7608 in relation to its sister species *dubliniensis* and *tropicalis,* the next related species of this genus, Orf19.7608 seems to have dropped into the genome with none of the neighboring genes highly expressed in *Candida albicans* biofilms [31]. Although Orf19.7608 belongs to a family of genes with regards to function, it's locus is not in a cluster in the genome. The crystalline structure of this protein still remains unknown as the predicted structure from AlphaFold gives a poor predicted structure made of mostly loops, a non-preferred 3D structure for a protein. AlphaFold does not distinguish between instrinsically disordered proteins and poorly ordered predictions, so no insight was given on the structure of Orf19.7608. Uniprot, InterPro and other protein databases do not have any annotation for

its biological function and experimentally backed evidence for its localization. These preliminary analyses provide some information for Orf19.7608.

Fluorescent microscopy can be used for identifying the organelle a protein is localized in with the aim of understanding its function. It can define the subcellular resolution of proteins enhancing the understanding of cell physiology [58]. This technique has been used in various studies on subcellular localization in yeast [89,90]. When Orf19.7608 was tagged with GFP and visualized, as seen in figure 8, a punctate pattern distributed all over the cell surface was seen only in biofilm cells hence the name Punctate Pattern Protein 1 (Ppp1) for Orf19.7608. This difference in protein expression follows the pattern we expect from the differential analysis of biofilm genes [31]. This punctate pattern was specific to the yeast cells and even when cells were grown in biofilm inducing conditions as seen in figure 9, the punctate pattern is absent. In figure 10, cells imaged at time points representative of the different stages of biofilm formation show these unique puncta expression highly defined in the mature phase where the extracellular matrix (ECM) has formed. This suggest that Ppp1 is a mature biofilm protein with a possible role in the formation of the ECM.

The puncta expression pattern of Ppp1 in yeast cells is consistent with specific organelles in the model yeast *S. cerevisiae* namely the eisosome of the plasma membrane, the actin cytoskeleton, and the Golgi. All these organelles have potential roles in biofilm formation with the plasma membrane helping in endocytosis so *Candida* can get nourishment from the host [81]; the actin cytoskeleton involved in hyphae formation for cells to invade host tissues and cause infection [80] and the Golgi for protein packaging and transport [55]. The Golgi also helps in glycosylation for the synthesis of glycoproteins [50] which forms 55% of the extracellular matrix [64]. Markers of these components were used for analysis of the subcellular localization of Ppp1: Sur7 for the eisosome, Arc35 for the actin cytoskeleton and Sed5 for the Golgi. Gene disruption and colocalization were used to identify the organelle Ppp1 resides in. From analysis of knockouts and tags of *SUR7, ARC35* and *SED5* in *PPP1-GFP*, Ppp1 seems to appear in the same subcellular location as Sed5. Figure 11 shows the puncta distribution when Sur7 and Arc35 were knocked out. In *sur7Δ/Δ PPP1-GFP*, the

punctate pattern of Ppp1 remains the same and merging the GFP and Scarlet signals of *SUR7-SCARLET PPP1-GFP* shows the scarlet pattern is localized in the periphery of the cell unlike Ppp1. Further analysis shows that these two patterns do not colocalize, indicating Ppp1 is not resident of the plasma membrane. The same is true in *arc35Δ/Δ PPP1-GFP* and *ARC35-SCARLET PPP1-GFP* but not Sed5. *SED5* is an essential gene so we were unable to construct *sed5Δ/Δ PPP1-GFP.* In *SED5-SCARLET PPP1-GFP*, the GFP and Scarlet signals overlap forming a yellow color. As seen in fig 12, merging Sed5-Scarlet and Ppp1-GFP indicating colocalization between the two signals. This was further quantified using the Mander's Colocalization Coefficient (MCC) in figure 13 which measured a value of about 0.95 for the overlap of GFP with Scarlet and 0.9 for the overlap of Scarlet with GFP. This coefficient values depict perfect colocalization between Ppp1 and Sed5 suggesting that Ppp1 is a Golgi resident protein [91]. The Golgi is involved in protein glycosylation, sorting and trafficking to cellular destinations [50]. Glycosylation helps in proper folding of protein which makes them functional and aids their proper localization. This glycosylation process adds sugar residues to the proteins, critical for the synthesis of the glycoproteins. Glycosylation also protect transported proteins from rapid proteolytic degradation [92]. Glycoproteins make up more than half of the ECM of Candida biofilms [87]. The ECM acts as a capsule conferring antifungal and immunological resistance to biofilm cells [87]. These findings suggest Ppp1 is part of the Golgi and is highly expressed in the maturation stage of biofilm formation thus annotating its subcellular localization and candidate biological function.

An N- signal peptide is a characteristic of both secretory proteins and residents of the secretory pathway. We tagged a secretory protein Sap2 that has an N-signal peptide to validate our findings on Ppp1's localization. The choice of Sap2 gave us the opportunity to rely on extensive studies already done and conclusions made on this secreted protein [41,93]. After confirming that tagging this gene does not have an effect on its function using the YCB-BSA assay to evaluate growth on nitrogen as the sole protein source, we visualized the protein pattern and compared to Ppp1. From fig 15, we see Sap2 forms neither puncta nor the evenly distributed pattern of Ppp1, instead its signal is localized in the center of the cell. The

non-unique protein pattern of Sap2 comes as no surprise as this protein has been extensively studied decades ago yet there is no publication of its protein expression pattern using fluorescent microscopy. This difference in protein distribution supports our suggestion that Ppp1 is a Golgi resident protein.

Annotation by phenotypic analysis of mutant strains is a powerful, standard technique for functional characterization of proteins [94]. This can be used to identify drug targets and study the effect of different stress responses on the cell [95]. In this study, we constructed a mutant and complemented strain of Ppp1. The complemented strain, *PPP1* reinserted at its native locus, helps in determining if phenotypical differences are from Ppp1 or other genetic factors such as gene duplication [96]. With the subcellular localization and biological function of Ppp1 annotated, we investigated cells lacking *PPP1* for responses to antifungals and environmental factors such as oxidative and osmotic stress (see figure 16A). When cells were treated with H_20_2 for oxidative stress, we did not observe an altered growth in the mutant and complement strain compared to the WT. Under osmotic conditions of NaCl, $CaCl_2$, and glycerol we did not observe any defect in the cell morphology of *ppp1 Δ/Δ* and *PPP1*. Stress factors affect cell wall composition and *Candida* biofilm virulence [28]. The absence of altered growth when cells are treated with stress factors suggests that Ppp1 may not involved in stress responses in *Candida albicans*.

Drug resistance is a major area of medical research due to the high mortality rate associated with *Candida* infections. Screening Ppp1 with fluconazole, 5-flucytosine, caspofungin and amphotericin B representing the four classes of antifungals will study the impact of disrupting various biological processes on the cell integrity and viability of *Candida albicans*. In figure 16B, this assay shows that these antifungals did not affect the morphology and/or viability of Ppp1. We also measured biofilm formation when Ppp1 is deleted using the crystal violet assay. As seen in figure 18, there is no significant difference in biofilm formation when Ppp1 is knocked out. Since invasiveness is an important step in pathogenicity, we measured how invasive the cells are in the absence of Ppp1 using the plate washing assay [67]. As seen in figure 17, there was no invasiveness in the presence or

absence of Ppp1. This suggests that the role of Ppp1 in host pathogen interaction is not invasiveness. We can infer from these results that Ppp1 independently does not have an observable effect on the cell and may be part of a network of cells so analysis of only Ppp1 may not yield any useful information.

Microscopic analysis of cells can reveal morphological changes that are too small to be observed in the colonies of biochemical analysis. We used confocal microscopy to observe biofilm formation in Ppp1 mutant strains. Deletion of Ppp1 does not have an effect on biofilm formation but the complement strain shows a slight change in phenotype. In figure 19, *PPP1* forms slightly more biofilm formed compared to *ppp1 Δ/Δ* and WT in RPMI and when supplemented with 10% FBS in figure 20, it forms less hyphae. This reinforces the idea that phenotypes expressed by Ppp1 may be as a result of its genetic network.

To explore the genetic network of Ppp1, we looked at Orf19.6274 and Orf19.4654, proteins with similar characteristics as Ppp1. Orf19.6274 (Pbr1) and Orf19.4654 are small proteins with N-signal peptides highly upregulated in *Candida albicans* biofilms [36]. Pbr1 is a protein of 175aa with orthologs in *dubliniensis* and *tropicalis* while Orf19.4654 is protein of 109aa with an ortholog in *dubliniensis* [75]. Both have no annotated subcellular location and have similar log2 differential biofilm expressions: 8.79 for Pbr1 and 7.2 for Orf19.4654 [31]. As seen in figure 21, the protein expression pattern of Pbr1 and Orf19.4654 do not resemble Ppp1. Although these proteins share similarities to Ppp1, their protein expression patterns suggest that their localizations are different from Ppp1.

Overall, our study provides insight into the localization of the protein product and the biological function of a biofilm induced gene. We found that Ppp1 forms a punctate pattern distributed across the cell and this pattern colocalizes with that of Sed5, a Golgi resident protein. This was validated by protein visualization of Sap2, a secreted protein. We also identified the maturation stage as the phase of biofilm formation Ppp1 is involved in, thus speculating its function in glycoprotein synthesis for biofilm ECM. Biochemical assays to

test antifungal resistance and stress response on *ppp1 Δ/Δ* does not identify any alteration in growth but images from confocal microscopy show morphological differences in the complemented *PPP1* strain, suggesting that Ppp1 may belong to a genetic network. Images of Pbr1 and Orf19.4654, proteins similar to Ppp1 showed a different pattern to Ppp1. These findings show that Ppp1 has a pattern specific to the Golgi.

Future directions:

Ppp1 is a small protein with a signal peptide and no annotated retention motif. To verify that it is indeed retained in the secretory pathway, we can do a western blot of the supernatant of Ppp1 and check for post translational modifications like glycosylation on the extracted protein by treating with deglycosylase and comparing the protein size before and after treatment.

To further validate the localization of Ppp1, we could do more double tags with other markers of the Golgi. Getting same colocalization values will further solidify our conclusion on Ppp1's localization. We could also disrupt the Golgi without destroying it by using dCas9 to silence the expression of all other Golgi proteins and see if we still see the punctate pattern of Ppp1.

The Golgi is a complex structure made up of cisternae and lumen and involves SNARE proteins for intracellular transport. Our scope of analysis does not cover identifying which part of the Golgi Ppp1 resides in. This will be useful in further characterization of this protein and understanding what role it may play in biofilm formation. We could use the high resolution of electron microscopy to differentiate Golgi structural components and see where Ppp1 is localized or a super-resolution technique like photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) using immunofluorescence. We could explore its Golgi localization and function further by identifying when exactly Ppp1 is expressed in the Golgi, either early or late Golgi. Although our findings localize Ppp1 to the Golgi, radioactive pulse experiment measuring the transport of this protein will further validating our findings.

Deletion and complement strains of Ppp1 tested with different antifungals and stress responses did not show altered growth so we could investigate to see if Ppp1 is part of a complex, perhaps a family of proteins and how this affects its responses to different cellular and environmental factors. We can also investigate how this affects its response to biofilm formation. Construction of double mutants may provide insight on this, providing distinct phenotypes that can be identified [97,98].

Candida virulence through biofilm formation is a huge concern in healthcare especially as these cells are resistant to most antifungals in the market. High throughput subcellular localization can characterize the Orf biofilm genes uncovering more information like a biofilm complex starting with Pbr1 and Orf19.4654. This information can be relevant in antifungal therapeutics development specifically targeting and eliminating biofilm cells.

Bibliography

1. Woese C.R., Kandler O., & Wheelis M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*. Jun;87(12):4576-9. https://doi.org/10.1073/pnas.87.12.4576
2. Fricker M., Boddy L., & Bebber D. (2007). Biology of the fungal cell. *Springer.* pp. 309–330. https://link.springer.com/chapter/10.1007/978-3-540-70618-2_13
3. Editors B. D. (2017). Heterotroph - definition and examples. *Biology Dictionary*. https://biologydictionary.net/heterotroph/
4. Freitas F., Roca C., & Reis, Maria A.M. (2015). Fungi as Sources of Polysaccharides for Pharmaceutical and Biomedical Applications. In Thakur, VK; Thakur, MK (eds.). *Handbook of Polymers for Pharmaceutical Technologies: Vol 3, Biodegradable Polymers. Scrivener Publishing*. pp. 61–103. https://doi.org/10.1002/9781119041450.ch3
5. Tarkka, M. T., Drigo, B., and Deveau, A. (2018). Mycorrhizal microbiomes. *Mycorrhiza*, 28(5-6), 403–409. https://doi.org/10.1007/s00572-018-0865-5
6. Libretexts. (2021). The role of microbiota in human physiology. *Biology LibreTexts* 13.2. https://bio.libretexts.org/Courses/Mansfield_University_of_Pennsylvania/BSC_3271%3A_Microbiology_for_Health_Sciences_Sp21_(Kagle)/13%3A_The_Human_Microbiota/13.02%3A_The_Role_of_Microbiota_in_Human_Physiology
7. Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I. (2014). The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res.,42(1): D26-31. https://doi.org/10.1093/nar/gkr947

8. Petruzzello, M. (2023). Ascomycota. *Encyclopedia Britannica*. https://www.britannica.com/science/Ascomycota
9. Centers for Disease Control and Prevention. (2023). Invasive Candidiasis Statistics. *Centers for Disease Control and Prevention.* https://www.cdc.gov/fungal/diseases/candidiasis/invasive/statistics.html
10. Davis, C.P. (1996). Normal Flora. *Medical Microbiology 4th Edition*. chapter 6.https://www.ncbi.nlm.nih.gov/books/NBK7617/
11. Odds, F. C. (1988). *Candida* and Candidosis, a Review and Bibliography. *W. B. Saunders 2nd Edn*. https://doi.org/10.1002/jobm.3620300522
12. Mba, I. E., & Nweze, & E. I. (2020). Mechanism of *Candida* pathogenesis: revisiting the vital drivers. European journal of clinical microbiology & infectious diseases. *European Society of Clinical Microbiology*, 39(10), 1797–1819. https://doi.org/10.1007/s10096-020-03912-w
13. Wisplinghoff, H., Bischoff, T., Tallent, S. M., Seifert, H., Wenzel, R. P., & Edmond, M. B. (2004). Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*, 39(3), 309–317. https://doi.org/10.1086/421946
14. Noble, S. M., Gianetti, B. A., & Witchley, J. N. (2016). *Candida albicans* cell-type switching and functional plasticity in the mammalian host. *Nature Reviews Microbiology*, 15(2), 96–108. https://doi.org/10.1038/nrmicro.2016.157
15. Atriwal, T., Azeem, K., Husain, F. M., Hussain, A., Khan, M. N., Alajmi, M. F., & Abid, M. (2021). Mechanistic understanding of *Candida albicans* biofilm formation and

approaches for its inhibition. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.638609

16. Taff, H. T., Mitchell, K. F., Edward, J. A., and Andes, D. R. (2013). Mechanisms of Candida biofilm drug resistance. *Fut. Microbiol*. 8, 1325–1337. https://doi.org/10.2217/fmb.13.101
17. Perfect, J.R. (2017). The Antifungal Pipeline: A Reality Check. *Nature Reviews Drug Discovery*, 16(9), 603–616. https://doi.org/10.1038/nrd.2017.46
18. Seman, B.G., Moore, J.L., Scherer A.K., Blair B.A., Manandhar S., Jones J.M., & Wheeler R.T. (2018). Yeast and Filaments Have Specialized, Independent Activities in a Zebrafish Model of *Candida albicans* Infection. *Infection and Immunity*.;86(10):e00415-18. https://doi.org/10.1128/iai.00415-18
19. Malcok, H. K., Aktas, E., Ayyildiz, A., Yigit, N., & Yazgi, H. (2009). Hemolytic activities of the Candida species in liquid medium. *The Eurasian Journal of medicine*, 41(2), 95–98. PMID: 25610076. PMCID: PMC4299835
20. Deepa, K., Jeevitha T., & Michael A. (2015). In vitro evaluation of virulence factors of Candida species isolated from oral cavity. *J Microbiol Antimicrob* 7:28–32. http://dx.doi.org/10.5897/JMA2015.0337
21. Naglik, J. R., Gaffen, S. L., & Hube, B. (2019). Candidalysin: discovery and function in *Candida albicans* infections. *Current Opinion in Microbiology*, 52, 100–109. https://doi.org/10.1016/j.mib.2019.06.002
22. Tiwari, S., Thakur, R., & Shankar, J. (2015). Role of heat-shock proteins in cellular function and in the Biology of Fungi. *Biotech Res Int 11.* https://doi.org/10.1155/2015/132635

23. Westman, J., Moran, G., Mogavero, S., Hube, B., & Grinstein, S. (2018). *Candida albicans* Hyphal Expansion Causes Phagosomal Membrane Damage and Luminal Alkalinization. *mBio*, 9(5), e01226-18. https://doi.org/10.1128/mBio.01226-18

24. Vylkova, S., & Lorenz, M. C. (2014). Modulation of phagosomal pH by *Candida albicans* promotes hyphal morphogenesis and requires STP2P, a regulator of amino acid transport. *PLoS Pathogens*, 10(3). https://doi.org/10.1371/journal.ppat.1003995

25. Lewis, K.I.M. (2001). Riddle of biofilm resistance. *Antimicrob Agents Chemother*. 45:999–1007. https://doi.org/10.1128/aac.45.4.999-1007.2001

26. Donlan, R.M. (2001). Biofilms and device-associated infections. *Emerg Infect Dis.* 7:277–81. https://doi.org/10.3201/eid0702.010226

27. Nobile, C.J., & Johnson, A.D. (2015). *Candida albicans* Biofilms and Human Disease. *Annu Rev Microbiol*.;69: 71–92. https://doi.org/10.1146/annurev-micro-091014-104330

28. Pemmaraju, S. C., Padmapriya, K., Pruthi, P. A., Prasad, R., & Pruthi, V. (2016). Impact of oxidative and osmotic stresses on *Candida albicans* biofilm formation. *Biofouling*, 32(8), 897–909. https://doi.org/10.1080/08927014.2016.1212021

29. Kumar, D., Kumar, A. (2023). Molecular Determinants Involved in *Candida albicans* Biofilm Formation and Regulation. *Mol Biotechnol*. https://doi.org/10.1007/s12033-023-00796-x

30. Blankenship, J. R., & Mitchell, A. P. (2006). How to build a biofilm: A fungal perspective. *Current Opinion in Microbiology*, 9(6), 588–594. https://doi.org/10.1016/j.mib.2006.10.003

31. Nobile, C.J., Fox, E.P., Nett, J.E., Sorrells, T.R., Mitrovich, Q.M., Hernday, A.D., Touch, B.B., Andes, D.R., & Johnson, A.D. (2012). A recently evolved transcriptional network

controls biofilm development in *Candida albicans*. Cell;148(1-2):126-38. https://doi.org/10.1016/j.cell.2011.10.048

32. Talapko, J., Juzbašić, M., Matijević, T., Pustijanac, E., Bekić, S., Kotris, I., & Škrlec, I. (2021). *Candida albicans*—the virulence factors and clinical manifestations of infection. *Journal of Fungi*, 7(2), 79. https://doi.org/10.3390/jof7020079

33. Atriwal, T., Azeem, K., Husain, F. M., Hussain, A., Khan, M. N., Alajmi, M. F., & Abid, M. (2021). Mechanistic understanding of *Candida albicans* biofilm formation and approaches for its inhibition. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.638609

34. Richard, M.L., Nobile, C.J., Bruno, V.M., & Mitchell, A.P. (2005). *Candida albicans* biofilm-defective mutants. *Eukaryot Cell*, 4:1493-1502.

35. Nobile, C. J., Nett, J. E., Hernday, A. D., Homann, O. R., Deneault, J.-S., Nantel, A., Andes, D. R., Johnson, A. D., and Mitchell, A. P. (2009). Biofilm matrix regulation by *Candida albicans* ZAP1. *PLoS Biology*, 7(6). https://doi.org/10.1371/journal.pbio.1000133

36. The UniProt Consortium, UniProt: the Universal Protein Knowledgebase in 2023, Nucleic Acids Research, Volume 51, Issue D1, 6 January 2023, Pages D523–D531, https://doi.org/10.1093/nar/gkac1052

37. Wild, K., Halic, M., Sinning, I., & Beckmann, R. (2004). SRP meets the ribosome. *Nature Structural & Molecular Biology*, 11(11), 1049–1053. https://doi.org/10.1038/nsmb853

38. Keenan, R. J., Freymann, D. M., Stroud, R. M., & Walter, P. (2001). The signal recognition particle. *Annual Review of Biochemistry*, 70, 755–775. https://doi.org/10.1146/annurev.biochem.70.1.755

39. Cassone, A., De Bernardis, F., Mondello, F., Ceddia, T., & Agatensi, L. (1987). Evidence for a correlation between proteinase secretion and vulvovaginal candidosis. *The Journal of infectious diseases*, 156(5), 777–783. https://doi.org/10.1093/infdis/156.5.777

40. El-Maghrabi, E. A., Dixon, D. M., & Burnett, J. W. (1990). Characterization of *Candida albicans* epidermolytic proteases and their role in yeast-cell adherence to keratinocytes. *Clinical and experimental dermatology*, 15(3), 183–191. https://doi.org/10.1111/j.1365-2230.1990.tb02069.x

41. Hube, B, Sanglard, D, Odds, F.C., Hess, D., Monod, M., Schäfer, W., Brown, A.J., & Gow, N.A. (1997). Disruption of each of the secreted aspartyl proteinase genes SAP1, SAP2, and SAP3 of *Candida albicans* attenuates virulence. *Infect Immun*.;65(9):3529-38. https://doi.org/10.1128/iai.65.9.3529-3538.1997

42. Boone, C., Sdicu, A.M., Laroche, M., & Bussey, H. (1991). Isolation from *Candida albicans* of a functional homolog of the *Saccharomyces cerevisiae* KRE1 gene, which is involved in cell wall beta-glucan synthesis. *Journal of Bacteriology*, 173, 6859 - 6864. https://doi.org/10.1128/jb.173.21.6859-6864.1991

43. Kuranda, M. J., & Robbins, P. W. (1991). Chitinase is required for cell separation during growth of Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, 266(29), 19758–19767. https://doi.org/10.1016/S0021-9258(18)55057-2

44. Cukierman, E., Huber, I., Rotman, M., & Cassel, D. (1995). The ARF1 gtpase-activating protein: Zinc finger motif and Golgi complex localization. *Science*, 270(5244), 1999–2002. https://doi.org/10.1126/science.270.5244.1999

45. Faridmoayer, A., & Scaman, C. H. (2007). Truncations and functional carboxylic acid residues of yeast processing alpha-glucosidase I. *Glycoconjugate Journal*, 24(8), 429–437. https://doi.org/10.1007/s10719-007-9035-2

46. Rose, M. D., Misra, L. M., & Vogel, J. P. (1989). KAR2, a karyogamy gene, is the yeast homolog of the mammalian BiP/GRP78 gene. *Cell*, 57(7), 1211–1221. https://doi.org/10.1016/0092-8674(89)90058-5

47. Tachikawa, H., Takeuchi, Y., Funahashi, W., Miura, T., Gao, X.D., Fujimoto, D., Mizunaga, T., & Onodera, K. (1995). Isolation and characterization of a yeast gene, MPD1, the overexpression of which suppresses inviability caused by protein disulfide isomerase depletion. *FEBS Lett* 369 2-3):212-6. https://doi.org/10.1016/0014-5793(95)00750-4

48. Orlean, P. (1990). Dolichol phosphate mannose synthase is required in vivo for glycosyl phosphatidylinositol membrane anchoring, O mannosylation, and N glycosylation of protein in *Saccharomyces cerevisiae*. *Molecular and cellular biology*, 10(11), 5796–5805. https://doi.org/10.1128/mcb.10.11.5796-5805.1990

49. Nuoffer, C., Jenö, P., Conzelmann, A., & Riezman, H. (1991). Determinants for glycophospholipid anchoring of the *Saccharomyces cerevisiae* GAS1 protein to the plasma membrane. *Molecular and Cellular Biology*, 11, 27 - 37. https://doi.org/10.1128/mcb.11.1.27-37.1991

50. Mowbrey, K. & Dacks, J. (2009). Evolution and diversity of the Golgi body. *FEBS Letters*, 583, (23) 3738-3745. https://doi.org/10.1016/j.febslet.2009.10.025

51. Losev, E., Reinke, C. A., Jellen, J., Strongin, D. E., Bevis, B. J., & Glick, B. S. (2006). Golgi maturation visualized in living yeast. *Nature*, 441(7096), 1002–1006. https://doi.org/10.1038/nature04717

52. Wooding, S., & Pelham, H. R. (1998). The dynamics of Golgi protein traffic visualized in living yeast cells. *Molecular biology of the cell*, 9(9), 2667–2680. https://doi.org/10.1091/mbc.9.9.2667

53. Bonifacino J. and Glick B. (2004). The Mechanisms of Vesicle Budding and Fusion. *Cell*,116(2): 153-166. https://doi.org/10.1016/S0092-8674

54. Bonifacino, J. S., & Lippincott-Schwartz J. (2003). Coat proteins: shaping membrane transport. *Nature Reviews Molecular Cell Biology*, 4.5: 409-4. https://doi.org/10.1038/nrm1099

55. Rothman, J. E. (1994). Mechanisms of intracellular protein transport. *Nature*, 372.6501, 55-63. https://doi.org/10.1038/372055a0

56. Banfield, D. K. (2011). Mechanisms of protein retention in the Golgi. *Cold Spring Harbor perspectives in biology*, 3(8), a005264. https://doi.org/10.1101/cshperspect.a005264

57. Yuste, R. (2005). Fluorescence microscopy today. *Nature methods*, 2(12), 902–904. https://doi.org/10.1038/nmeth1205-902

58. Combs, C. A., & Shroff, H. (2017). Fluorescence microscopy: A concise guide to current imaging methods. Current Protocols in Neuroscience, 79, 2.1.1–2.1.25. https://doi.org/10.1002/cpns.29

59. Landmann, L., & Marbet, P. (2004). Colocalization analysis yields superior results after image restoration. *Microscopy research and technique*, 64(2), 103–112. https://doi.org/10.1002/jemt.20066

60. Dunn, K. W., Kamocka, M. M., & McDonald, J. H. (2011). A practical guide to evaluating colocalization in biological microscopy. *American journal of physiology*. Cell physiology, 300(4), C723–C742. https://doi.org/10.1152/ajpcell.00462.2010

61. Babbey, C. M., Ahktar, N., Wang, E., Chen, C. C., Grant, B. D., & Dunn, K. W. (2006). Rab10 regulates membrane transport through early endosomes of polarized Madin-Darby canine kidney cells. *Molecular biology of the cell*, 17(7), 3156–3175. https://doi.org/10.1091/mbc.e05-08-0799

62. Parmryd, I., Adler, J., Patel, R., & Magee, A. I. (2003). Imaging metabolism of phosphatidylinositol 4,5-bisphosphate in T-cell GM1-enriched domains containing Ras proteins. *Experimental cell research*, 285(1), 27–38. https://doi.org/10.1016/s0014-4827(02)00048-4

63. Bolte, S., & Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, 224(3), 213–232. https://doi.org/10.1111/j.1365-2818.2006.01706.x

64. Min, K., Ichikawa, Y., Woolford, C.A., & Mitchell, A.P. (2016). *Candida albicans* gene deletion with a transient CRISPR-Cas9 system. *mSphere* 1(3):e00130-16. https://doi.org/10.1128/mSphere.00130-16

65. Walther, A., & Wendland, J. (2003). An improved transformation protocol for the human fungal pathogen Candida albicans. *Current Genetics*, 42(6), 339–343. https://doi.org/10.1007/s00294-002-0349-08

66. Dunkel, N., & Morschhäuser, J. (2011). Loss of heterozygosity at an unlinked genomic locus is responsible for the phenotype of a *Candida albicans sap4Δ sap5Δ sap6Δ* mutant. *Eukaryotic cell*, 10(1), 54–62. https://doi.org/10.1128/EC.00281-10

67. Rashid, S., Correia-Mesquita, T. O., Godoy, P., Omran, R. P., & Whiteway, M. (2022). SAGA complex subunits in *Candida albicans* differentially regulate filamentation, invasiveness, and biofilm formation. *Frontiers in Cellular and Infection Microbiology*, 12. https://doi.org/10.3389/fcimb.2022.764711

68. Galdiero, E., de Alteriis, E., De Natale, A., D'Alterio, A., Siciliano, A., Guida, M., Lombardi, L., Falanga, A., & Galdiero, S. (2020). Eradication of *Candida albicans* persister cell biofilm by the membranotropic peptide GH625. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-62746-w

69. Ejaz, H., Ramzan, M. (2019). Determination of minimum inhibitory concentrations of various antifungal agents in clinical isolates of *Candida species*. *J Pak Med Assoc*. Aug;69(8):1131-1135. PMID: 31431766.

70. Roscetto, E., Contursi, P., Vollaro, A., Fusco, S., Notomista, E., & Catania, M. R. (2018). Antifungal and anti-biofilm activity of the first cryptic antimicrobial peptide from an archaeal protein against *Candida spp*. clinical isolates. *Scientific Reports*, 8(1). https://doi.org/10.1038/s41598-018-35530-0

71. Sharma, A., Solis, N. V., Huang, M. Y., Lanni, F., Filler, S. G., & Mitchell, A. P. (2023). HGC1 independence of biofilm hyphae in *Candida albicans*. *mBio*, 14(2). https://doi.org/10.1128/mbio.03498-22

72. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J.,

Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: An open-source platform for biological-image analysis. *Nature Methods*, 9(7), 676–682. https://doi.org/10.1038/nmeth.2019

73. Cavalheiro, M., & Teixeira, M. C. (2018). *Candida* biofilms: Threats, challenges, and promising strategies. *Frontiers in Medicine*, 5. https://doi.org/10.3389/fmed.2018.00028
74. Chandra, J., Kuhn, D.M., Mukherjee, P.K., Hoyer, L.L., McCormick, T., & Ghannoum, M.A. (2001). Biofilm formation by the fungal pathogen *Candida albicans*: development, architecture, and drug resistance. *J Bacteriol*;183(18):5385-94. https://doi.org/10.1128/jb.183.18.5385-5394.2001
75. Skrzypek, M.S., Binkley, J., Binkley, G., Miyasato, S.R., Simison, M., Sherlock. G. (2017). The *Candida* Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. *Nucleic Acids Res 45 (D1)*; D592-D596. https://doi.org/10.1093/nar/gkw924
76. U.S. National Library of Medicine. (n.d.). Blast: Basic local alignment search tool. *National Center for Biotechnology Information*. https://blast.ncbi.nlm.nih.gov/
77. Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology*, 157(1), 105–132. https://doi.org/10.1016/0022-2836(82)90515-0
78. Kyte-Doolittle Hydropathy Plots. Pearson, William R. 1998. U of Virginia. 28 Feb 2002 <http://fasta.bioch.virginia.edu/fasta/grease.htm>.
79. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler,

J., … Hassabis, D. (2021). Highly accurate protein structure prediction with alphafold. *Nature*, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

80. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., & Velankar, S. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic acids research, 50(D1), D439–D444. https://doi.org/10.1093/nar/gkab1061
81. Douglas, L. M., Wang, H. X., Li, L., & Konopka, J. B. (2011). Membrane compartment occupied by CAN1 (MCC) and eisosome subdomains of the fungal plasma membrane. *Membranes*, 1(4), 394–411. https://doi.org/10.3390/membranes1040394
82. Bean, B. D. M., Whiteway, M., & Martin, V. J. J. (2022). The MyLO CRISPR-Cas9 toolkit: a markerless yeast localization and overexpression CRISPR-Cas9 toolkit. *G3 (Bethesda, Md.)*, 12(8), jkac154. https://doi.org/10.1093/g3journal/jkac154
83. Epp, E., Walther, A., Lépine, G., Leon, Z., Mullick, A., Raymond, M., Wendland, J., & Whiteway, M. (2010). Forward genetics in *Candida albicans* that reveals the Arp2/3 complex is required for hyphal formation, but not endocytosis. *Molecular microbiology*, 75(5), 1182–1198. https://doi.org/10.1111/j.1365-2958.2009.07038.x
84. Roemer, T., Jiang, B., Davison, J., Ketela, T., Veillette, K., Breton, A., Tandia, F., Linteau, A., Sillaots, S., Marta, C., Martel, N., Veronneau, S., Lemieux, S., Kauffman, S., Becker, J., Storms, R., Boone, C., & Bussey, H. (2003). Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery.

Molecular microbiology, 50(1), 167–181. https://doi.org/10.1046/j.1365-2958.2003.03697.x

85. Ruis, H. & Schuller, C. (1995). Stress signaling in yeast. *Bioassays* 17, 959–965. https://doi.org/10.1002/bies.950171109

86. Kühn, C., & Klipp, E. (2011). Zooming in on yeast osmoadaptation. *Advances in Experimental Medicine and Biology*, 293–310. https://doi.org/10.1007/978-1-4419-7210-1_17

87. Lohse, M. B., Gulati, M., Johnson, A. D., and Nobile, C. J. (2018). Development and regulation of single-and multi-species *Candida albicans* biofilms. *Nat. Rev. Microbiol*. 16, 19–31. https://doi.org/10.1038/nrmicro.2017.107

88. Wilson, D., Mayer, F. L., Miramón, P., Citiulo, F., Slesiona, S., Jacobsen, I. D., & Hube, B. (2014). Distinct roles of *Candida albicans*-specific genes in host-pathogen interactions. Eukaryotic Cell, 13(8), 977–989. https://doi.org/10.1128/ec.00051-14

89. Chang, S., Joyson, M., Kelly, A., Tang, L., Iannotta, J., Rich, A., Castilho Coelho, N., & Carvunis, A. R. (2023). Unannotated Open Reading Frame in *Saccharomyces cerevisiae* Encodes Protein Localizing to the Endoplasmic Reticulum. *Micropublication Biology*, 2023, 10.17912/micropub.biology.000992. https://doi.org/10.17912/micropub.biology.000992

90. Augstein, A., Barth, K., Gentsch, M., Kohlwein, S. D., & Barth, G. (2003). Characterization, localization and functional analysis of Gpr1p, a protein affecting sensitivity to acetic acid in the yeast *Yarrowia lipolytica*. *Microbiology (Reading, England)*, 149(Pt 3), 589–600. https://doi.org/10.1099/mic.0.25917-0

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

92. Bonifacino, J.S., Glick, B.S. (2004). The Mechanisms of Vesicle Budding and Fusion. *Cell*,116,153–166. https://doi.org/10.1016/s0092-8674(03)01079-1

93. Dunkel, N., & Morschhäuser, J. (2011). Loss of heterozygosity at an unlinked genomic locus is responsible for the phenotype of a *Candida albicans* sap4Δ sap5Δ sap6Δ mutant. *Eukaryotic cell*, 10(1), 54–62. https://doi.org/10.1128/EC.00281-10

94. Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., et al. (1999). Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 285: 793–972 https://doi.org/10.1126/science.285.5429.901

95. Scherens, B., & Goffeau, A. (2004). The uses of genome-wide yeast mutant collections. *Genome biology*, 5(7), 229. https://doi.org/10.1186/gb-2004-5-7-229

96. Wagner, A. (2000). Robustness against mutations in genetic networks of yeast. *Nature genetics*, 24(4), 355-361. https://doi.org/10.1038/74174

97. Tong, A. H., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Pagé, N., Robinson, M., Raghibizadeh, S., Hogue, C. W., Bussey, H., Andrews, B., Tyers, M., & Boone, C. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science (New York, N.Y.)*, 294(5550), 2364–2368. https://doi.org/10.1126/science.1065810

98. Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G. F., Brost, R. L., Chang, M., Chen, Y., Cheng, X., Chua, G., Friesen, H., Goldberg, D. S., Haynes, J., Humphries, C., He, G., Hussein, S., Ke, L., …Boone, C. (2004). Global

mapping of the yeast genetic interaction network. *Science (New York, N.Y.)*, 303(5659), 808–813. https://doi.org/10.1126/science.1091317

APPENDIX

Oligonucleotides Used in the Study

Name	**Sequence (5' to 3')**
***SUR7*-SCARLET_sgRNA_F**	CGTATATTAAATATACCAATgttttagagctagaaatagcaagttaaa
***SUR7*-SCARLET_sgRNA_R**	ATTGGTATATTTAATATACGcaaattaaaaatagtttacgcaagtc
***SUR7*-SCARLET_repair_F**	CAGGCGGTATTAGATTCTTCAAAATCAAAAGAAACCAAAAAGTTTCCGATGATGAATCAGTAggtgctggcgcaggtgctatggtcagtaaaggggaagc
***SUR7*-SCARLET_repair_R**	TAAAGATTCCAATAATGGTAATACTGATAATAATAATAATAATAACAATGATGATTTTAGTAATAGTAGTAACAGTtctgatatcatcgatgaattcgag
***SUR7*-SCARLET_diag_ext_F**	GTCTCATTGCCCTTGCATTCAGTG
***SUR7*-SCARLET_diag_ext_R**	CACAATCCATATGTAACTCATGTCACG
Diag_*SCARLET*_R2	Ggcatttgcacaggtttcttagc
Diag_*URA3*-F1	Gaaactcatgcctcaccagtagc
***ARC35*Δ/Δ_sgRNA_F**	TGGCAGCTCCATTTGAACAAgttttagagctagaaatagcaagttaaa

***ARC35*Δ/Δ_sgRNA_R**	TTGTTCAAATGGAGCTGCCAcaaattaaaaatagtttacgcaagtc
***ARC35*Δ/Δ_repair_F**	TTGAAAACTCATAATTTCTCCATTCACCATTAACTTCTAGCA ACTCATATCAATTGATCTTTTATCCCCTACTACCgaagcttcgtacgct gcaggtcga
***ARC35*Δ/Δ_repair_R**	GGTTAGGTTAGGTTAGTTAATATATATCTTGTACATAAATTGA TACACACATAACATAGTACTCTTTTTTTTTtctgatatcatcgatgaattcg agctc
***ARC35*Δ/Δ_diag_ext_F**	TAGTTTTGGTCATAAAAGCCAACCTTTC
***ARC35*Δ/Δ_diag_ext_R**	TTCTGCATTAAGAAAATGTGCCAGA
Diag_*ARG4*_F	Accaaattttatgactaactgtgggtcg
Diag_*ARG4*_R	Cagcttttctaacacattcaccagatat
***ARC35*-SCARLET_sgRNA1_F**	ACACTATTTCTTTCTATTGTgttttagagctagaaatagcaagttaaa
ARC35-SCARLET_sgRNA1_R	ACAATAGAAAGAAATAGTGTcaaattaaaaatagtttacgcaagtc
***ARC35*-SCARLET_sgRNA2_F**	AACCTAACCTACACTATTTCgttttagagctagaaatagcaagttaaa
***ARC35*-SCARLET_sgRNA2_R**	GAAATAGTGTAGGTTAGGTTcaaattaaaaatagtttacgcaagtc
***ARC35*-SCARLET_repair_F**	ATGAAGGTAAAGTTATTGAAAATAGAAAAACCGCTAGTGGT AGAAGATTTGATATTAGAGCGggtgctggcgcaggtgctatggtcagtaaagg ggaagc

***ARC35*-SCARLET_repair_R**	TTGTATTGGTTAGAGGTGGTAGAGTTCAAGATGTTCCTGGGGTTAAATATCATTTGGTTAGAGGTGCTTTTGATtctgatatcatcgatgaattcgag
***ARC35*-SCARLET_diag_ext_F**	GAATTTGTTGATGCTAGGAAAAGATC
***ARC35*-SCARLET_diag_ext_R**	CATTCAATCAAGTTCGTCATTCAT
***SED5*-SCARLET_sgRNA1_F**	GACATGGATTGGTCGTGTAAgttttagagctagaaatagcaagttaaa
***SED5*-SCARLET_sgRNA1_R**	TTACACGACCAATCCATGTCcaaattaaaaatagtttacgcaagtc
***SED5*-SCARLET_repair_F**	GGTTATTTTTGAAAATATTTGGCGTTTTAATAGTATTTTTCTTCTTATGGGTATTAGTTAGTggtgctggcgcaggtgctatggtcagtaaaggggaagc
***SED5*-SCARLET_repair_R**	AAAATCTTGAATCTAGTTGATTATTTATCTAAAAAGACAAATGGAAGATTGAGCGACCAAAAAAAAAAAGAAAATG tctgatatcatcgatgaattcgag
***SED5*-SCARLET_diag_ext_F**	CCTTCGCAAACACAACAAATGTTACT
***SED5*-SCARLET_diag_ext_R**	TTTGATTTGGATGTATTTCCGGGATG
***SAP2*-SCARLET_sgRNA1_F**	TTTTACTTACTTGAAAGGAGgttttagagctagaaatagcaagttaaa
***SAP2*-SCARLET_sgRNA1_R**	CTCCTTTCAAGTAAGTAAAAcaaattaaaaatagtttacgcaagtc

***SAP2*-SCARLET_sgRNA2_F**	TTAGTTTTTACTTACTTGAAgttttagagctagaaatagcaagttaa
***SAP2*-SCARLET_sgRNA2_R**	TTCAAGTAAGTAAAAACTAAcaaattaaaaatagtttacgcaagtc
***SAP2*-SCARLET_repair _F**	ATAATGAAATTTCTTTGGCTCAAGTCAAATATACTTCTGCTTCCAGTATTTCTGCCTTGACCggtgctggcgcaggtgctatggtcagtaaaggggaagc
***SAP2*-SCARLET_repair _R**	TATATATTTCTAAATATATTATGTACTTTTAATTTAAATGGTATTTTAATGAAATTTAAGAATCGAATCGAAAAAAtctgatatcatcgatgaattcgag
***SAP2*-SCARLET_diag_ext_F**	TCGATGTTCTTGTGGATTCAGGTACCA
***SAP2*-SCARLET_diag_ext _R**	ACATAATTTATGGGAAACCGTTGAAACC
***ORF19.4654*-SCARLET_sgRNA1_F**	CTTCTCTGTACATTAAGTTAgttttagagctagaaatagcaagttaaa
***ORF19.4654*-SCARLET_sgRNA1_R**	TAACTTAATGTACAGAGAAGcaaattaaaaatagtttacgcaagtc
***ORF19.4654*-SCARLET_sgRNA2_F**	CTACTCATTGGTTGTTCTAGgttttagagctagaaatagcaagttaaa
***ORF19.4654*-SCARLET_sgRNA2_R**	CTAGAACAACCAATGAGTAGcaaattaaaaatagtttacgcaagtc
***ORF19.4654*-SCARLET_repair_F**	ATATTCAATGGAAATTATTATTAGGTTGTTTTTTATTTGCAATTGTAAGTTTACTTGCAATGggtgctggcgcaggtgctatggtcagtaaaggggaagc

***ORF19.4654*-SCARLET_repair_R**	TAGAACTGAATCTGGGTATATTTTTTGTTAATTTCGGTCCGATTTAGCTATTTGCTATGGTATTAAGTTCCCCATTctgatatcatcgatgaattcgag
***ORF19.4654*-SCARLET_diag_ext_F**	CAGCAACAACATTGTCAACCACCAA
***ORF19.4654*-SCARLET_diag_ext_R**	CGCGCGTTCTTTTCTCCTCCT
***ORF19.6274*-SCARLET_sgRNA_F**	GTTTAATTAATTGTTAAGTTgttttagagctagaaatagcaagttaaa
***ORF19.6274*-SCARLET_sgRNA_R**	AACTTAACAATTAATTAAACcaaattaaaaatagtttacgcaagtc
***ORF19.6274*-SCARLET_repair_F**	ATCTGCCACCAAAGAACTCCTACAAAGTTAGTATCTACGGTCGTCTCAATTGGGCTGTCTTGggtgctggcgcaggtgctatggtcagtaaagggaagc
***ORF19.6274*-SCARLET_repair_R**	TTATATATATATGTACACATAGTATAAGAACTAATAAATAAGCAACGAGAAGAAACGAAATAAAACAACGACAACGtctgatatcatcgatgaattcgag
***ORF19.6274*-SCARLET_diag_ext_F**	GATCAAAAGAGTGGTGAAATCAAA
***ORF19.6274*-SCARLET_diag_ext_R**	GGGCTCAAGGGTAACTCTTCTTT

*Capital letter represent genomic sequence

*Small letter represent plasmid sequence, marker or linker