Optimising recombinant protein secretion in wild strains of *Saccharomyces cerevisiae*

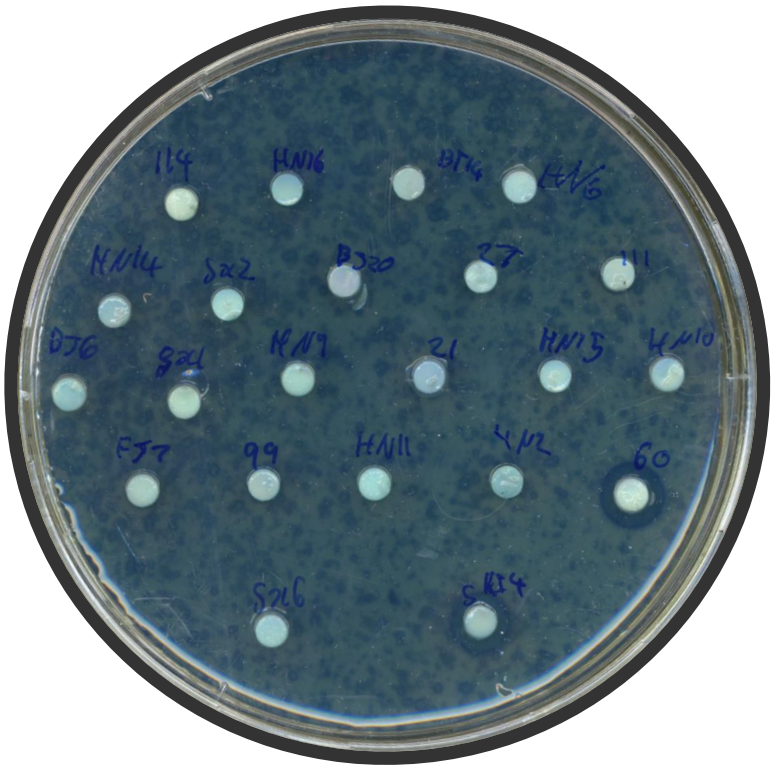
Andrew Strange

Biochemistry MSc

School of Biosciences

University of Kent

Supervisor: Professor Mick Tuite



Declaration

No part of this thesis has been submitted in support of an application for any degree or qualification of the University of Kent or any other university or institute of learning.

Andrew Strange

Acknowledgements

I wish to first thank Professor Mick Tuite, who has supervised thought this project, providing invaluable oversight and wisdom

I am extremely grateful both Gemma Staniforth and David Beal for putting up with my constant questions and for the guidance and help without which I would have been lost, truly they both have been founts of knowledge during this project.

And to everyone in Kent Fungal Group who made me feel welcome and surrounded by friendly helpful faces, I am thankful.

Finally, I would like to thank my every caring parents, and my eternally understanding partner Laura Carman whose boundless willpower is a constant inspiration to me, I owe the success of this project, and in my higher education, to them.

Contents

Abstract

Yeasts such as *Saccharomyces cerevisiae* and *Pichia pastoris,* are challenging the dominance of *Escherichia coli* as the preferred unicellular system for the manufacture of biopharmaceuticals, primarily because of their ability to produce and secrete complex proteins containing disulphide bonds. The most significant bottleneck in the secretion of recombinant proteins from yeast cells is the relatively low abundance of key processing enzymes (such as Sil1 or Lhs1), and chaperones in the secretion pathway yet nevertheless laboratory strains of yeasts have been successfully engineered to secrete a wide range of complex recombinant proteins. In this project I am asking whether a group of unmodified wild strains are efficient at secreting large amounts of recombinant protein. To do this I am examining the efficacy of protein secretion in a range of novel wild strains some of which have been discovered in remote uninhabited regions around the globe. To achieve this aim, I have first characterised the ability of various wild strains to secrete the yeast dsRNA virus-encoded 'killer toxin' (KT). A 'halo' assay has been used to quantify the amount of KT secreted by each strain. To this end technology to insert the KT gene into strains lacking auxotrophic makers was developed. This analysis has also included characterising the secretome of the strains. 8 strains were identified as possible hosts for recombinant protein production, with 3 of these strains showing a high level of secretion. A set of industrial used 'domesticated 'strains were also used as a control. These strains had KT secretion tested as well as gaussia luciferase secretion. There was a marked difference between the strains and between the two different proteins used, in terms of secretion levels

Abbrviations

µg	Microgram
µl	Microliter
A.U.	Arbitrary units
Amp	Ampicillin
CFU	Colony Forming Units
CHO cells	Chinese hamster ovary cells
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic Acid
E.coli	Escherichia coli
ER	Endoplasmic reticulum
Et al	Et alia
EtBr	Ethidium Bromide
g	Gram
H_2O	Water
hrs	hours
IBs	Inclusion bodies
kDa	KiloDalton
KT	Killer toxin
Leu	Leucine
LiAc	Lithium Acetate
mins	minutes
ml	Millilitre
mm	milimeter
O.D600	Optical Density at a 600 nanometre wavelength
P. pastoris	Pichia pastoris
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PTM	Post-translational modification
RNA	Ribonucleic Acid
ROS	Reactive oxygen species
Rpm	Revolutions per minute
S.cerevisiae	Saccharomyces cerevisiae
SD	Synthetic Defined
SDS PAGE	Sodium Dodecyl Sulfate Polyacrylamide
sgRNA	Single guide RNA
TEMED	Tetramethylethylenediamine
tRNA	transfer RNA
Ura	Uracil
V	Volt
w/v	Weight per volume
wt	Wild type
YEPD	Yeast Extract Peptone Dextrose

1. Introduction

The harnessing of single-celled organisms to produce recombinant proteins revolutionised the production of proteins, both in an industrial and research setting, and has led to hugely more efficient production methods. Whereas before many tons of bio matter were required to purify very small amounts of proteins such as insulin, the use of single-celled organisms as recombinant protein expression hosts has allowed less expensive and rapid production of a wide variety of useful biopharmaceuticals, enzymes and other useful bio-products. In the following, I review the improvements made over the last few decades that have increased the efficiency of secretion of recombinant proteins when using yeasts as host organisms. I also look at both *E. coli* and Chinese hamster ovary cells as host for recombinant protein expression.

1.1 Recombinant protein expression systems

Escherichia coli was the first microbial species to be used for recombinant protein production and still remains the preferred cellular host with 50% of all commercial proteins being produced In *E. coli*[1] . The reasons for its dominant position are as a recombinant host system; that it is easy to culture and has been optimised in different ways to inexpensively produce useful products[2,3]. This is achieved whilst having a very high protein yield in comparison to eukaryotic systems[4]. However there are highly notable downsides to using an *E. coli* host system, some of which are explored below.

1.1.1 Uncommon codon interference in recombinant protein production in *Escherichia coli*

There are 61 mRNA codons possible with the genetic code, however these codons are not all represented evenly in the genome of a given organism. In fact different organisms display different biases[5], and as a result, the levels of tRNA available to a cell reflect its codon bias.

The tRNA levels in *Escherichia coli* are no exception and reflect the levels of use of endogenous codon usage[6] to optimise the decoding speed of endogenous proteins. However, many recombinant proteins are modified proteins originating in eukaryotes (e.g. monoclonal anti-bodies) which have a different set of codon biases. These different biases leads to a drop in efficiency of decoding, frameshifts, swapping of similar amino acids (e.g. lysine for arginine), or even stalling or premature terminator of translation, leading to a truncated protein[7].

Current approaches to this problem include codon optimisation; the process of altering the codons in a given gene, to better complement the native tRNA levels in the host system[8,9,10]. Whilst codon optimisation has shown an increase protein production speed in *E. coli*[9], it is a costly and time consuming process. Depending on the exact codon makeup of the original gene used, large regions may have to be altered for efficient protein expression. Snajder et al found that to generate high levels of Pernisine (a protein used in industrial cleaning), 25.3% of the gene was altered[10] to allow efficient production levels.

Whilst codon optimisation may be required in yeast to expresses proteins that originate in bacteria[11], the increase in production of mono-clonal anti-bodies, and other biopharmaceuticals with their origins in eukaryotes makes this requirement much less common than seen in *E.coli* systems.

1.1.2 The inclusion body problem

Inclusion bodies (IBs) are the result of protein aggregates and have often been observed in recombinant *E. coli* since the beginning of its use as an expression host. IBs are formed of aggregated proteins, as a result of unbalance in the aggregation/solubilisation equilibrium. This equilibrium is pushed towards the crystallisation of proteins, when large volumes of

non-native proteins are produced, the results of strong promoters or high inducer concentrations. This large volume of protein far outstrips the cell's ability to properly fold them, and so the excess proteins begin to aggregate, and form IBs[12].

IBs act as a large bottle neck to protein secretion, having been observed to occupy up to 20% of the total cellular volume at peak protein production[13]. Large volumes of active bio product can become trapped inside IBs, and are wasted. whilst catalytic activity has been observed in IBs and processes have been developed to recover bio-active products from IBs, the yields these processes are often low, and the processes themselves are often labour intensive[14].

1.1.3 Post-translational modifications in *Escherichia coli*

Many recombinant proteins have complex post-translational modifications (PTM); these range from N- and O-linked glycosylation, to multiple disulphide bonds. These modifications are added by families of enzyme (such as PDI, or NMT-1), and often require co-factors. Eukaryotic cell have sets of chaperones, within membrane-bond organelles, which ensure the conditions to correctly create these PTMs are generated. For example the formation of disulphide bonds, both intra- and inter-protein, requires an oxidising environment, which is not found in the cytoplasm of wt *E.coli*, and proteins must be exported to the periplasm to be properly folded. This lack of folding increases the chance of the formation of inclusions bodies (see above), especially in disulphide-rich proteins, such as Fab anti-body fragments.

Many recombinant proteins require glycosylation for proper folding and function. Non-glycosylated versions of glycoproteins tend to be misfolded, biologically inactive or quickly broken down by the cell[15,16]. *E. coli* and other prokaryotes lack much of the machinery

required for glycosylation of any kind, and this can present a major issue in recombinant protein production[17].

1.2 Chinese hamster ovary cells as protein expression hosts

Chinese hamster ovary (CHO) cells are often used as an expression host for recombinant proteins. Being eukaryotic cells, CHO cells have the ability to produce a large range of PTMs, and so can be used to produce a large range of recombinant proteins especially those needed by biopharma companies. For example, 3D6 single chain Fv-Fc anti-HIV-1[18], an anti-HIV antibody or recombinant human (rh)Thrombin[19]. As CHO cells are genetically similar to humans, containing many of the same chaperones and enzymes, it follows that CHO cells are used for some of the more complex biopharmaceuticals. However, whilst the native machinery in yeasts is less complex and possible less efficient than in mammalian cells, such as CHO, these are not the only factors to consider when looking at overall recombinant protein secretion. It has been shown that 'less complex' proteins such as human serum albumin are secreted in much greater quantities by yeasts than by CHO cells due to several factors such as cellular replication time, and maximum cell density[20].

There are also economic factors to consider: CHO cells require more complex media in which to grow in. This tends to be very expensive, compared to the inexpensive media used for yeasts such as *S. cerevisiae* or *P. pastoris*. CHO cell lines are also much more expensive to buy than yeasts or bacterial hosts, raising the initial investment required.

CHO cells have demonstrated sensitivity to environmental conditions such as heat and Ph, relative to yeasts or bacterial hosts. CHO cells have also show a vulnerability to viral infection[21], which can lead to the loss of entire experimental culture, and requires decontamination, or an entirely new culture to avoid re-infection.

1.3 Yeasts as a recombinant protein expression host

Yeasts such as *Saccharomyces cerevisiae* and more recently *Pichia pastoris,* are challenging the dominance of *E. coli*. Yeast has been used to produced recombinant proteins by the biopharmaceutical industry since 1977, when Novo Nordisk used *S. cerevisiae* to produce human insulin to replace the traditional pig insulin that was used at the time[21]. By 2012, approximately 20% of all protein-based biopharmaceuticals (by mass) on the market were produced in *S. cerevisiae*[22].

There are several reasons to use a yeast cells as the host over a bacterial one when trying to express complex mammalian proteins, the most notable of which is the lack of complex PTM and post-translational processing within prokaryotes [23], such as complex folding, glycosylation, phosphorylation and the removal of signal sequences.[24] All these processes take place in yeasts.

There has been a great deal of effort over the last 25 years or so, to improve *E. coli* as a host. Yet despite this, the complex cellular machinery required to properly fold proteins including the addition of sugar groups (i.e. glycosylation) or other complex PTMs such as disulphide bond formation , remain the domain of eukaryotes such as the yeasts[23]. Whilst there is strong evidence that prokaryotes can perform simple PTMs such as Ser-Thr phosphorylation[25], the lack of dedicated 'quality control' centres such as the Golgi apparatus and endoplasmic reticulum (ER), mean proteins containing more complex PTMs such as glycosylation or multiple disulphide bridges, are either of low quality, or are simply impossible to produce in a correctly modified and folded state[26]. The lack of complex protein folding and chaperoning systems in prokaryotes also causes the build-up of aggregated, misfolded proteins in inclusion bodies. Eukaryotes such as yeast, however, can perform

these functions, due to their complex organelle systems. For these reasons, we see an increasing number of recombinant proteins are being expressed in yeast, especially for larger more complex proteins[21] such as MUC1, or human podoplanin, a glycoprotein notable for the extensive O-linked glycosylation of its extracellular domain and a platelet-aggregating factor respectively[27].

1.4 Expansion of the abilities yeast to produce PTM in recombinant proteins

Whilst yeasts can produce many of the more complex PTMs seen in human proteins, they still lack many of the enzymes required to fully replicate the full range of PTMs found on human proteins, most notable of which are the complex glycans. This has necessitated the use mammalian cells in these cases, which comes with many associated difficulties (**Section 1.2**). It would therefore be ideal to be able to take some of the mammalian enzyme pathways and insert them into a host that is easier to work with. It is this line of thought that has led to the 'humanisation' of the yeast protein folding machinery[28].

The first few stages of N-glycosylation are identical between the human and the yeast cell: it begins with the transfer of N-acetylglucosamine onto dolichol phosphate, on the cytoplasmic face of the ER, before it is transported through the membrane, to the lumen face by a flipase. More mannoses are added, until Glc3Man9GlcNAc2-Pdolichol is generated, at which point the structure is transferred to the target protein at a N-X-S/T motif, and the whole protein is sent to the Golgi apparatus. At this point, human cells use the action of several 1,2-mannosidases[29], which cleave the Man8GlcNAc2 glycan present down to Man5-GlcNAc2. From there, further modification turns the glycan into a sialylated glycan. The lack endogenous 1,2-mannosidases in yeasts means they can only keep adding mannose sugars, producing hyper-mannosylated glycans. Hamelton et al[29] demonstrated that this hyper-

mannosylation can be prevented by mutating the genes encoding the enzymes responsible for the first mannose added, i.e. the *OCH1* and *MNN1* genes. By using this as a starting point, a novel path was engineered into *P. pastoris*, that produced sialylated glycan albeit in a slightly different way to mammals[29]. A summary of this path is shown in **Figure 1**.

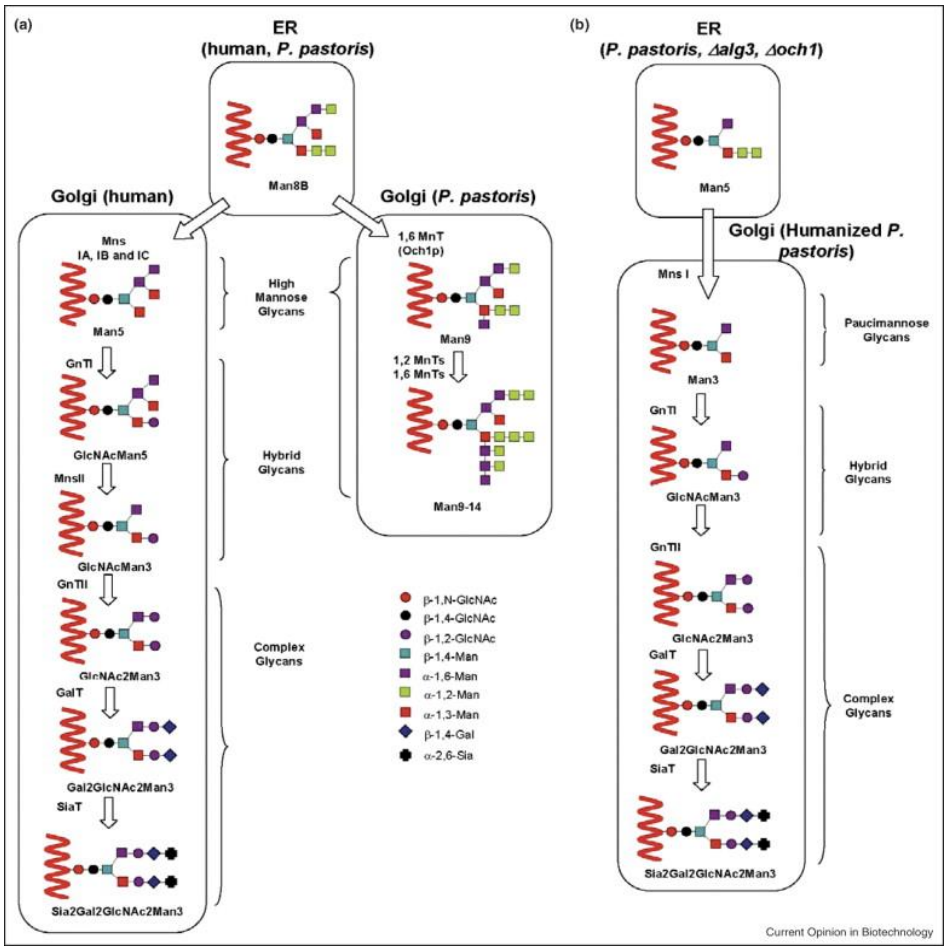


Figure 1: **A comparative pathway set N-linked glycosylation in human and *P.pastoris*(a) and humanised *P.pastoris* (b).** Inactivating the *ALG3* and *OCH1* genes stops additional mannose being added in *P.pastoris*, which would normally lead to hyper-mannosylation. Instead, upon reaching the Golgi, the simple glycan is converted immediately into GlcNAcMan3, bypass a some of the steps seen in human, but ultimately producing the same glycan structure. Figure taken from Hamilton et al[29]

The development of novel pathways for complex N-linked glycosylation has allowed the generation of various more complex biopharmaceuticals: for example, there has been much research concerning the production of human monoclonal antibodies for pharmaceutical

application. Traditionally, recombinant antibodies have been produced by batch fed mammalian cell methods, which are both costly, and inefficient[30]. Attempts have been made to remedy this by the development of transgenic plants and animals for monoclonal antibody production with the protein being successfully expressed in plants[30] and the milk of transgenic goats[31]. However, we now have the ability to produce different monoclonal antibodies in a range of different yeast hosts for both research and pharmaceutical purposes[1,32,33].

1.5 The use of different yeast species.

A major improvement in recombinant protein production came in 1985[34] with the harnessing of alternatives to the original *S. cerevisiae* i.e. the methylotrophic yeast *Pichia pastoris*. This species of yeast holds a few advantages over *S. cerevisiae*. Firstly (and primarily) its ability to use methanol as a carbon source, which is driven by two genes: *AOX1* and *AOX2*[35] both of which are strongly induced by the presence of methanol in the culture medium. Using either of these genes allows any plasmid inserted to have a built in, easily activated, high efficiency promoter. *P. pastoris* can also be easily genetically manipulated[35]. Since the mid-1980s numerous yeast species have been used as expression hosts, a few examples of which can be seen in **Table 1**, each with specific advantages and draw backs. These yeast species can all be genetically manipulated relatively easily, which helps make them useful as expression hosts.

Table 1: A selection of yeasts species and the date when they were first used as a host for recombinant proteins. Each strain has a specific trait, or traits, which renders it useful in a particular situation.

Species name	First described	Example protein expressed
Saccharomyces cerevisiae	1977	Insulin
Pichia pastoris	1985[34]	single-cell protein
Schizosaccharomyces pombe	1994[36]	Human lipocortin I
Hansenula polymorpha	1994[37]	human urinary plasminogen activator
Candida boidinii	1996[38]	Fungal glucoamylase
Pichia methanolica	1998[39]	Human glutamate decarboxylase

1.6 Strategies for optimisation protein production

1.6.1 Manipulation of the quality control system

The major bottleneck in recombinant protein production in yeasts is in their secretion from the cell which can reduce the effective level of protein production by as much as 1000-fold[24]. One of the main factors affecting this secretion is the so called 'quality control'[40] systems that are so important to the production of complex PTMs. These quality control systems involve multiple complex enzyme pathways, often in the ER, which ensure proteins are correctly translated and folded, and function at both the mRNA[41] and protein level. The large number of enzymes and chaperons involved in these pathways make it difficult to modify effectively. Despite this there has been much work published on the subject of improvement of the quality control system, much of it focusing on the optimisation of *S. cerevisiae*, the

most prominent of the yeast hosts. The simplest way of speeding up the whole system is to over express the enzymes involved: For example, up regulation of certain genes linked to ATPase activity (such as the chaperone protein producing genes: *SIL1, LHS1, JEM1*, and *SCJ1*) via random mutagenesis, leads to enhanced secretion of recombinant human serum albumin[42] from *S. cerevisiae*. Overexpression and mutation of the disulphide bond forming protein PDI showed a similar increase in protein production in numerous proteins[43,44,45], presumably caused by the increase in the cells ability to produce disulphide bonds at a high speed.

1.6.2 Manipulation of growth conditions

There have been a number of reports on the effects of stress on yeasts and the vast metabolic changes that occur when a cell is stressed, whether that is thermal stress, oxidative stress or osmotic stress. Harnessing these changes is a fruitful and rapidly evolving avenue of research. Heat stress is the most common target for investigation, as the heat shock proteins are heavily involved in refolding of proteins that have become denatured during heat stress. Mutating endogenous genes, to exploit these stress responses, can increase the speed of protein folding, and so the speed of protein production. For example, Hou et al. [46] showed, that by mutating the heat shock gene *HSF1* to be constituently active, they induced the cell's natural heat shock response. By putting the cell into such a stress recovery state, the folding of some proteins, such as β-amylase (an enzyme that can be used in industrial ethanol production[47]) can be increased by up to 75%[46]. However, this does not work for every protein; for example, human insulin precursor shows little to no improvement in production levels[46] in the same mutant strain.

Growth conditions such as temperature may also have an effect on recombinant protein expression in yeast: growth in colder temperatures (20 °C rather than the more common 30 °C for yeast growth) has been show to increase the production of Fab fragments, by as much as 3 fold[48]. Low Ph has shown a similar increase in protein production in *P.pastoris*[49]

There is also a consideration of how the yeast cells should be grown, as most laboratories use batch method, on small scales, whereas large industrial growths usually uses fed batch methodologies. The effect of this can be quite dramatic with Scheidle et al reporting a 53-fold increases of GFP secretion in *H. polymorpha,* when using an batch fed method versus a regular batch method[50].

1.6.3 Plasmid copy number optimisation

The link between gene copy number and protein production and secretion has been long established, whether by using a plasmid-based strategy or chromosomal integration of multiple gene copies[51]. However, putting a large number of plasmid copies into a cell causes a great deal of metabolic stress, and means more resources are diverted away from normal metabolic operation, such as reproduction[52]. Unfortunately, the exact effect of strain ploidy, copy number, plasmid burden, selection marker, and promoter selection on the host cell all remain uncertain.[53]

Despite the lack of a precise understanding of the underlying mechanism of metabolic stress caused by large numbers of plasmids, much reserch has been done to establish the exact copy number at which the highest level of the target protein can be produced. This number is dependent on the host, the plasmid used and the relative complexity of the protein in question. For example, in *S. cerevisiae ,* expressing porcine insulin precursor (PIP), the

optimum plasmid copy number is 12[54], whereas expressing PIP in *P. pastoris* shows an increased level of secretion up to (and possibly above) 18 copies[55].

1.6.4 High-throughput screening for high secretion levels

The strategies so far described have involved discrete and often pre-planned changes, each using a limited number of species and growing up cells in batches. These types of methods for the improvement of recombinant protein secretion in yeasts are not the only way to improve secretion of desired proteins. An alternative strategy is to screen a great many individual strains for the desired trait i.e. high through-put screening (HTS). For example each member of a large library of cells could have random mutations introduced by one of serval methods (degenerated oligonucleotides, chemical mutagenesis or, the most popular, error-prone PCR), with these mutation often being targeted to a specific gene[56]. The main limit of this process is the rate at which the mutant library can be screened, which determines the size of the library generated, and so the chance of identifying a novel strain at the high end of the secretion rate bell curve[56]. HTS can have a drastic effect on secretion levels of recombinant proteins, Huang et al showed a 6 fold increase in the secretion of recombinant β-amylase, in strains of *S. cerevisiae* mutated with UV light, against a control strain[57].

1.6.5 Use of non-domesticated yeast strains

An alternative approach to the problem of generating more efficient recombinant protein secreting strains is looking at understudied wild strains of yeast, which have had minimal human exploitation, and as a result have undergone no genetic manipulation. Analysis of these strains has shown to have a large genetic diversity[58]. Such strains may show a great efficiency in their secretion abilities; these strains may naively secrete the endogenous

protein Killer toxin (KT) (**Section 1.7**), which would lead them to have the necessary machinery in place, which could be harnessed in recombinant protein secretion.

By looking at the secretion related phenotypes of these strains, using a model secreted protein we can identify strains with enhanced protein secreting abilities. It is from this core idea that this current project comes; by comparing these strains in to some 'standard wild strains' I hope to identify a set of useful strains, which are able to efficiently secrete complex disulphide bonded proteins. Once this set of strains has been identified, the next stage of later projects would be to perform genetic and biomolecular analysis and identify the core differences responsible for these phenotypes.

1.6.6 The wild strains used in this project

For this project a collection of 21 wild strains were used. The majority of these strains come from parts of China; largely from the mountainous forest covered island Hainan, just off the south-western coast of the Guangdong province and have been isolated from tree bark. In addition there are strains isolated from the Ecuador and Slovakia, from river beds and insects. This diverse range of yeasts should produce a large range of phenotypes, and is a fertile ground to start looking for new 'super secreting' strains.

1.7 Killer toxin

This project required a model protein that was secreted from the cell, have a simple and reliable bioassay, and contain post translation modifications. The pore-forming toxin[59] K1 Killer Toxin, a 19.088 kDa[60] heterodimer containing 3 disulphide bonds was chosen. K1 is one of three killer toxins that are found in *S. cerevisiae*, the others being K2 and K28. K1 KT also has a simple assay, called the halo assay, which is described in **section 2.5.1**

1.7.1 Biosynthesis of the mature killer toxin

The pre-pro toxin (**Figure 3A**) consists of 5 components: a 'pre' signal sequence, a small delta chain and the pro-toxin, which is made up of the alpha beta and gamma chains. Production of the mature protein requires processing in the Golgi apparatus, and ER. In the ER the alpha and beta chains[61] are bonded with disulphide bonds at residues 92 to 248, 95 to 312 and 107 to 239 (**Figure 3B**) and the gamma chain is N-glycosylated in three places. The pro-toxin is then transported to the Golgi where it is cleaved into each individual chain by the enzyme Kex2p, and then the C terminus of the β-chain is trimmed by the carboxypeptidase Kex1p, to create the secreted toxin (**Figure 4**). This pathway is shared by all 3 toxins, though the amount of disulphide bonds differs between each toxin[62].

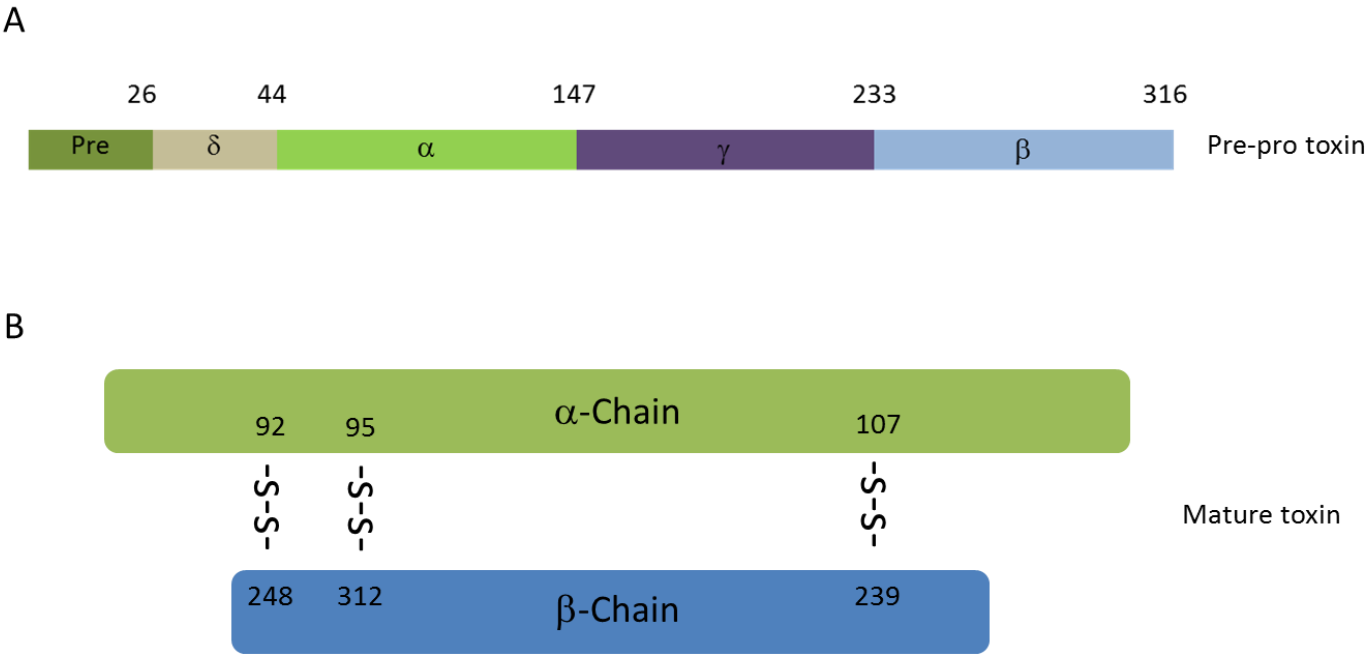


Figure 3; **Schematic diagram of the K1 pre-pro toxin as well as the mature toxin.** A; The unprocessed pre pro toxin as translated from the mRNA. The start residue for each region are noted. B; The mature K1 killer toxin, with position of disulphide bounds noted.

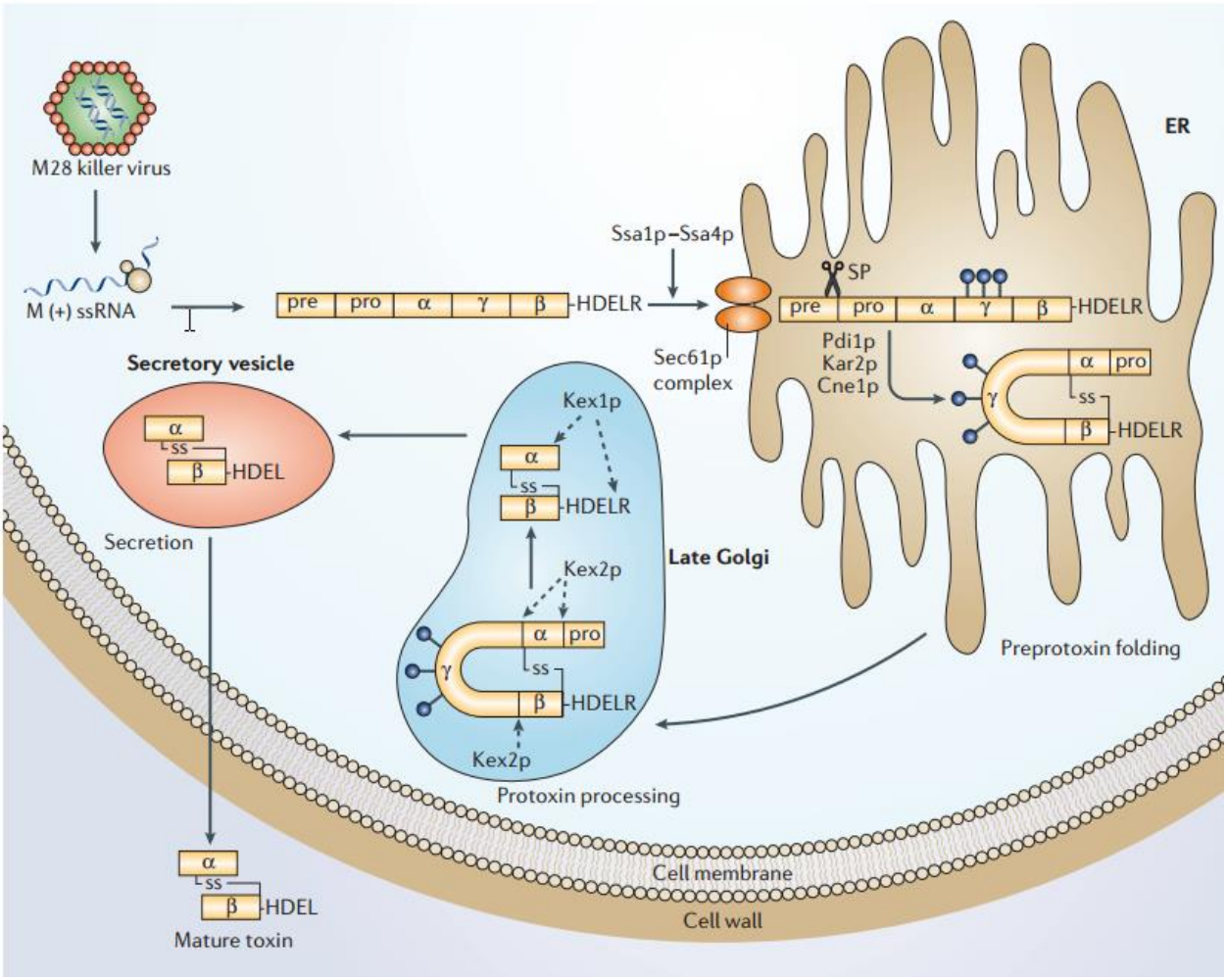


Figure 4; **Schematic of the processing undergone by killer toxin.** This process is shared between the three different yeast killer toxins. Figure taken from Schmitt et al[62]

The gamma, immunity, chain, undergoes 3 N-linked glycosylation events (residues 181, 203 and 216) and, whilst part of the pro-toxin, confers immunity against the toxin, by binding to the cytoplasmic side of the transmembrane protein which is targeted by the secreted toxin, blocking the toxins action.

Immunity to the toxin being co-expressed by the same gene as the toxin is a key factor of choosing KT. Whilst some toxins require a separate gene to encode an immunity protein, this

is not the case in KT, meaning a single cassette containing both toxin and immunity chain can be transformed into a given strain, and so toxicity from the product is not something we have to consider. This is not the case in all recombinant proteins, as there are examples of recombinant proteins that have toxic effects on the host system[23].

The complex post translational processing required by this protein makes it a good model protein for testing the secreting abilities of these strains, as the ability to make such modifications is one of the key reasons to use yeasts for recombinant protein secretion.

1.7.2 Mechanism of action for K1 Killer toxin

The secreted, K1 KT binds to the cell wall glucan receptor β-1,6-D[63], which transports the toxin inside the target cell. Once inside the cell, the toxin binds to the plasma membrane receptor Kre1p, and forms a cation-selective ion channel in the membrane, which leads to cell death. At higher concentrations (>10pM), the toxin may bind to Tok1p, and which in turn activates Yca1p and Dnm1p, which leads to mitochondrial ROS generation which eventually leads to cell death via Reactive oxygen species build up and stimulating apoptosis [63].

1.7.3 Killer toxin dsRNA virus

KT is the only endogenous protein that is naturally secreted by some strains of *S. cerevisiae* which contains complex PTMs, making it an ideal candidate for testing in the wild strains: it may be the case that some of these wild strains already have the toxin, and so have the machinery in place to secrete proteins with PTMs. Natively, K1 KT is encoded on a dsRNA virus which infects *S. cerevisiae* cells called M1, with the K2 and K28 toxin being encoded on the M2 and M28 viruses respectively. M1, M2 and M28 are satellite virus maintained by a second dsRNA virus called the L-A virus. the viral pair are coated by 60 76-kDa dimer proteins called Gag, and a 171-kDa fusion protein consisting of Gag and a second protein

Pol[62]. The viral RNA is replicated and a single stranded RNA is extruded into the cytoplasm. The Gag-Pol fusion protein binds to this RNA and acts as a nucleating point to form a new viral particle. In this way the virus replicates inside the cell, without causing cell lysis.

1.8 Aims of the project

By using the Halo assay to measure KT, a quantitative value for the secretion abilities of the wild strains can be generated. To do this technology to insert a cDNA 'gene' of K1 KT was developed. These technologies included a method of insertion into strains, lacking auxotrophic markers. The assay for KT was also developed, along with analysis of the secretome of transformed strains. A set of domestic strains were also assayed for their secretion levels, using both for KT, and Gaussian luciferase.

2. Materials and Methods

2.1 Yeast strains

Table 2.1 Summary of all standard strains of *S. cerevisiae* used

Strain name	**Genotype**	**Notes**
S6	*MATα his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0	Lab strain, Killer toxin sensitive, from Kent fungal group
Ski4 mutant	*MATα, ura3Δ0, ski4-1*[64]	Lab strain, Super Killer producing strain, from Kent fungal group
BY4742	*MATα his3*Δ1 *leu2*Δ0 *lys2*Δ0 *ura3*Δ0	Standard Lab strain, from Kent fungal group
3883	*MATa ura3Δ0 leu2Δ0 lys2Δ0 met15Δ0*	Modified wild-type isolate, European wine strain
3893	*MATa ura3Δ0 leu2Δ0 lys2Δ0 met15Δ0*	Modified wild-type isolate, West African strain
3903	*MATa ura3Δ0 leu2Δ0 lys2Δ0 met15Δ0*	Modified wild-type isolate, North American Strain
3913	*MATa ura3Δ0 leu2Δ0 lys2Δ0 met15Δ0*	Modified wild-type isolate, Sake strain
3923	*MATa ura3Δ0 leu2Δ0 lys2Δ0 met15Δ0*	Modified wild-type isolate, Malaysian Strain

Table 2.2 Summary of all experimental wild strains of *S. cerevisiae* used		
Strain name	**Geographic source**	**Biological source**
21	Ecuador, Los Rios	Beetle
27	Ecuador, Rumipamba forest	Grass
60	Ecuador, Podocarpus national park	Hemipteran insect
99	Slovakia, Bratislava	Danube river water
111	Slovakia, Bratislava	Plum tree leaf
114	Slovakia, Bratislava	Peach tree leaf
BJ20	China; Northern china, Secondary forest	Fruit
BJ6	China; Northern china, Secondary forest	Fruit
BT14	Unknown providence	Unknown providence
FJ7	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN10	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN11	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN14	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN15	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN16	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN2	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN6	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
HN9	China; Hainan, Primeval forest	Rotten wood and bark of fagaceae trees
SX1	China; Shaanxi province	Oak isolates
SX2	China; Shaanxi province	Oak isolates
SX6	China; Shaanxi province	Oak isolates

Strains were grown to stationary phase, (72 hrs in a shaking incubator at 30°C) and stored in suspensions of 500µl of 40% glycerol and 500µl of YEPD in sterile cryo-tubes and then stored at -80°C. Strains were then revived by inoculating into a liquid YEPD media, and grown at 30°C for 48 hrs.

NB: Throughout this thesis, the novel wt strains are collectively referred to as the 'Wild strains' whilst strains 3883, 3893, 3903, 3913, 3923 will be referred to as the 'domesticated strains'. Strains BY4742, Ski4 and S6 are standard laboratory strains

2.2 Plasmids used

Table 2.3 All plasmids used, including their notable feature(s) and source. Plasmid maps can be found in **appendix 1**

Plasmid name	**Main features**	**Selection**	**Source**
pVT100u	Killer toxin gene (*ADH1* promoter)	*URA3* (Yeast), Ampicillin Resistance (*E. coli*)	H. Bussy, McGill University, Montreal Canada
pML107	Cas9/gRNA plasmid	*LEU2* (Yeast), Ampicillin Resistance (*E. coli*)	John Wyrick Lab (Unpublished)
RHx702	Hygromycin B drug resistance gene	Hygromycin B drug resistance (Yeast), Ampicillin Resistance (*E. coli*)	Randal Hoffman
pUKC3546	Killer toxin gene (*ADH1* promoter), Hygromycin B drug resistance	Hygromycin B drug resistance, Amp.	Novel plasmid generated in this study: modified

	gene	Resistance (*E. coli*)	RHx702
pCG495	Gaussia luciferase (*ADH1* promoter)	*URA3* (Yeast), Ampicillin Resistance (*E. coli*)	Campell Gourly lab, Kent Fungal Group

2.3 Growing yeast and *Escherichia coli*

2.3.1 Yeast medium

Table 2.4 gives a summary of the medium used for growing yeast. All media was sterilised in an autoclave for 45 mins (Classic Media, prestige medical) before use. Where agar plates were required, 2% granulated agar was added before autoclaving. For auxotrophic selection, Synthetic Defined (SD) media was used, however, instead of synthetic complete mixture; 2% of the relevant mixture was substituted in (e.g., ura drop out mixture for –ura plates).

Table 2.4 Summary of all media used for yeast

Media	**Composition**
Synthetic Defined (SD)	2% glucose (G/0500/61, Fisher scientific) 0.67% Yeast nitrogen base (291940, BD) 2% synthetic complete mixture (DSCK1009, Formedium LTD)
YEPD	2% glucose (G/0500/61, Fisher scientific) 1% yeast extract (212750, BD) 2% bactopeptone (211677, BD)

2.3.2 *E.coli* medium

Table 2.5 gives a summary of the components in the medium used for growing *E.coli*. All media was sterilised in an autoclave before use. Where agar plates were required, 2% granulated agar was added. Ampicillin resistance was often used as a selective marker for transformations, in which case 100 μg/ml filter sterilised Ampicillin was added after autoclaving. For long term storage, a stock solution of 100mg/ml, suspended in distilled water was made

Table 2.5 Summary of all medium used for *E.coli*	
LB	1% tryptone (211705, Becton, Dickinson and company (BD)) 0.5% yeast extract (212750, BD) 1% NaCl (S/3160/60, Fisher scientific)

2.4 Molecular Biology techniques

2.4.1 Yeast transformations

Yeast transformation were done using a standard LiAc protocol[65], before being grown on agar plates with a different selection marker depending on the plasmid used. For this, cells were taken from a single colony on an existing agar plate and grown to log phase in YEPD liquid media for 24 hrs at 30°C in a shaking incubator. 2ml of culture was taken, and spun at 4000 rpm for 5 mins, and the supernatant removed via pipetting. The cells were resuspended in 240μl of 50% PEG, and transferred to sterile 1.5ml Eppendorf tubes. A master transformation mix was made up (**Table 2.6**), and 85.5 μl of this mix was added to each Eppendorf tube. The pellet was re-suspended. The cells were then incubated at 30°C for 45 mins, before being incubated for a further 45 mins at 42°C. Cells were then spun down

at 2000rpm for 5 mins, and the supernatant was removed. The cells were resuspended in 150μl of sterile water and plated onto appropriate selective agar plate.

Table 2.6; the components required for a single yeast transformation mix.	
Component	**Amount**
1M LiAc	36 μl
Sterile distilled water	34 μl
Single stranded DNA	10μl
Plasmid	3μl (300-600ng)
2-Mercaptoethanol	2.5ul

2.4.2 *E.coli* transformations

To increase the amount of any of the plasmids used , 20μl of Super Competent cells[66] from a -80 °C stock was taken and kept on ice. 200ng of the plasmid was then added. Initially, cells were left for 5 mins on ice after the addition of the DNA. The cells were then plated onto LB-amp plates, and grown at 37 °C for 24 hrs. This method, however produced very low efficiency (100 cfu/μg) when using any of the plasmids listed in **table 2.3.** To counter this, after the plasmid DNA was added, the cells were gently mixed and then left on ice for an hour. 5μl of LB was then added and the cell suspension was left for a further 2 hrs. This was enough to raises the efficiency to typically 10^9 cfu/μg.

2.4.3 Plasmid purification from *E.coli*

Successfully transformed E.coli cells were inoculated into 20ml of LB media and incubated overnight at 37° C in a shaking incubator. The Qiagen QIAprep Spin Miniprep Kit was used to extract and purify DNA of interest.

2.4.3 PCR

2.4.3.1 Primer design, and PCR conditions

PCR was done using pairs of oligonucleotide primers (**table 2.7**) designed with the online tool 'Primer3'[67] with any required cut sites added in as necessary. A Techne TC-312 PCR Thermal Cycler was used for the amplification. The annealing temperature and extension time of each reaction was demined by the primers and products respectively. The annealing temperature was set at 5°C below the melting temperature of the section of the primers homologues to the plasmid; this does not include any added cut sites. The extension time was calculated at 1 min per kB of the final product.

Table 2.7 The sequences used in PCR as primers

Primer name	Sequence
KT_clone_pci1_f	ACATGTTATGGAAAAACGCCAGCAAC
KT_clone_xho1_r	CTCGAGCAAAAGCGCTCTGAAGTTCC
CRISPR_RT_P1a	CAACGACCGCAAAAAGGTATTTAAATATACCACAACCGGGAAAAGATTTGATTGCATTTTCT GCCAACAACTTCGCCTTAAGTTGAACGGAGTCCGGAACTCT
CRISPR_RT_P1b	CAAAAGCGCTCTGAAGTTCC
CRISPR_RT_P2a	CAAAAGCGCTCTGAAGTTCCTGGTGTGAAAATGTTCTATTTAGAAACAGGGGAATTGCTTAT TAACGAAATTGCCCCAAGGCCTCACAACTCTGGACATTATACCATT
CRISPR_RT_P2b	CAACGACCGCAAAAAGGTAT

2.4.3.2 Killer toxin cloning PCR

KT_clone_pci1_f and KT_clone_xho1_r are primer pairs cloning the killer toxin gene into RHx702. Both of these primers consisted of a stretch of 80 bases that pair to the region either side of the cut site in the yeast genome(marked in red in **table 2.7**), as well as adding a 6 base pair cut site (*Pci*1 and *Xho*1 respectively), to be added (marked in green in **table 2.7**). For this reaction, 100 ng of pvt100, 5μl of 10x taq reaction buffer, 200 μM dNTPs, 0.2uM of both KT_clone_pci1_f and KT_clone_xho1_r, were added to 1.25 units of taq polymerase. The mix was then made up to 50μl with sterile distilled water. PCR conditions were as follows:

Stage	Temperature	Time (sec)
Initial denaturation	95°C	300
30 cycles of:		
Denaturation	95.0°C	30
Annealing	53.4°C	45
Extension	68.0°C	210
Final extention	68.0°C	600

To check the size of, and purify the PCR product, the completed reaction was run on an agarose gel (Biolaboratories LTD , batch number :F25073). The gel was run at 75 volts for 45 mins, and used ethidium bromide (final concentration 0.5 μg/ml) to visualise the DNA, on a transilluminator. Once the band was compared in size to a marker, and confirmed to be of the size predicted by Primer3, the band containing the product was cut out, and gel purified

2.4.3.3 Generation of CRISPR repair template with PCR

The repair templates used in the CRIPSR genome editing, were generated by PCR using the two primer pairs CRISPR_RT_P1a/ CRISPR_RT_P1b and CRISPR_RT_P2a/ CRISPR_RT_P2b.The sequence for these primers can be found in **Table 2.7**, with the region that pairs with the KT-encoding sequence of PVT100 coloured in blue. The target sequence is highlighted in red. The region highlighted in green is an inserted stop codon. See **Figure 2.1** for a diagram of how the primers paired to both the genome and the plasmid. The mix for both sets of PCR were 100 ng of pvt100, 5µl of10x taq reaction buffer, 200 µM dNTPs, 0.2uM of both forward (CRISPR_RT_P1a and CRISPR_RT_P2a) and reverse primers(CRISPR_RT_P1b and CRISPR_RT_P2b) , were added to 1.25 units of Taq polymerase before being made up to 50µl with autoclaved distilled water. The conditions for each PCR were as follows:

Stage	Temperature	Time (sec)
Initial denaturation	95°C	300
30 cycles of:		
Denaturation	95.0°C	30
Annealing	53.4°C	45
Extension	68.0°C	210
Final extention	68.0°C	600

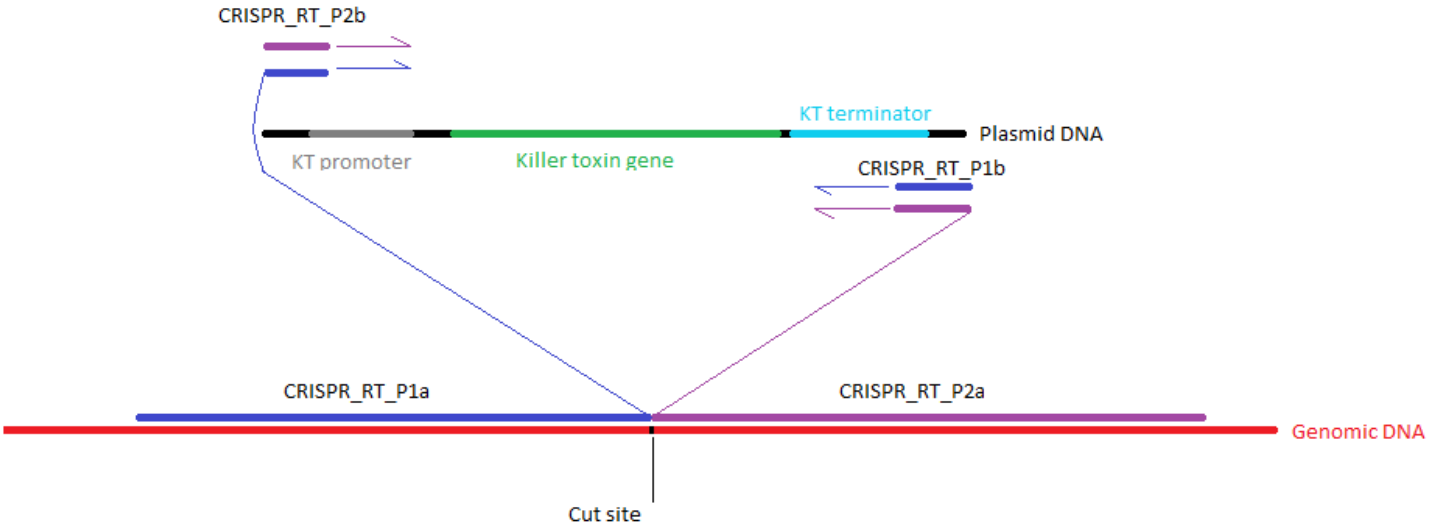


Figure 2.1; **Generation of CRISPR repair template with PCR.** Primer map in terms of homology to both the killer toxin gene (in this case, in pVT100u) and the target loci in the yeast genome. Primer pairs have the same colour .Primers P2a and P1a have homology to both pVT100u and to the target in the genome. Primers p1a and p1b leads to the generation of a repair template with an inserted stop codon: once integrated into genome, this will produce a truncated AIR carboxylase (the product of the *ADE2* gene) protein.

2.4.4 CRISPR target site and plasmid

When designing the CRISPR experiment, the target locus for gene insertion was identified using 'CHOPCHOP'[68]. CHOPCHOP outputs the best guide RNA to use to target specific loci.

The guide RNA oligo was designed with a pair of restriction enzyme cut sites (*Bcl*I/*Swa*I) which allowed insertion into the pML107 plasmid at the correct locus. Both the plasmid and the primers were digested in the same way: *Swa*I was used first, incubated at 25°C for 1 hr, before being heat inactivated at 65°C for 40 mins. *Bcl*I was then added and incubated at 50°C for a further 1 hr. The cut fragment was then run on an agarose gel (75 volts for 45 mins, and used ethidium bromide (final concentration 0.5 μg/ml) to visualise the DNA, on a transilluminator), and gel purified using a GeneJET Gel Extraction Kit (ThermoScientific). The

two fragments were then ligated with 1 Weiss units of T4 ligase, in a molar ratio of 3:1 inert to vector, made up to 30μl with water. The reaction was incubated for 16 hrs at 25°C.

To generate the repair template, PCR using the primer pairs p1a/p1b and p2a/p2b were separately run, using pVT100 as a template, cloning of the killer toxin region. Once generated, the repair template and the Cas9/gRNA plasmid were transformed into the 5 wild strains, which were then plated out on SD plates deficient in leucine, and grown overnight at 30°C

2.4.5 Agaroses gel electrophoresis

During this project, agaroses gel electrophoresis was used for a range of reasons (PCR product checking, Ligation/restriction digest checking, etc.). Whilst the reasons for its use may differ the underlying method remained the same.

A 1-2% w/v (with concentrations near 2% being used for short DNA sequences, and nearer 1% for larger sequences) solution of agarose and TBE buffer (10 M EDTA, 0.04 M Tris-acetic acid [v/v]) was prepared. Typically a final volume of 50ml was used. This solution was heated in a microwave, until the agarose was completely dissolved (typically 30-45 seconds), then the solution was left to cool to 50 °C. Ethidium bromide was then added to a final concentration of approximately 0.2-0.5μg/mL, before the gel was poured into a gel cast, with a comb to form the wells. Bubbles were popped with a clean metal spatula, to ensure a smooth, even gel.

Once the gel had set, it was submerged in TBE, until completely covered. Samples were then loaded into the wells, along with an appropriate marker (in this project, GeneRuler 1 kb Plus DNA Ladder from ThermoFisher scientific was used) and the gel was run at 95 volts for one

hour, or until the dye front had reached the bottom of the gel. The DNA in the gel was then visualised using a UV- transilluminator.

2.4.6 Gel purification

Once useful fragments were identified (generally via comparing the size to an expected value, e.g. plasmid size), the DNA was purified from the gel. This was done using a QIAquick Gel Extraction Kit, following the protocol included with the kit. This typically produced a DNA concentration of ~ 50-300ng/µl, (depending on the initial concentration run on the gel) suspended in elution buffer. This concentration was tested on a NanoDrop microvolume UV-Vis fluorospectrometer(NanoDrop 3300, Thermo Scientfic).

2.4.7 Restriction digests

To digest a plasmid, the correct restriction enzyme was identified using serial cloner. Then a mixture was made up as following: 1 µg DNA,1 µL of each Restriction Enzyme, 3 µL 10x digest buffer, up to 30µL with distilled H_2O.

The buffer used was one of the NEBuffers depending on the enzymes being used. This was identified using the New England Biosciences web tools. The mix was then incubated at room temperature (25 °C) overnight (8-18 hours). Enzymes were heat inactivated at 65°C. To check for successful digestion, the mix was run on an agarose gel, along with a sample of intact DNA as a control. Gel was then inspected using a UV- transilluminator.

2.4.8 Ligation

To insert fragments into a vector, T4 DNA ligation was used. For this, a mix was made on ice of a ~1:3 mass ratio of vector to insert, along with 2 µl T4 DNA Ligase Buffer (10X), 1 µl of T4 DNA Ligase. The T$ ligase was added last, and kept at -20 until needed, before being quickly returned to the -20 freezer The solution was made up to 20 µl with dH_2O. Mix was

incubated at room temperature (25 °C) overnight (8-18 hours). Ligation was check via gel electrophoresis, using an unligated sample as a control.

2.5 Assays for secreted proteins

2.5.1 Halo assay for Killer toxin

To quantitate the level of KT secreted from a given producing strain, a 'halo' assay was used. This assay works on the susceptibility of a sensitive strain of *S. cerevisiae* (in this case the S6 Strain; **Table 2.1**); 25ml of a top agar (a full break down of components is listed in **Table 2.8**) containing the $5x10^6$ cells/ ml in mid log phase of a sensitive strain was poured over 100 ml of a base (a full break down of components is listed in **Table 2.8**) containing 0.15M Methylene blue.

Initially a 'cleared media' approach was tried: test strains were grown in an SD selective media, for 12 hours. The OD_{600} was measured, and each test strains was diluted down to OD_{600} = 0.8 . The media was then cleared of cells via centrifugation (4000rpm, 5 mins) and 15ul of this cleared media was pipetted into 6mm filter disks, and incubated for 3 days at 25°C. This method however, produced inconsistent results, with halos often being very small. In an attempt to produce more consistent results, an alternative 'washed cell' method was tried; 1 ml of the test strain (at OD_{600} = 0.8, i.e. log growth phase) was taken, as before, spun down and resuspended in 10μl of ddH_2O. This suspension was then pipetted onto 6mm filter disks and incubated at 25°C for 3 days (**figure 2.2**). This method produced more consistent results, so was used for the remaining assays.

During the 3 day incubation of the washed cell method, the tests strains secrete KT, which causes the cell from the sensitive strain to die. The methylene blue is then taken up by the dying cells, producing a dark blue ring surrounding the zone of clearing. Plates were then scanned and the image analysed with "imageJ", a Java-based program which allows the size of components of an image to be analysed and compared. For each test plate, untransformed BY4742 and SKI4 strains (**Table 2.1**) were used as negative and positive controls respectively. The recipe for the agar used can be found in **table 2.8**

Figure 2.2; construction of the zone of clearing assay plates. Base agar is mixed with Methylene blue, and the top agar is doped with S6 sensitive yeast.Dyijng cells in the top layer absorbed the blue from the bottom, leading to the formation of hales around the test strains on the filter disks. 6mm disks were used, with 10µl of OD_{600} 0.8 test strain.

Table 2.8 The two agars used in the zone of clearing assay	
Zone of clearing assay base agar (pH 4.6 - 4.8)	0.5% yeast extract (212750, BD) 0.5% bactopeptone (211677, BD) 1.5% Citric Acid(251275-500G, Sigma-Aldrich) 1.5% agar (214530, BD) 2% glucose (G/0500/61, Fisher scientific) 0.625% 0.15M Methylene blue(added after autoclaving, filter sterilised)

	1.42% Na_2HPO_4 (S3264-500G)
Zone of clearing assay top agar (pH 4.6 - 4.8)	0.5% yeast extract (212750, BD) 0.5% bactopeptone (211677, BD) 1.5% Citric Acid (251275-500G, Sigma-Aldrich) 0.75% agar (214530, BD) 2% glucose (G/0500/61, Fisher scientific) 1.42% Na_2HPO_4 (S3264-500G, Sigma-Aldrich) $5x10^6$ cells /ml S6 (added after autoclaving, and cooling to 45 °C.)

2.5.2 Gaussia Luciferase luminesce assay

Once strains were transformed with plasmid pCG495 (**Table 2.3**) they were grown overnight at 30°C. A Pierce Gaussia Luciferase Glow Assay Kit (Catalogue number: 16160) was then used. Protocol was followed as included in the kit. A 96 well plate was used (**Figure 2.3**); 20 μl of cell cleared media was taken and added to 50 μl Working Solution (100:1 Gaussia Luciferase Glow Assay Buffer: 100X Coelenterazine) in each test well, and was then incubated at 25°C for 10 mins.

Once the substrate (coelenterazine) was added, florescence readings at 482 nm (the peak emission of the catalysed reaction Gaussia Luciferase) were taken in a plate reader (FLUOstar Omega Microplate Reader - BMG LABTECH) every 10 mins for 1 hr. Values were exported to Excel for further analysis; repeat runs were averaged, and standard error was calculated.

	1	2	3	4	5	6	7	8	9	10	11	12
A	T1	T1	T1	T1	T1	T1	B	B	B	B	B	B
B	T2	T2	T2	T2	T2	T2	B	B	B	B	B	B
C	T3	T3	T3	T3	T3	T3	B	B	B	B	B	B
D	T4	T4	T4	T4	T4	T4	B	B	B	B	B	B
E	T5	T5	T5	T5	T5	T5	B	B	B	B	B	B
F	Cp	Cp	Cp	Cp	Cp	Cp	B	B	B	B	B	B
G	Cn	Cn	Cn	Cn	Cn	Cn	B	B	B	B	B	B
H	B	B	B	B	B	B	B	B	B	B	B	B

Figure 2.3; layout of the 96 well plates for Gaussia Luciferase assay. T1-5 was each test strain (in order: 3903, 3913, 3923, 3883, and 3893). Cp was the positive control: BY4742 transformed with the plasmid. Cn was the negative control: BY4742 with no plasmid. B represents blank wells which remained empty.

2.5.3 Hygromycin B resistance assay

Hygromycin B is a anti-biotic which is effective against both prokaryotic and eukaryotic cells functions to stop translation of proteins by binding to the S30 ribosomal sub unit. The residue it binds to (H44), has been implicated in the movement of tRNA from the A site to the P-site, which in turn cases a stalling of protein production.[69]

To determine the level of Hygromycin B that should be used when selecting for transformation by the plasmid pUKC3546 (**Table 2.3**) each strain was grown on 3 different YEPD plates, each with a different concentration of Hygromycin B added after autoclaving; 200 ug/ml, 1000 ug/ml, 2000 ug/ml. Each strain was grown overnight, and then diluted to OD_{600} 0.01. 10μl of each strain was then plated out, and grown for 72 hrs at 30 °C.

Some strains had low numbers of colonies growing on 2000 μg/ml. To test the ability of these strains to produces resistant colonies, 10^5 cells (determined using a haemocytometer) were suspended in 10μl YEPD and plated out onto a 2000 μg/ml hygromycin B plate, and grown for 6 days. The amount of colonies that displayed resistance was then scored.

2.5.4 Plasmid stability assay

To test the stability of pvt100u within yeast strains, an assay was developed where transformed strains would be grown in non-selective liquid media (YEPD) for 24 hrs at 30°C. 10µl of media were taken and plated out onto both SD –ura and SD agar plates. Plates were incubated for 3 days at 30°C. Colonies for an individual strain, on both plates were counted and compared.

2.6 secretome analysis

To check the secretome of the domesticated strains, both untransformed and transformed with pvt100, the strains were grown in YEPD liquid media for 72 hrs at 30°C in a shaking incubator. The culture was then centrifuged at 14000 rpm for 15 mins, and 30 ml of culture was taken, before undergoing TCA protein precipitation and was then run on an SDS Tris-Tricine gel.

Table 2.9: Composition of 40ml 1x sample buffer

Component	**amount**
1 M Tris-HCl pH 6.8	2.5 ml
SDS	1.0 g
0.1% Bromophenol Blue	0.8 ml
100% glycerol	4 ml
14.3 M 2-mercaptoethanol (100% stock)	2 ml
H_2O	up to 40ml

2.6.1 TCA protein precipitation

30ml growth culture was taken and cleared of cells by 10 min of centrifugation at 4000 rpm in sterile 50 ml falcon tubes. 25% by volume (7.5ml in the case of 30 ml cultured media) of 100% w/v TCA was added. This solution was incubated for 60 mins at 4°C and then tubes were spun at 14000 rpm for 60 mins at 4°C. Supernatant was removed, leaving the white protein pellet intact. The pellet was then washed with 500µl of -20°C acetone. Tubes were then spun down at 14000 rpm for 5 mins, and the acetone removed via pipetting. This wash was then repeated. Once acetone had been removed again, the pellet was re-suspended in 50µl of 1x sample buffer (**table 2.9**), and transferred to sterile 1.5ml Eppendorf tubes. Samples were then boiled at 95°C for 15 mins before analysis by SDS-PAGE.

2.6.2 SDS PAGE analysis

Protein samples were loaded onto a Tris–Tricine gel (**Table 2.10**). The internal cavity of the gel tank was filled with a cathode buffer, and the exterior cavity with an anode buffer. The 1X cathode buffer was 100mM Tris, 100mM Tricine, and 0.1%SDS, at pH 8.25, stored as a 10X solution at room temperature. The 1X anode buffer consisted of 210mM Tris pH8.9, stored as a 10X solution at room temperature. The gels were run for 18 hrs at 25 volts and 30 milliamps. Coomassie brilliant blue was used to stain for an hour, and strong SDS distain (50% distilled water, 40% methanol, 10% glacial acetic acid) was applied until the gel underwent sufficient de-staining to produce visible bands (typically 6-12 hrs).

Table 2.10: Composition of the Tris–Tricine polyacrylamide gel (makes 2 gels)		
SOLUTIONS	**Separating gel**	**Stacking gel**
Acrylamide Solution (40% Acrylamide (36.5:1))	10 ml	1 ml
Gel Buffer (3 M TrisHCl pH8.45 + 0.3% SDS)	10 ml	3 ml
Glycerol	3 g	0 g
H_2O	Up to 30ml	Up to 12 ml
10% APS	100 µl	90 µl
TEMED	1.0 µl	9 µl

2.6.3 Lysed cell analysis

To check the intracellular protein components, both untransformed, and strains transformed with pvt100, were grown in YEPD liquid media for 24 hrs at 30°C in a shaking incubator. This may lead to a loss of plasmid in a transformed strained; however, the plasmid retention assay had already shown that a 24 hr growth in a non-selective media had little to no effect on pvt100u retention (**section 3.1.2**). 50ml of culture was taken and spun down at 4000rpm for 5 mins at 4°C. The liquid medium was poured off and the cells were resuspended in 12ml sterile distilled water. The cells were then spun down at 4000rpm for 5 mins, before being resuspended in 12ml of 1M sorbitol. The cells were once again spun down at 4000rpm for 5 mins, before being resuspended in 12ml SCE buffer (1 M sorbitol 100mM sodium citrate, 10mM EDTA (pH 8.0),0.125% β-mercaptoethanol). 120µl of 1M DTT and 60µl of 10,000 units/ml lyticase (SigmaAldrich) were then added, and the solution incubated for 1 hr at 30°C in a shaking incubator.

The tubes were spun down at 1500 rpm for 10 mins, and the supernatant removed. The resultant pellet of spheroplasts were resuspended in 2ml 20% SDS, and incubated for a further 10 mins at room temperature. The solution was then spun down at 4000 rpm and the supernatant taken and underwent TCA protein precipitation (**section 2.6.1**), and was run on an SDS Tris-Tricine gel (**Section 2.6.2**).

2.7 Computational analysis: ImageJ

To analyses the data from the plates generated in the halo assays (**Section 2.5.1**), the Java-based image analysing program ImageJ[70] was used. ImageJ allows users to measure the size of various features of a given image; by measuring the size of the halo and expressing it as a percentage of the size of the plate on the image, a measurement of the size of the halo could be made. The next step was to then measure the size of the actual plate, and using the percentage from before to calculate the size of the zones of clearing. By following this method, consistent measurements were made, and most importantly, measurements between different plates could be compared.

3. Results

3.1 Domestic strains: transformation and halo assay

To optimise the halo assay for KT secretion, the five domesticated strains were used as well as BY4742 and Ski4. They were first transformed with pVT100u using the LiAc method (**section 2.4.1**) and then plated onto the halo assay plates. The stability of this transformation was also tested, to check if the plasmid would be lost over time, if grown on non-selective media.

3.1.1 Successful transformation using the LiAc method

Strains were transformed with at least 200ng of pVT100u, following the method outlined in section 2. The selection plate used was a SD plate without uracil; this was because the domestic strains had the *URA3□0* genotype, meaning they could not produce uracil. If grown on or in a medium which lacked uracil these strains would die, however pVT100u contained the *URA3* gene, so any cells successfully transformed would be able to grow, whilst any cells that did not take up the plasmid would die. This method typically generated approximately $2x10^4$ successfully transformed cells per plate, which equates to approximately $1x10^5$ cfu/ug of plasmid DNA.

3.1.2 pVT100 plasmid retention assay

Since yeast strains grow much quicker in YEPD rather than SD -Ura , an experiment was performed to test if the strains would maintain pVT100u after a 24 hour growth in YEPD media (ie. a non-selective media). Each strain was tested in triplicate.

Figure 3.1 and **Figure 3.2** show the results of the plasmid retention assay. The clear trend across all strains is that, for a single 48 hour growth period in a non-selective environment, there is no significant loss of pVT100. All strains have over lapping error bars, meaning there is no statistically significant difference between the two sample sets. Note that strain 3893 show a lot less CFU from this experiment, but other tests did not continue this trend

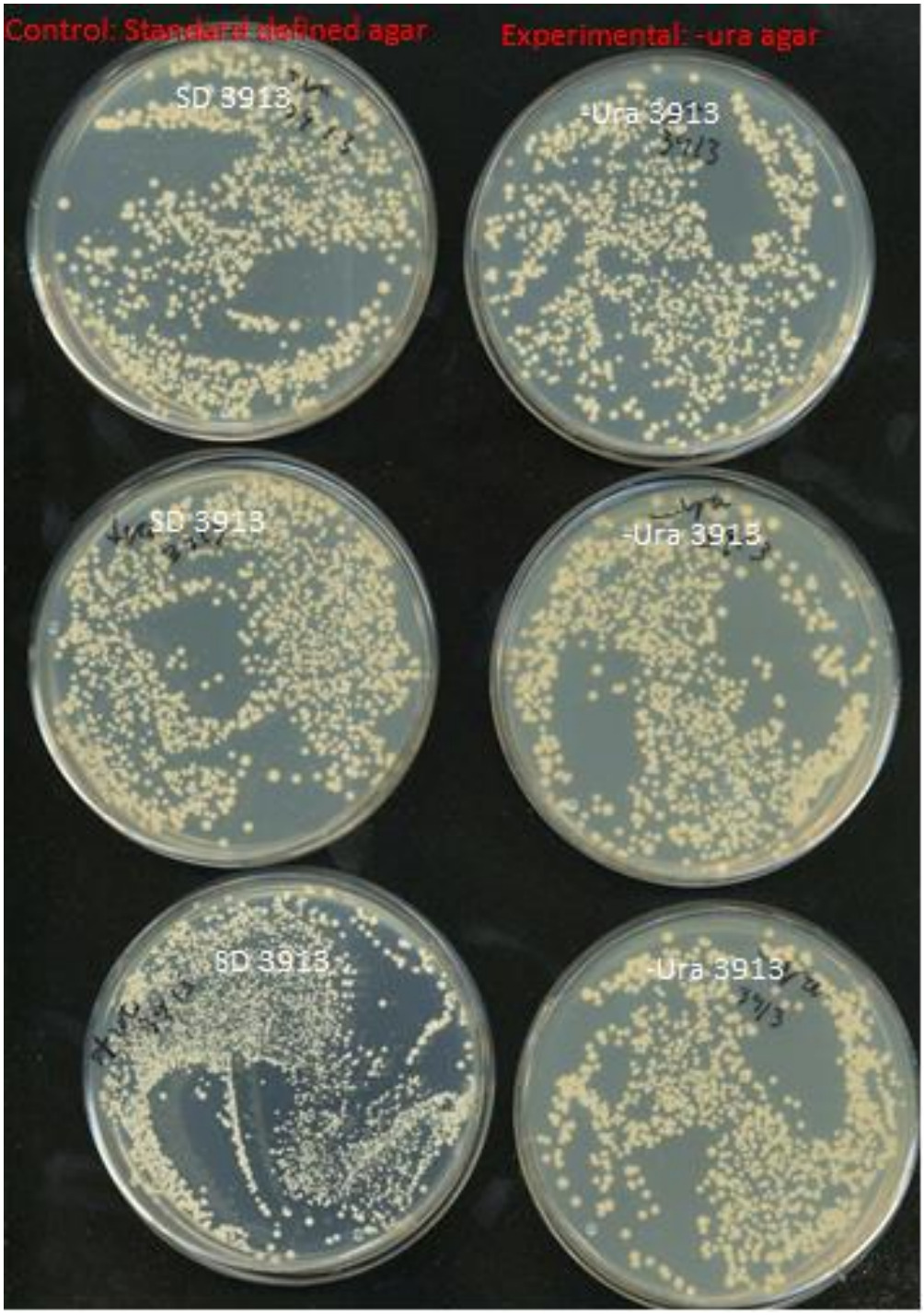


Figure 3.1 Set of results for the plasmid retention assay, containing 3 repeats for strain 3913. The left column is SD media, and acted as a control, whilst the right column is the experimental plates, containing no uracil.1000 cells were plated onto each plate (measured using a haemocytometer), and If both plates grew similar numbers of colonies, then the plasmid would have been retained, allowing the use of YPD as a temporary growth medium. The plates above were typical results of the whole experiment.

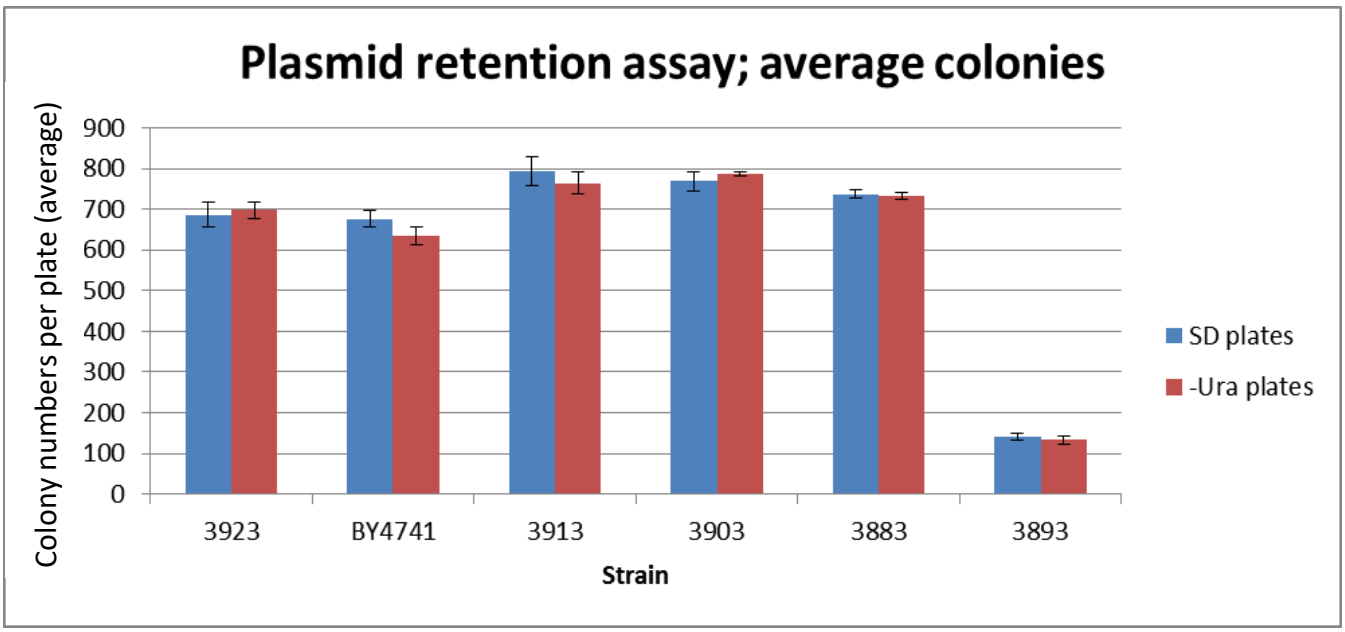


Figure 3.2 Plasmid retention assay results; Graph of the data from the colony count (Figure 3.1). Data is expressed as the number of separate individual colonies on each plate, averaged across each set. Each column is labelled to the corresponding strain name, and error bars shown are +/- 1 standard deviation, across three repeats. Plasmid is retained in all strains: control and test plates for each experiment show the same level of colonies, within the error bars.

3.1.3 Domestic strain halo assay

Once the washed cell method for the halo assays was decided on, rather than the cleared media method, the transformed domestic strains were used in the halo assay (see section 2 for full details) along with BY4742 and Ski4 as positive and negative controls. The graph of these results can be found in **Figure 3.3.** Each experiment was done in triplicate, and error bars are +/- 1 standard deviation.

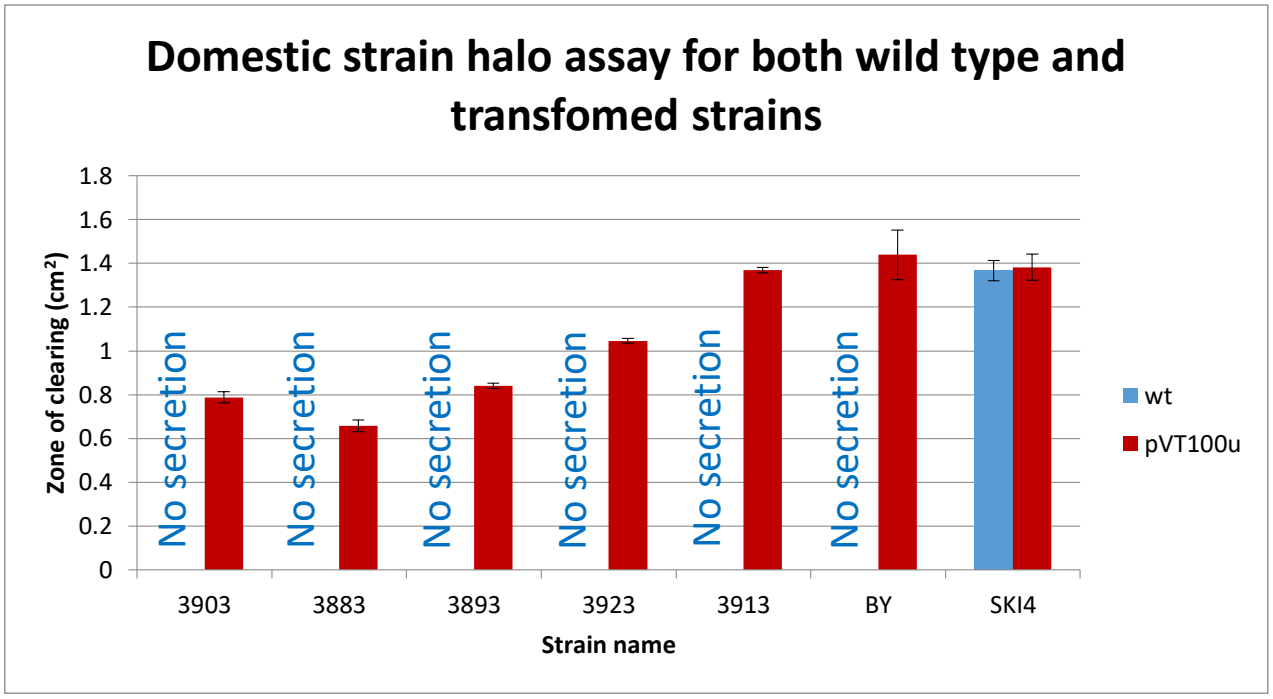


Figure 3.3 the results from the halo assay using the domesticated strains, with Ski4 and BY as controls. Blue bars represent the untransformed strains and red bars represent the strains transformed with pVT100u, expressed as the area of the zone of clearing each strain produced after 72hrs incubation. Whilst data for all strains was collected, only Ski4 showed secretion of killer toxin when untransformed. Each column is labelled to the corresponding strain name or number.

3.3 Hygromycin B Resistance assays

Hygromycin B was chosen as a selective marker due to its toxic effect on yeast; Hygromycin B binds to the mRNA decoding centre in the small (30S) ribosomal subunit of the 70S ribosome. This induces a conformational change which stops bound tRNA from progressing from the A binding site to the P site, which leads to a staling of protein synthesis.

However, there have been examples of yeast displaying hygromycin B resistance. To check the resistance of strains used in the current project, a resistance assay was performed**.**

Figure 3.4 shows the results of the hygromycin B resistance assay. The outermost column on the left and right sides were spotted with 10µl media containing around 10^6 cells. Each dilution was at 10X meaning that the centre most columns of spots initially contained 10^3 cells each. It should be noted that the domesticated strains were not used in this assay; disruption of the *HO* gene common to these strains renders them resistant to the effects of

the drug. All strains grew well on the medium containing 200μg/ml though there was slight variation in growth when 10^3 were initially put onto the plate. All strains were then plated onto the 1000μg/ml plates. This is where a clear difference emerged, with many strains (HN16, HN10, HN1, SX1, BJ20, 21, HN9, HN11, BJ6, and HN15) showing no growth after 72 hrs.

Interestingly, some strains showed slight resistance across all concentrations, but still showed retarded growth compared to the previous test plate (BT14, FJ7, 99, 60, SX2 SX6 and 114). Strain BY4742 showed a low level of resistance that was borne out in all 3 repeats, each from a different overnight culture, this resistance only showed at the lowest hygromycin B concentration (200mg/ml).

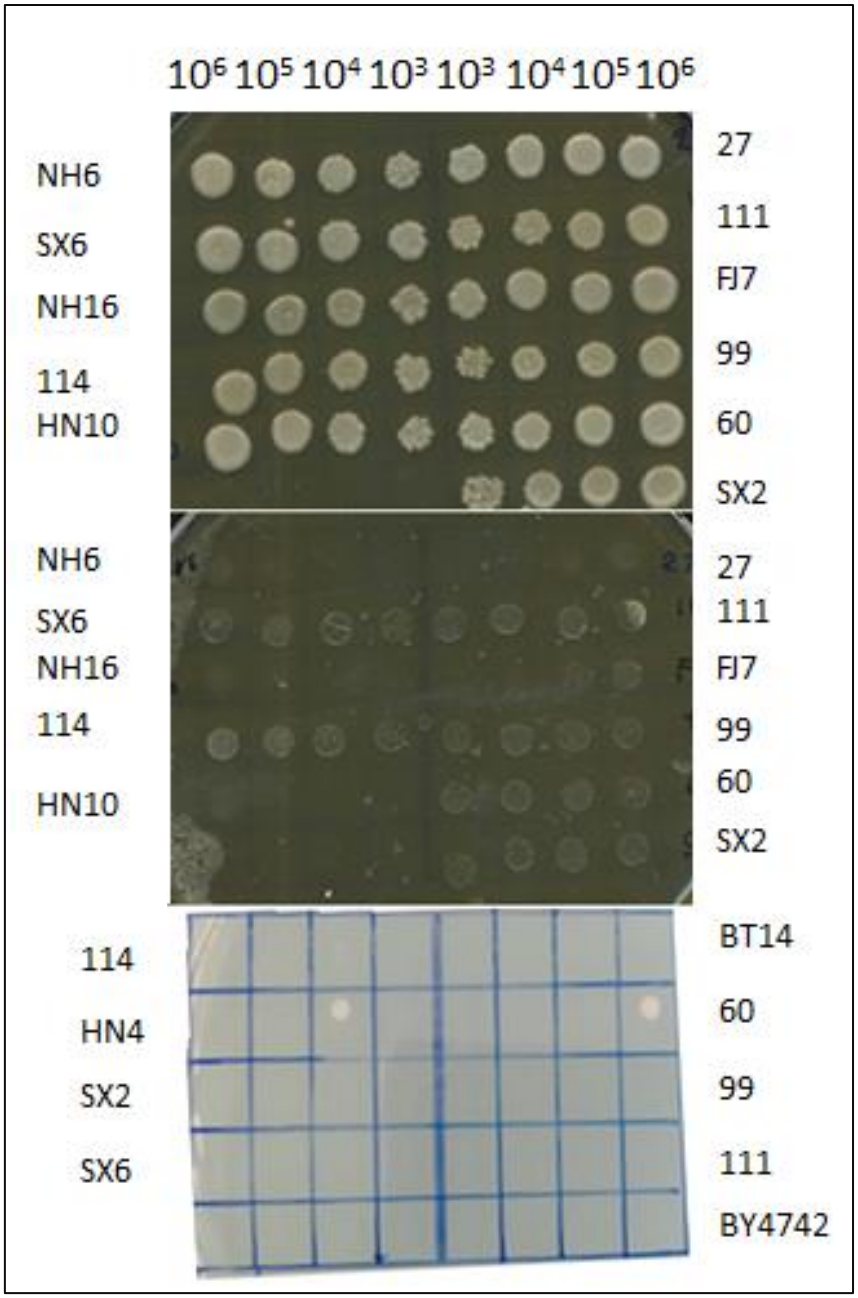


Figure 3.4: Serial dilutions of each untransformed strain plated onto a varying concentration of hygromycin B. The concentration of hygromycin B increases down the figure: top row; 200μg/ml, middle row 1000μg/ml bottom row; 2000μg/ml. The number of cells spotted in each column is labelled at the top, with strain name at the side. The bottom left row of each plate is empty.

Table 3.1 Summary of hygromycin B resistance assay.

Strain name	Highest concentration to show high growth (mg/ml)	Concentration used for transformation selection (mg/ml)
21	200	1000
27	200	1000
BJ20	200	1000
BJ6	200	1000
FJ7	200	1000
HN10	200	1000
HN11	200	1000
HN15	200	1000
HN16	200	1000
HN2	200	1000
HN6	200	1000
HN9	200	1000
SX1	200	1000
99	1000	2000
111	1000	2000
114	1000	2000
BT14	1000	2000
SX2	1000	2000
SX6	1000	2000
60	2000	2000
HN14	2000	2000

3.4 pUKC3564: plasmid construction and yeast transformation

To insert the KT gene into the wild strains, a plasmid was required that contained both a hygromycin B resistance gene and the KT gene. To this end, pUKC3564 was designed and created from two existing plasmids; pVT100u and RHx107.

3.4.1 Plasmid construction

To allow plasmid selection in yeast strains without built in auxotrophic mutations, such as *leu2Δ*0, a plasmid that had an alternative, positive selection marker was required, in this case a plasmid which conferred resistance to the drug hygromycin B. It should be noted that as part of their modification, strains 3883, 3893, 3903, 3913 and 3923 are resistant to hygromycin B, due to the knockout of one allele of the *HO* gene, full details of which can be found in Louvel *et al*[71]. This was confirmed by spotting out these strains on agar plates containing 2000mg/ml of hygromycin B.

The strategy used to construct a hygromycin B /Killer Toxin plasmid was to clone the whole K1 Killer Toxin gene (an artificial cDNA gene constructed from the RNA virus) from the pVT100 plasmid using PCR, and insert the fragment into RHx702, which contained a hygromycin B resistance gene (**Appendix 1b**). The primers for this were positioned so they also captured the promoter and terminator. In addition to cloning these genetic elements, each primer contained a restriction enzyme site: *Pci*I on the forward primer and *Xho*I on the reverse primer. Both the PCR product and RHx702 were digested overnight with *PciI* and *XhoI*. When separated on an agaroses gel the digested RHx702 separated into a 6085bp and a 2759bp fragment, which was consistent with what would be expected of a successful digest with these *PciI* and *XhoI*.

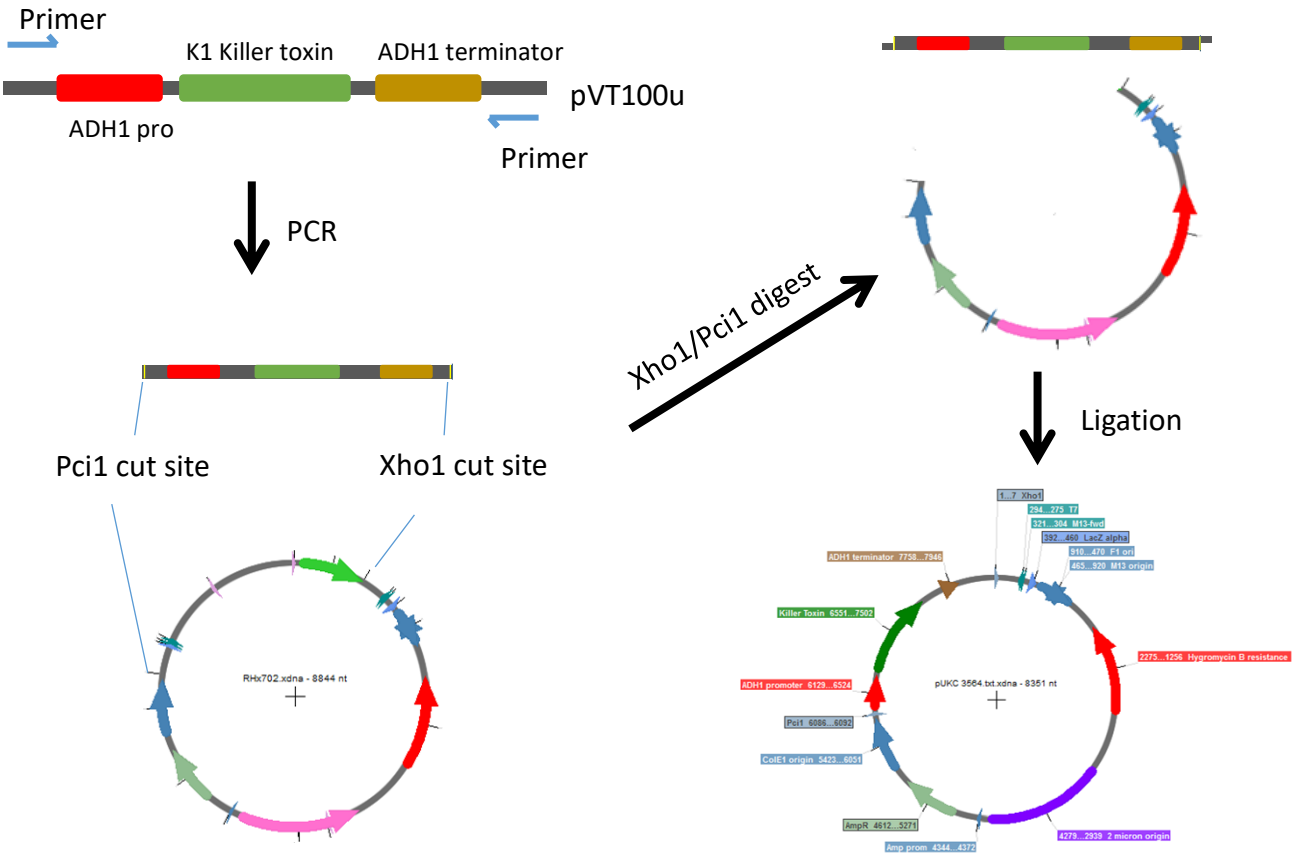


Figure 3.5: Summary Process of creation of pUKC3564. Final plasmid contained both killer toxin gene, and hygromycin B resistance. It also contained an ampicillin resistance gene for *E.coli* transformation. First the killer toxin in pVT100u was cloned out using PCR, with the promoter and terminators as noted. Then both the base plasmid (RHx702) and the PCR product were digested with *Pci1* and *Xho1*in a restriction digest. Then both the vector and insert we ligated together, and the final plasmid was bulked up in *E.coli.*

Once the PCR product and RHx702 had been digested, they were ligated using a Roche T4 DNA Ligase kit (1X reaction buffer, 1:1 approximate molar ratio of gene to plasmid (calculated as a 1:1 of weight added in Nano grams/size in bp. For example 3 fold more weight in mass would be required for an insert that is 3 fold smaller than the plasmid), 1ul T4 ligase left at 20°C for 16hr overnight, **Figure 3.5**). Product size was checked via gel

electrophoresis to confirm successful ligation, before being transformed into super-competent *E. coli* for bulking up. A full map of pUKC3564 can be seen in **Figure 3.6**, which includes both the KT gene, and the hygromycin B resistance gene. This map was produced in seral cloner by repeating the experiment *in-silico.*

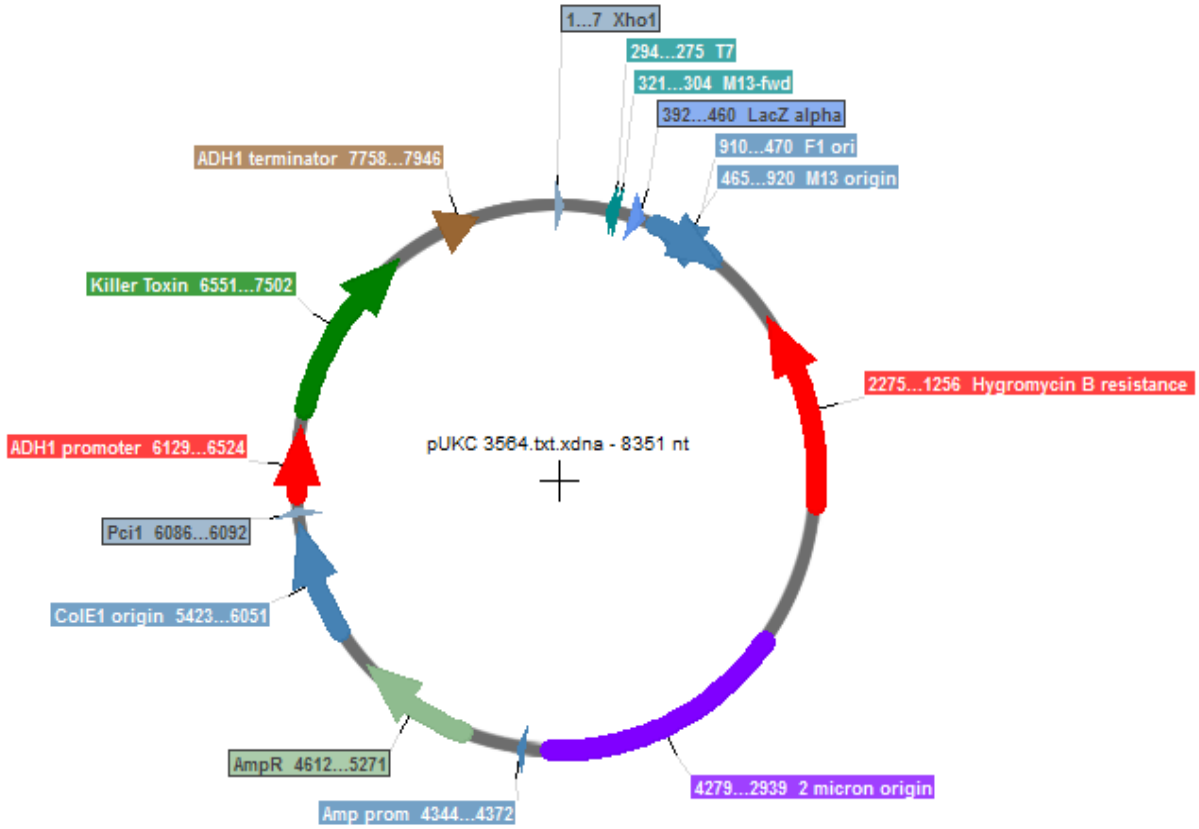


Figure 3.6: pUCK3564 plasmid map, labelled with relevant features. Figured produced in Serial cloner[98]

3.2.2 Transformation of wild strains of yeast

Once constructed, attempts were made to transform pUKC3564 into all wild strains (see **section 2.4.1** for methodology). It was decided to use different levels of hygromycin B for different strains, based on the results from the hygromycin B resistance assay (**table 3.1**)

However, after 48 hours, all plates transformed showed either none or very low levels of growth (1-2 small colonies). The plates that did show colonies took up to a further 48 hrs before the colonies were of a size to be easily picked for growth in liquid media. Many strains simply could not be transformed even after a variety of different attempts; selection plates showed no growth even after 3 days incubation at 30°C .The lack of transformers was seen across multiple repeats, and under different transformation conditions: including using longer incubation times (overnight), and longer heat shocks (up to 50 mins). Other possibilities are discussed in the discussion and further works section. It would appear that this intractability to transformation is a trait inherent to these strains; another lab working in parallel with these strains but using different plasmids, have reported the same difficulties[72].

Due to the difficulty in transforming wild strains, only a limited number of strains could be transformed with pUCK3564; 8 out of 21 total strains. The successfully transformed strains were: HN16, BT14, HN10, HN9, HN6, HN11, BJ6, and HN2 (**Table 3.1**). Of these only HN6 showed a similar level of transformants to a BY4742 control, transformed with pUKC3546, even then, the number of HN6 cfu was still much lower than BY4742. The other successfully transformed strains only showed between 1-15 colonies per plate, over multiple attempts

3.5 Halo assay for wild strains

Once the strains transformable with the LiAc method had been identified, they were grown in a shaking incubator for 48 hrs, before being diluted down to the same OD_{600}. Each strain was then plated in triplicate allowing for statistical significance to be worked out, as well as identification of any outliers. To identify any strains that may be native KT producers, all untransformed strains were plated out, with a Ski4 positive control. Two strains were identified as native KT producers: 60 and NH6. This is summarized in **Table 3.3.** **Figure 3.7** is

an example of a mature halo assay plate: it has been incubated for 3 days, showing good KT production for the transformed strains HN6 and HN11, and for the untransformed strain 60. However, the transformed strans of BT14, HN16 and HN9 show no KT production, and HN10 shows only very poor production levels. BY was used as a negative control and Ski4 as a positive control. **Figure 3.8** is a graph of the data obtained across the 3 repeats of the halo assays, with error bars of +/- 1 standard deviation. **Table 3.2** contains the average halo size data for each strain.

Table 3.2 Average size of halo produced in the halo assay for both the untransformed wild type, and the transformed strains. Strains BT14, NH16, HN9 and NH2 produced no secretion both when transformed and untransformed so have been omitted from this table.

strain name	**Wt (cm^2)**	**Transformed (cm^2)**
3903	0	0.788
3883	0	0.658
3893	0	0.841
3923	0	1.046
3913	0	1.368
BY	0	1.439
SKI4	1.367	1.383
HN6	0.497	1.661
60	1.167	0
HN11	0	1.401
HN10	0	0.124
HN15	0	1.231

Table 3.3 Summary of successful transformation, and native Killer toxin secreting ability of *S. cerevisiae* used		
Strain name	**Successfully Transformed?**	**Native Killer toxin producer?**
21	No	No
27	No	No
60	No	Yes
99	No	No
111	No	No
114	No	No
BJ20	No	No
BJ6	Yes	No
BT14	Yes	No
FJ7	No	No
HN10	Yes	No
HN11	Yes	No
HN14	No	No
HN15	No	No
HN16	Yes	No
HN2	Yes	No
HN6	Yes	Yes (very low levels)
HN9	Yes	No
SX1	No	No
SX2	No	No
SX6	No	No

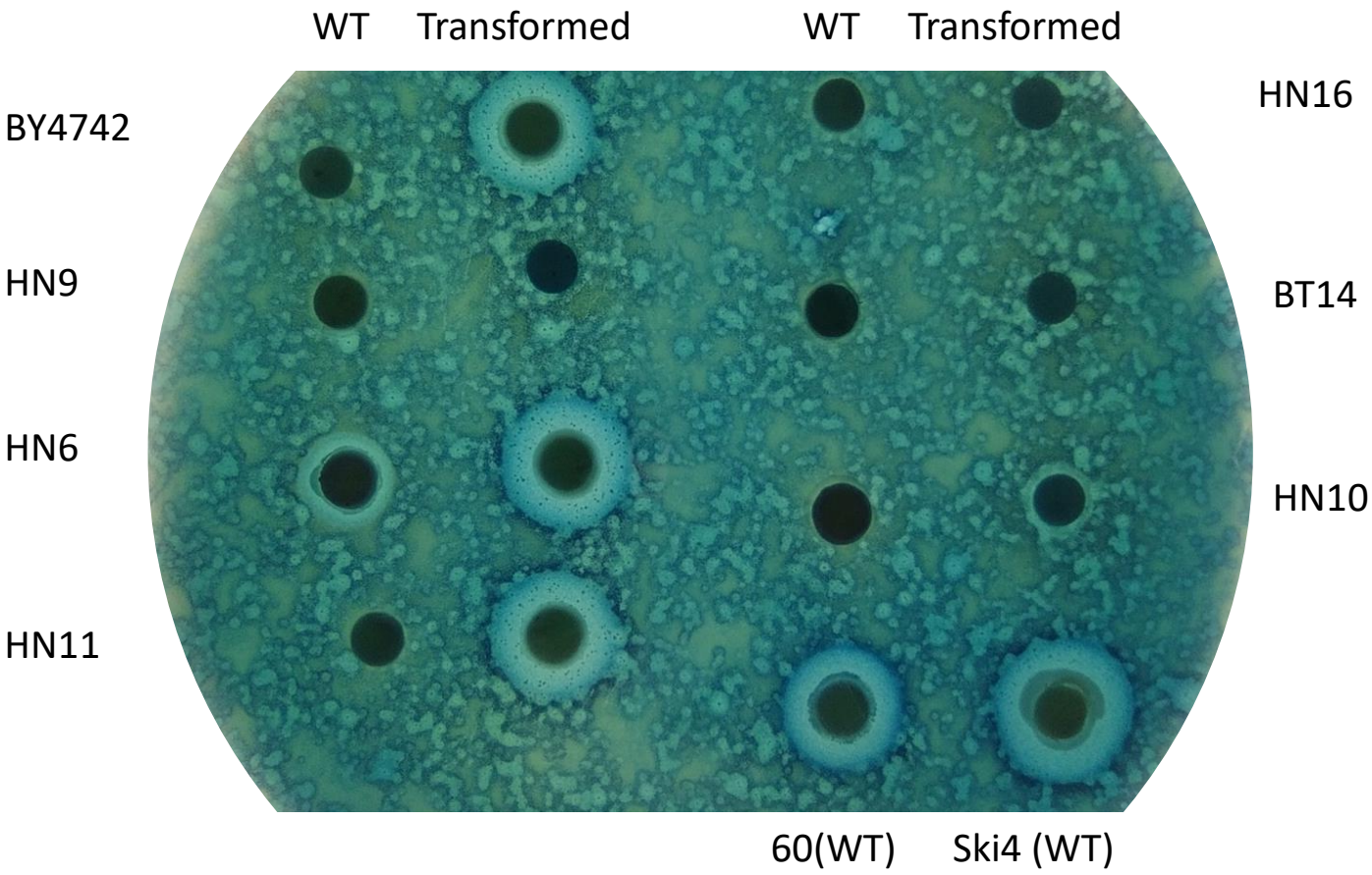


Figure 3.7. **An example of a mature halo assay plate**. In this case, 2 sets of strains have been tested, each in a pair of transformed (on the right of each column) and non-transformed (on the left). The top strain in the left hand column is BY4742, whilst the bottom strains in the right hand column are Ski4 and strain 60 (both untransformed)

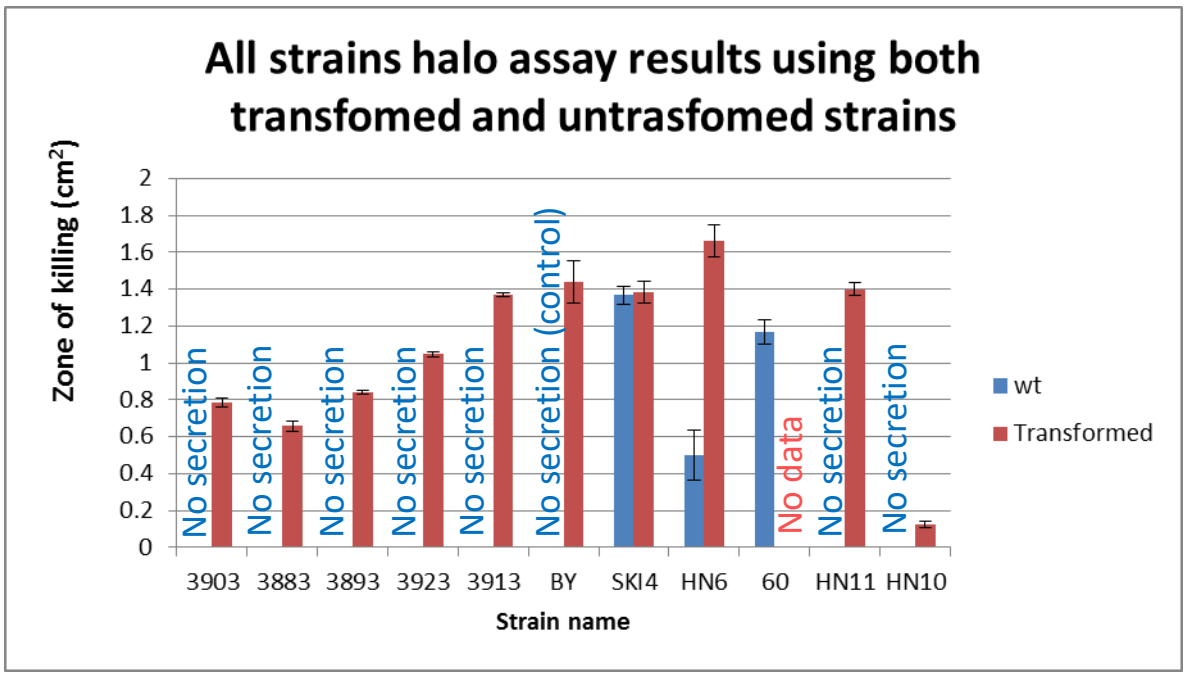


Figure 3.8 the results from the halo assay, including the domesticated and wild strains. Blue bars represent the untransformed strains and red bars represent transformed strains: pVT100u and pUKC3564 for the domesticated strains and the wild strains respectively. Strains BT14, NH16, HN9 and NH2 produced no secretion both when transformed and untransformed so have been omitted from this graph. Strain 60 was not transformed, and so there is no data for transformed strain 60, this does not mean it would not secrete, but the data was not obtained. Strains BY and Ski4 are negative and positive control respectively

3.6 CRISPR/Cas9 gene insertion

3.6.1 Selecting the target site for gene insertion

Running in parallel with the development of pUKC3564 was a set of experiments with the aim of inserting the K1 KT gene directly into the genome of the wild yeast strains. To do this a target site had to be chosen, which allowed the selection of strains with a successful genome integration event. By inserting the gene, as well as a stop codon into the middle of the ADE2, which encodes the enzyme Phosphoribosylaminoimidazole carboxylase, any

strains with a successful genome integration event would accumulate adenine precursors, which colour the cell a characteristic red. To check conservation of the chosen target site within the genome, BLAST[73] was used on the target sequence, against 50 *S. cerevisiae* reference genomes found on the Saccharomyces Genome Database[74] (**Appendix 2**).The target sequences were completely conserved between all the genomes tested. Using this sequence would also produce no off target cuts.

Once the target was decided, the sgRNA was designed and inserted into pML107 (**section 2.4.4**). PCR was performed as per **section 2.4.4**, and both the repair template and pML107 were transformed into the target strains. This was attempted 4 times, however, no red colonies were seen, and due to time pressure, it was decided that proceeding on the pUKC3564 front was more likely to produce results.

3.7 Secretome analysis

An alternative method to evaluate the proteins being secreted by a test strain was to precipitate the proteins found in the secretome, and visualises them on an SDS tris-tricine gel. This would allow basic identification of any products, and complement the data obtained from the halo assay. Given time, it could lead on to further analysis such as mass spectrometry .To this end, the secretome analysis was this was performed in parallel with the halo assay. To evaluate the secretome of the strains a protein precipitation approach was taken. Cells were grown overnight in YPD media, before the spent media was cleared of cells via centrifugation. Two approaches were tried to isolate secreted proteins from the medium: TCA protein precipitation, and Strataclean protein binding beads. Both methods ended with the samples being boiled in SDS sample buffer, before being loaded onto a tris-tricine gel, and then stained with commassie brilliant blue. **Figure 3.9** is the gel for Ski4/BY4742, using both methods of protein precipatation. **Figuere 3.10** shows the tris-

tricine SDS PAGE gel, both of the secreated proteins, and the cellular proteins, released by cell lysis. KT can be seen in neither gel

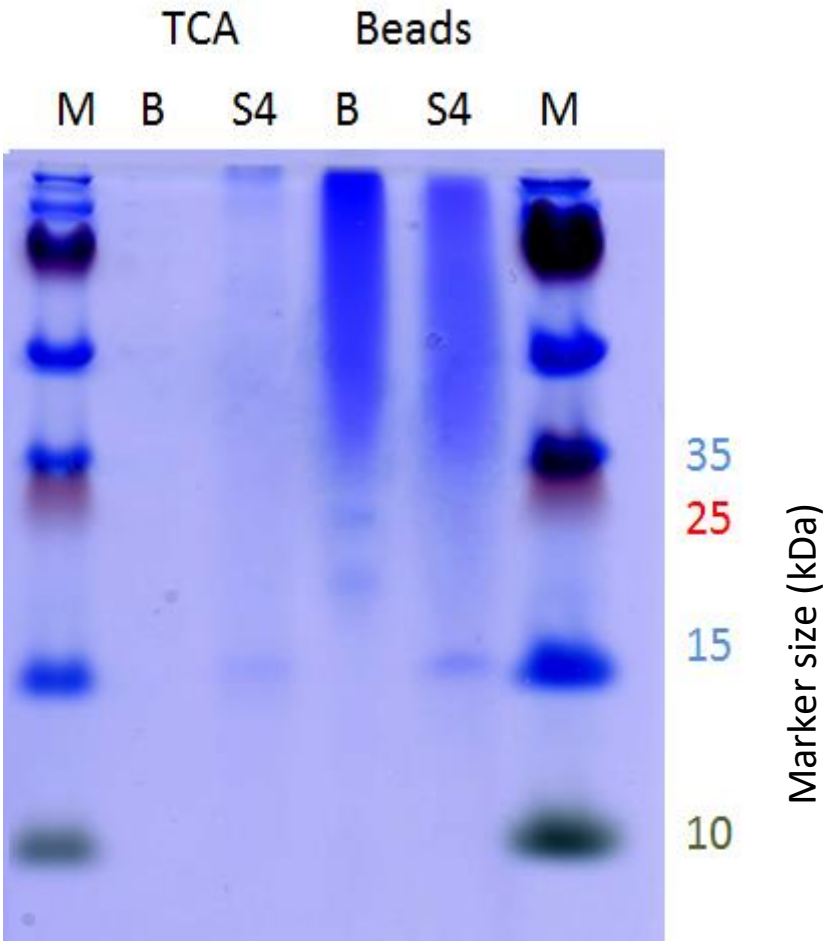


Figure 3.9 TCA / Strataclean bead gel stain with coomassie blue. Lane M: maker, PageRuler plus. Lane B: BY4742 media. LaneS4: Ski4 media. Size of marker bands noted in killer Dalton. Note the band at 15kDa; this is likely the killer toxin as previous work on KT has noted that it appears at a lower weight marker than expected, at around 15 kDa. However, mass spectrometry would be required to confirm the identity of the band.

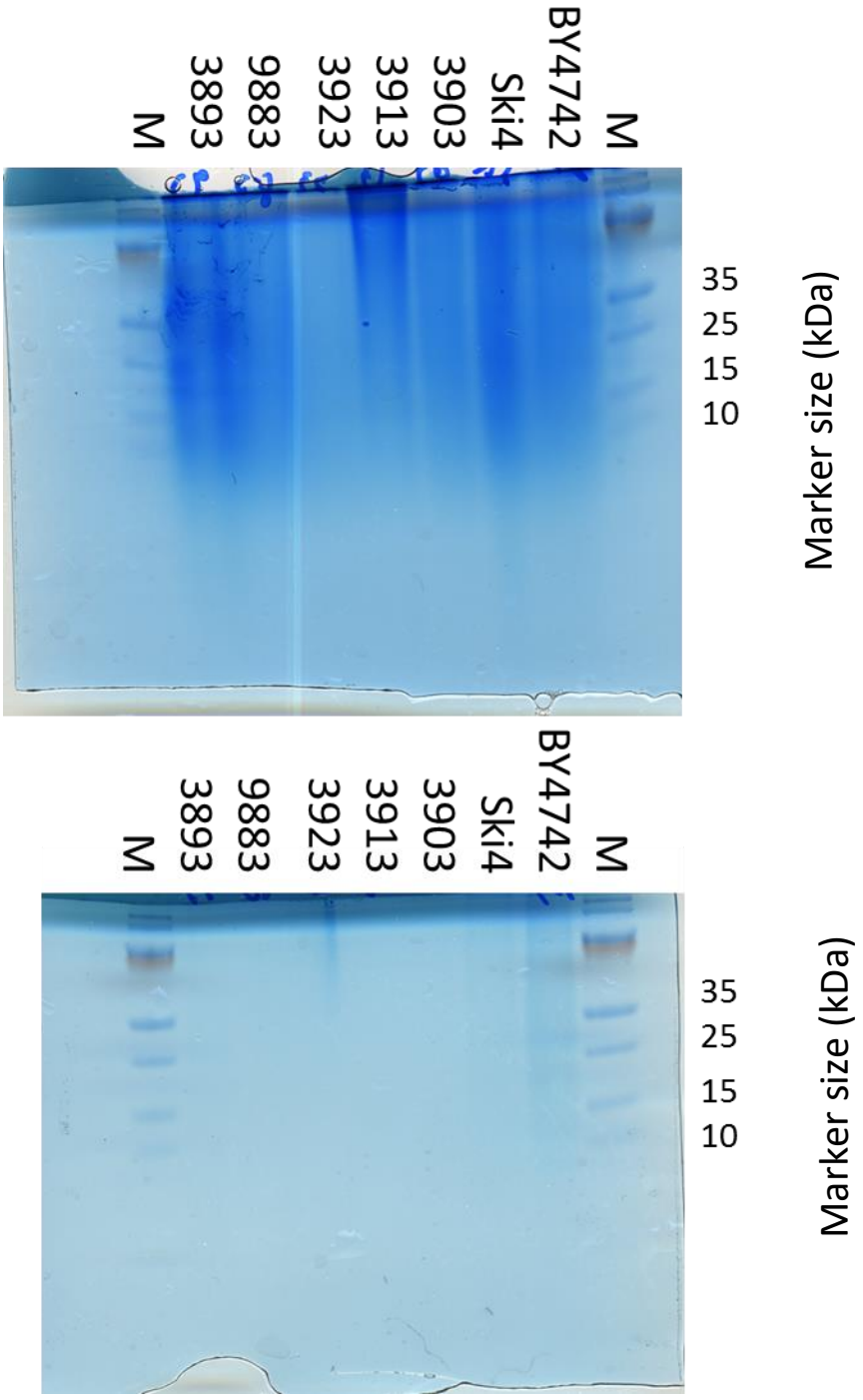


Figure 3.10 TCA protein precipitation tris-tricine gel stain with coomassie blue. Top panel; cellular proteins from lysed cells. Bottom panel; cell cleared media. Each lane is marked with the name of the strain run in that lane. The lanes marked M are the marker lanes.Ski4 and By4742 are controls. Unlike **figure 3.8**, neither gel shows killer toxin despite similar preparation method.

3.8 Gaussia luciferase assay

As well as the K1 Killer toxin, another protein was used as a secondary model protein; Gaussia luciferase (GL) , a secreted luciferase which uses coelenterazine as a substrate to produce light at 482 nm. For this assay, the 5 domestic strains as well as BY4742 were transformed with pCG495, a

GL containing plasmid, and tested with an untransformed BY4742 as a control (**Section 2.5.2**). This assay had measurements taken at 15, 30 and 45 minute intervals, **Figure 3.11** is the graph of these sets of measurements. We can see that the control strain (BY) transformed with the plasmid out performs all the domestic strain, which is consistent with

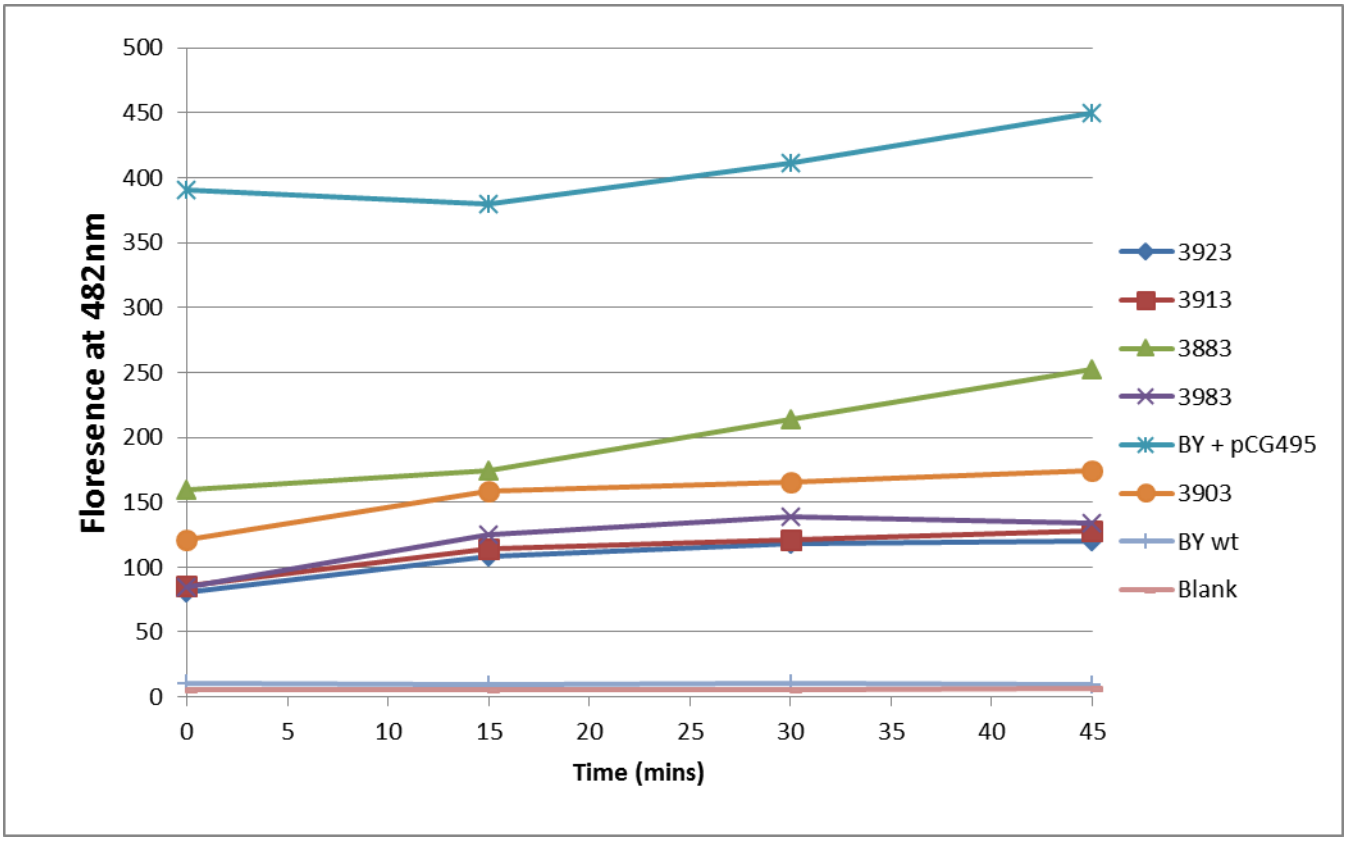


Figure 3.11 Graph of the GL data. The data is the average of 6 measurements for each strain, and readings were taken every 15 min, after 10 min incubation, as per the protocol of the kit used. Each strain is demarked by a different coloured line. Both blank, and a BY negative control are shown along the bottom, showing that all strain outperformed a control.

the halo assay results, however, this is where the similarities end: strain 3883 had shown a very low KT production, but was the second heist in this assay. This might indicate that each strain is better a particular set of proteins, and if this assay was generalised to all the strains, we might see a similar difference in results.

4. Discussion and future work

4.1 strategies for genetically engineering yeasts

During this project, a number of different methods have been used to insert the killer toxin genetic information. This 'gene' is an artificial creation originally of cDNA created from the killer toxin viral dsRNA particle. For sake of ease, in this document, the term killer toxin gene is used to refer to this DNA which encode the toxin, as well as the promoter and terminator used (ADH1 in this case).

4.1.1 LiAc transformations

Transforming the domesticated strains and BY4742 with pVT100u, pML107 and pCG495, was done without any problems. However, using the wild strains proved more problematic. The exact reason for this is unclear, however it is interesting to note, that of the 8 strains successful transformed, (NH16, BT14, NH10, HN9, HN6, HN11, BJ6 and HN2) all but one (BT14) came from china, and all bar BT14 and BJ6 came from bark of fagaceae trees in Hainan.

The yeast cell wall is extremely tough, comprising of cross linked layered branched mannose, Beta glycans, and chitin; yeasts also show the ability to changes its cells wall structure in response to environmental stresses[75]. As a direct result, getting DNA into a yeast cell is difficult, and requires the use of 30 minutes heat shock, as well as high concentrations of

lithium acetate. whilst the exact underlying molecular mechanism though which lithium acetate transformations function is unclear, some recent studies have looked at the structural processes that occur during transformation [76]. Polyethylene glycol (PEG) allows DNA to adhere to the cell walls, whilst lithium acetate and heat shock serve to help increase the permeability of the cell wall. This was shown with experimentation with transformation of spheroplasts (which lack a wall); this process was unaffected by the presents of lithium acetate, or by heat shock of the sample[77,78].

It is possible that the strains resistance to transformation either have fundamental differences in the nature of their cell walls (such as in the complexity of branched mannose) or a greater ability to adapt to a hostile environment (such as heat shock). Such adaptations might be serve as a direct resistance to foreign DNA; it has been theorised that one of the key roles of the cell wall in yeasts is as a protective measure against viral infection[79], which would certainly include a resistance to picking up large amounts of intact DNA from the environment. Alternatively, such resistance could simply be a side effect of environmental adaptation; for example, resistance to heat shock. This could be an interesting avenue of investigation: a range of phenotypic assays (heat shocks, extremes of PH or osmotic stress etc.) could help point to an underlying advantage which may explain this resistance to transformation.

4.1.2 CRISPR

The aim of the CRISPR experiment was to insert the KT gene into the genome of the test strain, allowing stable insertions without the worry of plasmid loss. Developing the plasmid required would also mean that genes for any test recombinant protein could be quickly inserted into the genome, by simply chaining the repair template. Overall, the CRISPR

experiment was not successful, this project has laid the ground work for future projects involving CRISPR mediated gene insertion in the yeast genome: the killer toxin gene was successfully inserted into pML107, along with the guide RNA to target the system to the ADE2 gene. One possibility for the unsuccessful outcome of this experiment is the length of the insert: much of the previous work done using Cas9/CRISPR, has focused on single base pair, site directed mutagenesis, or small (<100bp) insertions and deletions. Larger insertions have only recently been explored [80]. The method used here had followed the methodologies for small inserts, and would require modification to accommodate the larger 3kB insert required, most notably the use of homologous arms, much larger than were used for this project (around 10 fold longer).

4.1.3 Future work

Given more time on this project, development of the CRISPR/Cas9 technology would take priority: stable introduction of the KT 'gene', into the genome, would dispense with the requirement of markers such as hygromycin B resistance. Using larger homologous arms for homologues recombination may yield results: the Zhang lab has reported insertions of up to 5.3 kB with A 720bp right arm and 967bp left arm[80].

If the CRISPR/Cas9 approach cannot be successfully utilised, then I would proceed on the plasmid front. To overcome the transformation difficulties, I would use a single plasmid-multiple proteins approach: By using pUKC3564 as a starting point, multiple recombinant proteins could be inserted into the same plasmid, thus allowing for multiple secretion assays off a single transformation. Whilst this doesn't address the strains that have yet to transformed at all, it would save large amounts of time and resources that would otherwise be taken up with repeated transformation attempts.

To tackle the transformation issue, transforming spheroplasts via electroporation would be a way to get around any problems that may be caused by the cell wall: if this method was successful, it would confirm that it is the cell wall that causes the resistance to transformation seen. Spheroplasts would be generated using lyticase treatment, to break down the cell wall.

To analyse the underlying causes of the transformation problems, looking at the cell wall may yield useful results; previous yeast cell wall studies have utilised High pressure freezing Transmission electron microscopy (HPF-TEM) to look at size and structure of the yeast cell wall[81]. It might be possible to see difference between theses 'untransformable' strains, and wild type strains, which may reveal more about how the LiAc transformation exactly works.

4.2 Native Killer toxin in Wild strains

During the course of this investigation, a number of strains generated positive results from the halo assay without being transformed with a KT producing plasmid. This indicates they are native producers of KT; growing these strains on a lawn of S6 strain leads to a halo of killed cells.

4.2.1 Producing strains

The vast majority of wild yeast strain do not produce killer toxin. For reasons discussed below, it can be a disadvantage to host the virus. The Ski mutants (Ski4 being a member) are notable exceptions to this, as not only are they natural secretors, but also contain mutations (in the so called 'Ski genes') which makes them 'super producers'. This makes Ski4 a good benchmark against which to test the secreting strains. By using the halo assay, and measuring the size of the zones of clearing, we can get an idea of the strains relative abilities to secrete KT, and so infer there overall secretion abilities.

Of the wild strains, two were noted as generating positive results in the halo assay, implying they are producers of killer toxin, even when untransformed: 60 and HN6. Strain 60 was the most potent producer, almost at the level of Ski4, whereas HN6 produced significantly less. This small number is consistent with previous studies , which find very low occurrence of KT production in wild strains [82, 83]. This raises the question of why these strains might be natural producers; producing KT can be a selective advantage, by killing off other strains competing for the same ecological niche. However, the KT virus does come with disadvantages; the fact that natural the gene is encoded on a dsRNA virus means the cell hosting it must give up any RNA silencing capacity. If a cell within a carrier colony develops this ability, it will lose the immunity granted by the dsDNA virus, and be killed by its secreting neighbours meaning that once infected, the strain self-selects for the virus[84]. Of course this also mean infection with the KT virus renders the strain, and any descendants, vulnerable to attack from other viruses, due to the removal of the RNAi systems. The end result is that strains that acquire the KT virus do have an advantage over other, sensitive strains, but give up a measure of their ability to defend themselves. It has been suggested that this trade is somewhat unfavourable, and that this might explain the low numbers of wild strain hosts[84] .

The fact that the two different strains appear to produce different levels of killer toxin could indicate their secretion abilities vary; strain 60 being far more efficient than NH10. Alternatively, they could be producing different forms of KT: there are 2 common variants of killer toxin, K1 and K28. It is possible that S6 is more sensitive to one variant than the other

4.2.2 Future work

It is very possible that these strains are not producing KT, but rather another protein that is toxic to S6. Given more time I would perform the same TCA protein precipitation that was

performed on the domesticated strains, to check if these strains were producing KT. Extracting all the dsRNA from each strains would also elucidate what, if any, dsRNA viruses infect these two strains.

This would also help clear up if the two strains are infected with different viruses and so are producing different toxins. This would also neatly segue into creating different plasmids with the different killer toxin to assay which one is more efficient at killing a variety of sensitive strains.

4.3 The gaussia luciferase secretion assay in the domestic strains

4.3.1 Assay results

The results of the gaussia luciferase assay show the domestic strains tested secreting a very low level of GL, which was borne out over both experiments. This is largely consistent with the work of the halo assay for KT; the domestic strains do not secrete proteins as well as BY4742. There is however one difference: the halo assay showed that 3813 secreted K1 KT at a similar level to BY4742, but 3813 produced an average of only 128 fluoresces units at T=45, almost 4 fold lower than BY4742 at the same time point.

4.3.2 Future work

The difference between the GL assay and the halo assay, when looking at strain 3913, highlights the importance of using different proteins; GL is much larger than KT and has more disulphide bounds, indicated that 3913 is better as secreting more simple proteins, but this ability is significantly less with larger, more complex proteins.

Inserting the GL gene into the wild strains, may show different strains having different abilities to secrete GL in comparison to K1 KT. If a fast and simple method could be developed for gene insertion, then a range of proteins of different size and complexities

containing different PTM, could be used to assay the full range of the wild strains abilities to secret recombinant proteins.

4.4 Protein secretion in yeasts

4.4.1 The dangers of ROS

Whilst this project has largely focused on using killer toxin as a model protein, we also have to look at the greater context, of general recombinant protein production. When thinking about protein secretion in yeasts, we have to consider how our product of interest affects and are affected by other process in the cell. On the most basic level, this might include looking at the toxicity of the product itself to the cell. We should also consider if the product or the process of making the product puts stress on the cell, which we must take into consideration when growing these cells for example the generation of reactive oxygen species by the production of disulphide containing proteins.

It has been well established that Reactive oxygen species (ROS) and cell apoptosis are linked[85,86,87] .in fact the mechanism of action for KT seems to involve ROS generation[88]. Molecular oxygen Is the terminal electron acceptor during the formation of disulphide bounds, though the action of Ero1 and PDI which in turn leads to the generation of peroxide, a strong ROS. This means that cells which produce large amounts of disulphide bound containing proteins, necessarily would have to deal with large volumes of ROS, and so the cell is limited in the volume it can produce by its ROS clearance mechanisms[89] (such as peroxidases).Normally this would not be a problem, however, by introducing a gene with a highly active promoter, the cells have a greater level of ROS to deal with. These cells are also more vulnerable to other sources of ROS, including exogenous ROS, and to any spontaneous mutations which affect the ROS system, directly or indirectly.

Chronic reductive stress manifests in the ER as futile cycles of proteins with disulphide bonds folding, when bound to chaperone proteins and members of the PDI protein family, and unfolding, as they are released. Not only do these cycles waste the cells resources, they also produce more ROS with each cycle. Cells that are producing large volumes of these proteins are more greatly affected: more proteins undergoing these cycles' means more ROS produced.

Hydrogen Peroxide also has a direct effect on the endoplasmic reticulum ability to fold proteins, by causing down regulation of the foldase GRP94 and GRP78[90,91]. It has been suggested that CYP2E1- derived oxidant stress was responsible for this down regulation[92], thought the exact mechanism is still unclear.

All these problems caused by ROS mean that there is a hard limit on protein production, which is related to the cells ability to deal with ROS, specifically peroxide. Due to the fact that many peroxidases are limited by the rate of diffusion [93] attempts to optimise their structure to increases the rate of catalysis would be fruitless. Instead to increases the production levels of disulphide bond containing proteins, work could be done to increase protein levels of these enzymes, for example. Looking at human cell which produce a large volume of disulphide containing proteins, such as plasma cells, may hold the key to this. By increasing a cells resistance to acute oxidative stress, we can increase the upper limit of production of our protein of interest. Yeast cells only have the Ero1-PDI pathway to produce disulphide bounds[94], and as a result an increases in disulphide bounds always leads to an increase in hydrogen peroxide formation. Mammals on the other hand contain tissue specific alternative pathways , such as the Ero1[94]and PRDX4 mediated[95] amongst others[96],

Figure 4.1, some of which uses hydrogen peroxide as the terminal electron acceptor, others are completely peroxide independent.

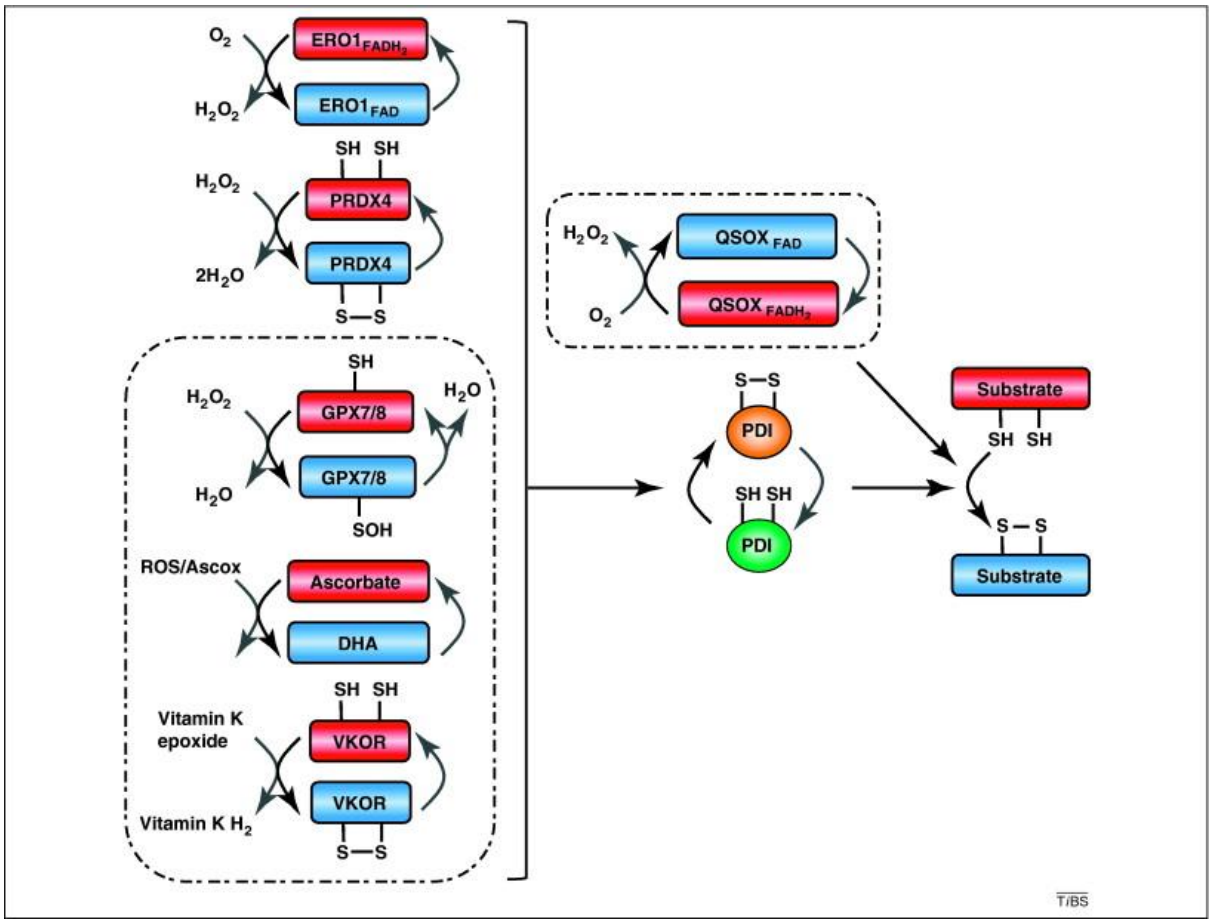


Figure 4.1 Summary of various disulphide bound forming pathways in mammalian cells. With the exception of the Ero1 pathway, these methods are unavailable to yeast cells, which must produces peroxide to create disulphide bounds. Taken from N. J. Bulleid and L. Ellgaard 2011[96]

It should also be noted that very few proteins are natively secreted by yeasts: in fact killer toxin is the only endogenous disulphide bound containing protein secreted by *S .cerevisiae*. The dangers of producing high levels of peroxide do not often counter balance the advantages brought by secreted proteins, as a result, infection by the KT virus is rarely see, and is defended against by native RNAi. KT may not be an exception to this rule, as we have

already discussed, there are other selection pressure selecting for the dsRNA virus than just environmental fitness.

4.4.2 Protein precipitation; secretome and cellular protein analysis

In an attempt to evaluate the secretome of the domestic strains, two methodologies were used: Trichloroacetic acid (TCA) protein precipitation and Strataclean beads precipitation both methodologies seem to work for the precipitation of killer toxin from the secretome, with clear bands. TCA precipitates less material , which should not have been much of a problem however, when the TCA precipitation was repeated, it did not produce substantial results, even with longer incubation times. The higher sensitivity of the bead should produce more clear results and would allow more strains to be evaluated. This is likely down to the low levels of KT secreted natively by cells, as the cellular proteins gel shows a very strong signal which is consistent with the large amounts of total protein seen inside of the cell.

The next stage would be to use mass spectrometry to confirm the identity of the protein band. If this was the case, then using the Strataclean beads, and band intensity analysis could be a powerful quantitative alternative to the halo assay for killer toxin.

4.5 hygromycin B resistance in the wild strains

It is worth briefly looking at the very high level of resistance displayed by the wild strains, to the anti-biotic hygromycin B; such high resistance is unusual in wild strains of *S.cerevisiae*. This is especially notice able at levels 10 fold more than what is typically required to inhibit growth (2000μg/ml required, whereas 200μg/ml is more typical in ws), where two strains; 60 and HN14 showed a single colony each. This was repeated 3 times, with single colonies developing on two of the repeats .This effect has been seen before in *Monilinia fructicola*, [97]

but little work has been done in *S.cerevisiae.* The mutants seen in *M. fructicola* showed reduced growth rate and reduced resistance to demethylation, but the underlying mutation was not elucidated. Testing these strains to find out the underlying causes, by using SNP arrays, and testing to see if the strains can adapt to any other anti-biotics, would be worthwhile, to see if these strains are more adaptable than other strains.

4.6 Wild strains as protein production and secretion

The main aim of this project was to test the viability of these strains as good producers of recombinant protein production, so the advantages and disadvantages should be looked at.

4.6.1 The advantages of using the wild strains tested here

Despite some resistance, it has been shown that at higher concentration levels of hygromycin B is a useable selective marker for plasmid selection, allowing plasmids to be easily designed to insert recombinant protein genes, or use CRISPR/Cas9 for further engineering.

Strain 60 and HN6 both showed they natively produce killer toxin, and so have the required machinery to produce correctly folded disulphide bound containing proteins; strain 60 had a halo area of 1.16 cm^2 when untransformed, larger than 4 of the domesticated strains.HN6 produced only 0.496 cm^2, far less than any of the transformed strains , however, when transformed this level was greatly increased to 1.66 cm^2 , implying the lower secretion rate is linked to the dsRNA KT virus itself (transcription speed for example), not the overall machinery of the strain .

Strains HN6, and HN11 showed that once transformed with pUCK3564 they could produce large volumes of killer toxin, rivalling even Ski4. Transformed Ski4 produced an average halo of 1.38 cm^2, whereas HN6 and HN11 produced 1.66 cm^2 and 1.40cm^2. Adding in error bars of

+- 1 standard deviation, HN11 and SKI4 become similar, but HN6 remains a larger secretor with a lower bound of 1.58 cm^2, compared to the upper bound of ski4 of 1.44 cm^2. Interestingly untransformed strain 60 has an average halo of 1.16 cm^2, which is only 0.20 cm^2 smaller than untransformed ski4, the so called 'super producer', indicating that strain 60 might have one of the Ski mutations.

Strains HN6 and HN11 would be the perfect candidates to take forwards for further testing. HN6 is of particular interest, as it showed a good level of transformation, though still less than BY4742 transformed under the same condition, whilst HN6 produced the largest halo of all strains tested .

4.6.2 The disadvantages of using the wild strains tested here

A massive hurdle to use of theses strains is the resistance to standard LiAc transformation. Only 9 of the 21 strains were not resistant, and those that were successfully transformed showed very low efficiency. Until alternative methodologies have been tried (transformation of spheroplasts via electroporation for example[76]), these strain cannot be tested for recombinant protein secretion.

Of the ten strains successfully transformed, only four gave observable evidence of secretion and of those (HN10) showed an extremely low level of secretion in the halo assay: around 9% of the level seen with Ski4. The five remaining strains (HN9, BJ6, HN16,BT14 and HN2) as well as HN10, have displayed limited or no ability to secrete killer toxin, so would be inappropriate to use in recombinant protein production.

Of course, until these strains have their genomes published, we are also unable to perform any analysis on any genetic variation showed by these strain: even if a massively successful secretor was found. Whilst these strains could be used without the need for genetic analysis,

if the underlying genetic variance responsible for the observed phenotype was known, it could be introduced into existing high producing strains, to enhance these industrial strains' protein production rate.

4.6.3 The overall view of the wild strain and summary of the project

Of the above disadvantages, two of them may be able to be overcome: genome sequencing technology is getting faster and cheaper year on year, meaning these strain could easily be fully sequenced, or if full sequencing was still too expensive, SNP arrays could be used to elucidate genetic variance in successful secretor, which could be used to explain the observed phenotype. As mentioned above, alternative transformation strategies may yield good results.

The simple fact that BT14 HN16, HN9 BJ6 and HN2 generated no secretion in the halo assay, and that HN10 secreted very little, means these strains are not useful to us. However strain 60, HN6, HN2 and HN11 show very good secretion, so could be carried forward for further testing (**figure 3.2**).

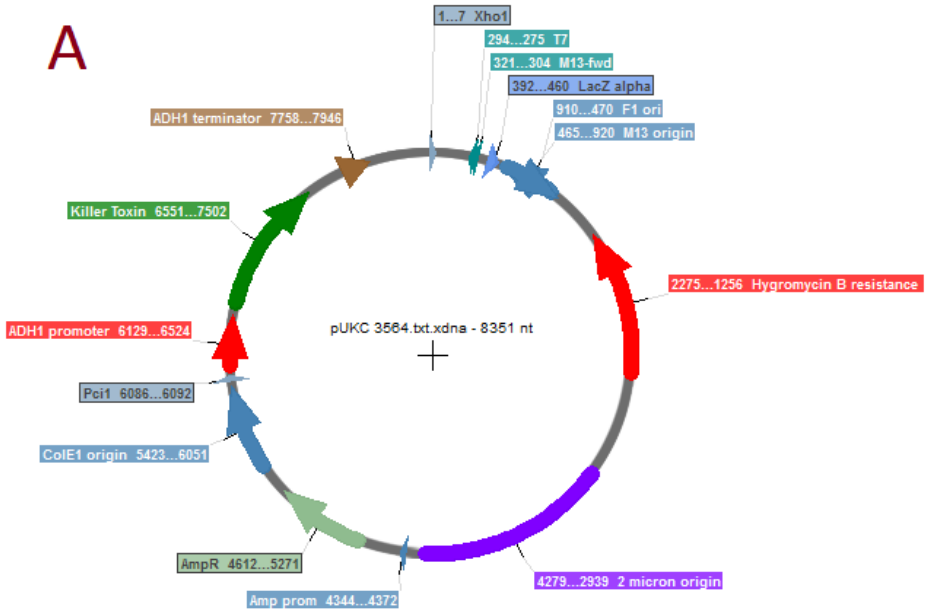
So to summaries, whilst many of the strain tested here proved to be either resistant to transformation, or produced no or little killer toxin when transformed, four of the tested trains showed good secretion. This shows that there is potential in testing a wide variety of wild strain. The simple fact is we don't know what potentially useful strain we might find, when testing wild strains.

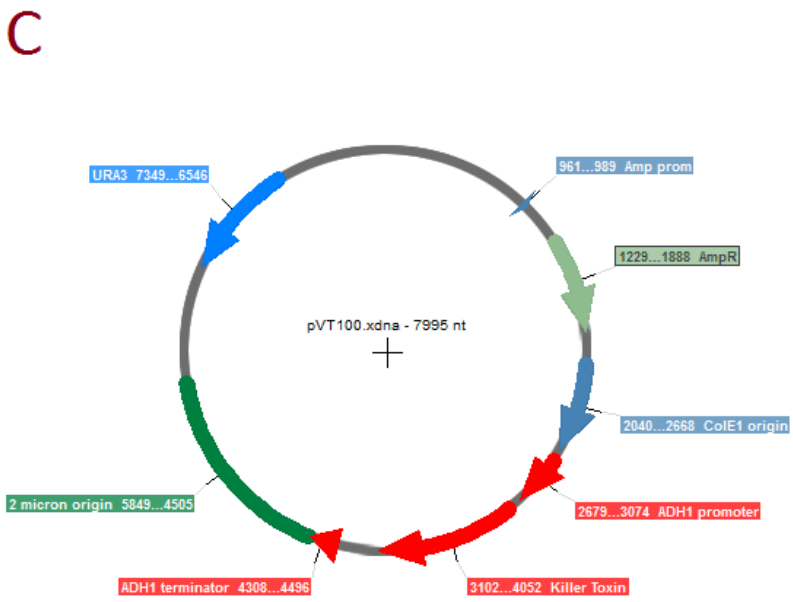
During this project, a plasmid based method was developed to insert the killer toxin gene into strains lack auxotrophic makers, as well as a method to assay for secreted killer toxin, namely the halo assay. The secretome was also looked at, and, using the strataclean beads, killer toxin extracted from media that had been used to grown transformed strains. There is

plenty of future work possible from this project, such as confirming the identity of the killer toxin with mass spec, and further attempts to transforms the wild strains.

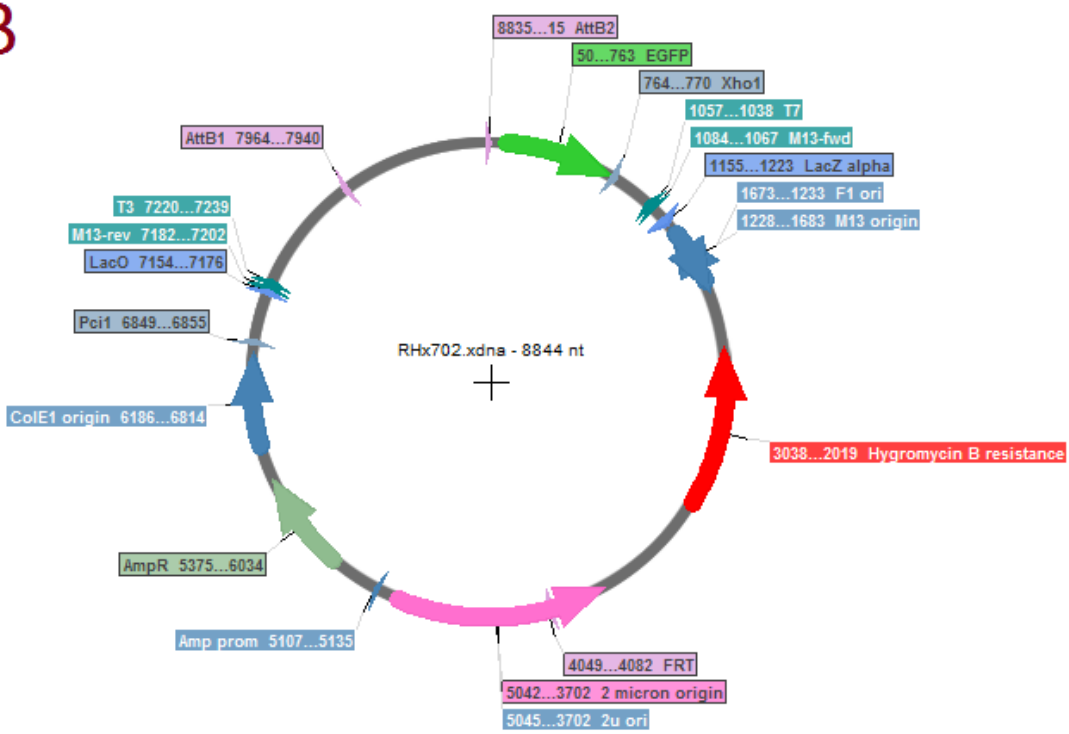
Appendix 1: Plasmid maps

Plasmid maps: **plasmid maps for all plasmids used**. A) pUKC3564; novel Killer toxin and hygromycin B resistance plasmid, B) RHx702; original hygromycin B resistance plasmid, C) pVT100; original Killer toxin, D) pML107; Cas9/Guide RNA combined CRISPR plasmid. E) pCG495; Gaussia luciferase containing plasmid

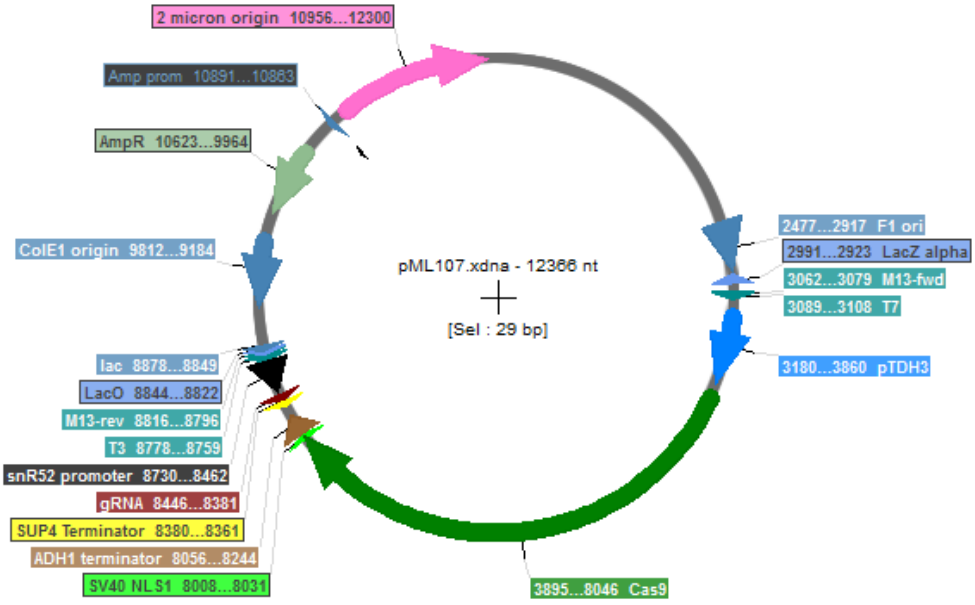




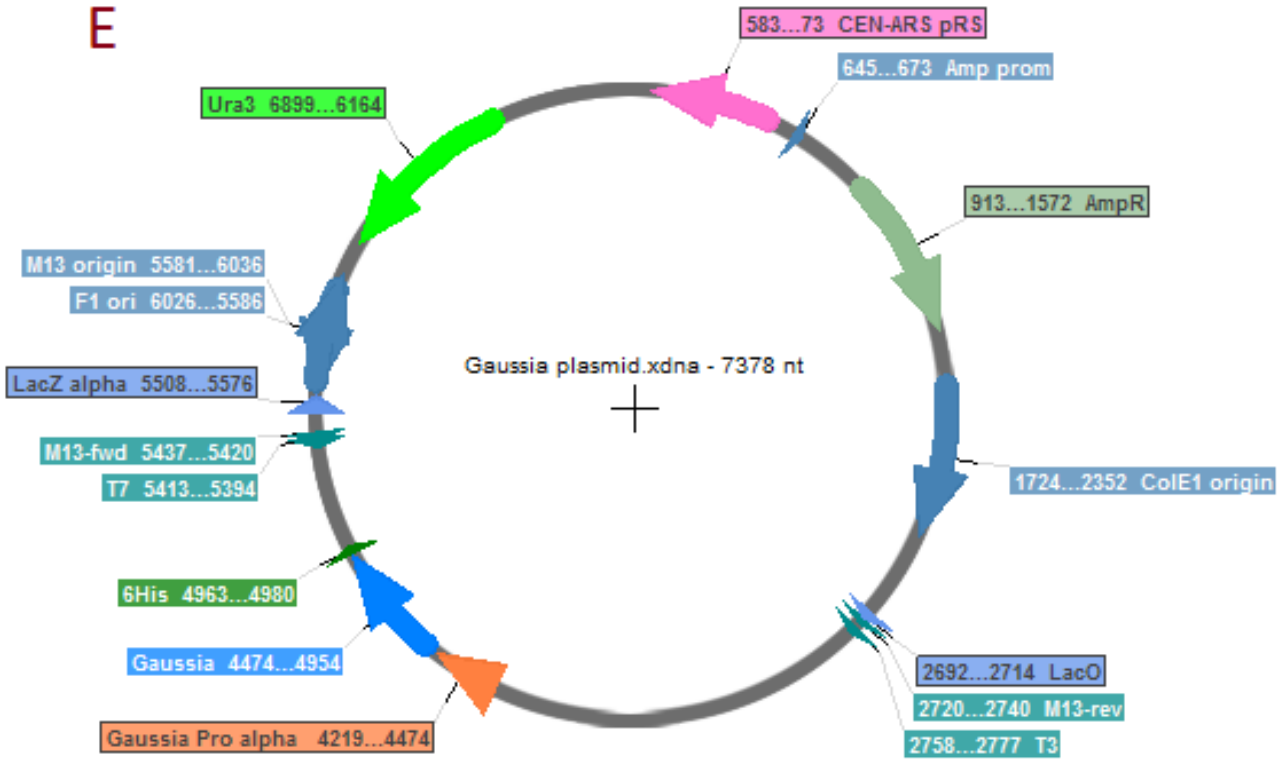
B


8835...15 AttB2
50...763 EGFP
764...770 Xho1
1057...1038 T7
1084...1067 M13-fwd
1155...1223 LacZ alpha
1673...1233 F1 ori
1228...1683 M13 origin
AttB1 7964...7940
T3 7220...7239
M13-rev 7182...7202
LacO 7154...7176
Pci1 6849...6855
RHx702.xdna - 8844 nt
ColE1 origin 6186...6814
3038...2019 Hygromycin B resistance
AmpR 5375...6034
Amp prom 5107...5135
4049...4082 FRT
5042...3702 2 micron origin
5045...3702 2u ori

D


2 micron origin 10956...12300
Amp prom 10891...10883
AmpR 10623...9964
2477...2917 F1 ori
ColE1 origin 9812...9184
2991...2923 LacZ alpha
pML107.xdna - 12366 nt
3062...3079 M13-fwd
[Sel : 29 bp]
3089...3108 T7
3180...3860 pTDH3
lac 8878...8849
LacO 8844...8822
M13-rev 8816...8796
T3 8778...8759
snR52 promoter 8730...8462
gRNA 8446...8381
SUP4 Terminator 8380...8361
ADH1 terminator 8056...8244
SV40 NLS1 8008...8031
3895...8046 Cas9

E


583...73 CEN-ARS pRS
645...673 Amp prom
913...1572 AmpR
1724...2352 ColE1 origin
2692...2714 LacO
2720...2740 M13-rev
2758...2777 T3
Gaussia Pro alpha 4219...4474
Gaussia 4474...4954
6His 4963...4980
M13-fwd 5437...5420
T7 5413...5394
LacZ alpha 5508...5576
M13 origin 5581...6036
F1 ori 6026...5586
Ura3 6899...6164
Gaussia plasmid.xdna - 7378 nt

Appendix 2: CRISPR Blast results

Blast results across 58 model strains of yeast. Sequence used: TTTTCCCGGTTGTGGTATATTTGGTGTGGAAATG which is the target sequence with 4 residues either side

Sequences producing significant alignments:	**Score (bits)**	**E value**
gb|AFDF01001248.1| Saccharomyces cerevisiae T73 Contig159.9, whole genome sh...	49.4	3.10E-06
gb|AFDG01001124.1| Saccharomyces cerevisiae CLIB382 Contig348.8, whole genom...	49.4	3.40E-06
gb|ABPD01002640.1| Saccharomyces cerevisiae YPS163 Contig394.1, whole genome...	49.4	3.40E-06
gb|AEWM01001999.1| Saccharomyces cerevisiae CLIB324 Contig286.3, whole genom...	49.4	3.40E-06
gb|AEWO01002210.1| Saccharomyces cerevisiae FL100 Contig282.4, whole genome ...	49.4	3.40E-06
gb|AFDD01000257.1| Saccharomyces cerevisiae UC5 Contig13.2, whole genome sho...	49.4	3.40E-06
gb|ABPC01001609.1| Saccharomyces cerevisiae M22 Contig180.6, whole genome sh...	49.4	3.50E-06
gb|ABSV01002165.1| Saccharomyces cerevisiae AWRI1631 chromosome 15 chr15_554...	49.4	3.50E-06
gb|AFDC01001012.1| Saccharomyces cerevisiae PW5 Contig38.54, whole genome sh...	49.4	3.50E-06
gb|AEWP01000422.1| Saccharomyces cerevisiae CLIB215 Contig72.2, whole genome...	49.4	3.50E-06
gb|AGSJ01000028.1| Saccharomyces cerevisiae EC9-8 contig00028, whole genome ...	49.4	3.50E-06
gb|AEWL01000738.1| Saccharomyces cerevisiae CBS 7960 Contig111.3, whole geno...	49.4	3.50E-06
gb|AEWN01000414.1| Saccharomyces cerevisiae YJM269 Contig77.2, whole genome ...	49.4	3.50E-06
gb|JRIG01000347.1| Saccharomyces cerevisiae DBVPG6044 scaffold-362, whole ge...	49.4	3.50E-06
gb|JRIE01000280.1| Saccharomyces cerevisiae YJM339 scaffold-295, whole genom...	49.4	3.50E-06
gb|JRIM01000838.1| Saccharomyces cerevisiae UWOPS05_217_3 scaffold-844, whol...	49.4	3.50E-06
gb|JRIK01000125.1| Saccharomyces cerevisiae L1528 scaffold-130, whole genome...	49.4	3.50E-06

gb|JRIJ01000006.1| Saccharomyces cerevisiae K11 scaffold-5, whole genome sho...	49.4	3.50E-06
gb|JRII01000136.1| Saccharomyces cerevisiae BC187 scaffold-139, whole genome...	49.4	3.50E-06
gb|JRIH01000113.1| Saccharomyces cerevisiae SK1 scaffold-145, whole genome s...	49.4	3.50E-06
gb|JRIB01000237.1| Saccharomyces cerevisiae YS9 scaffold-274, whole genome s...	49.4	3.50E-06
gb|JRIC01000085.1| Saccharomyces cerevisiae YPS163 scaffold-87, whole genome...	49.4	3.50E-06
gb|JRIF01000155.1| Saccharomyces cerevisiae Y55 scaffold-202, whole genome s...	49.4	3.50E-06
gb|JRID01000339.1| Saccharomyces cerevisiae YPS128 scaffold-353, whole genom...	49.4	3.50E-06
dbj|BABQ01000174.1| Saccharomyces cerevisiae Kyokai no. 7 DNA, chromosome 15,...	49.4	3.50E-06
gb|AEHG01000400.1| Saccharomyces cerevisiae CEN.PK113-7D contig213, whole ge...	49.4	3.50E-06
gb|JRIU01000141.1| Saccharomyces cerevisiae W303 scaffold-178, whole genome ...	49.4	3.50E-06
gb|JRIO01000092.1| Saccharomyces cerevisiae YPH499 scaffold-106, whole genom...	49.4	3.50E-06
gb|ACFL01000063.1| Saccharomyces cerevisiae JAY291 contig113_88_45, whole ge...	49.4	3.50E-06
gb|JRIV01000149.1| Saccharomyces cerevisiae CEN.PK2-1Ca scaffold-197, whole ...	49.4	3.50E-06
gb|JRIW01000060.1| Saccharomyces cerevisiae SEY6210 scaffold-70, whole genom...	49.4	3.50E-06
gb|ADXC01000074.1| Saccharomyces cerevisiae Vin13 chromosome XV VIN13_c15_74...	49.4	3.50E-06
gb|JRIY01000194.1| Saccharomyces cerevisiae D273-10B scaffold-266, whole gen...	49.4	3.50E-06
gb|JRIZ01000031.1| Saccharomyces cerevisiae JK9-3d scaffold-37, whole genome...	49.4	3.50E-06
gb|JRIR01000197.1| Saccharomyces cerevisiae BY4742 scaffold-262, whole genom...	49.4	3.50E-06
gb|JRIX01000100.1| Saccharomyces cerevisiae X2180-1A scaffold-126, whole gen...	49.4	3.50E-06
gb|JRIL01000028.1| Saccharomyces cerevisiae RedStar scaffold-51, whole genom...	49.4	3.50E-06
gb|ADVV01000079.1| Saccharomyces cerevisiae Lalvin QA23 chromosome XV QA23_c...	49.4	3.50E-06

gb|AAEG01000058.1| Saccharomyces cerevisiae RM11-1a cont1.29, whole genome s...	49.4	3.50E-06
gb|ADVS01000045.1| Saccharomyces cerevisiae AWRI796 chromosome XV AWRI796_c1...	49.4	3.50E-06
W303_contig00286 [organism=Saccharomyces cerevisiae W303] [strain=W303]	49.4	3.50E-06
gb|JRIS01000158.1| Saccharomyces cerevisiae BY4741 scaffold-202, whole genom...	49.4	3.50E-06
gb|JRIQ01000023.1| Saccharomyces cerevisiae 10560-6B scaffold-32, whole geno...	49.4	3.50E-06
gb|JRIN01000192.1| Saccharomyces cerevisiae FY1679 scaffold-244, whole genom...	49.4	3.50E-06
gb|AFDE01000030.1| Saccharomyces cerevisiae T7 Contig3.4, whole genome shotg...	49.4	3.50E-06
gb|AEHH01000073.1| Saccharomyces cerevisiae FostersB chromosome XV FOSTERSB_...	49.4	3.50E-06
gb|JRIP01000004.1| Saccharomyces cerevisiae RM11-1A scaffold-3, whole genome...	49.4	3.50E-06
gb|JRIT01000140.1| Saccharomyces cerevisiae FL100 scaffold-178, whole genome...	49.4	3.50E-06
gb|ACVY01000004.1| Saccharomyces cerevisiae Sigma1278b chromosome 15 chr15.2...	49.4	3.50E-06
gb|AEJS01000062.1| Saccharomyces cerevisiae VL3 chromosome XV VL3_c15_62, wh...	49.4	3.50E-06
gb|AEEZ01000095.1| Saccharomyces cerevisiae FostersO chromosome XV FOSTERSO_...	49.4	3.50E-06
gb|AAFW02000030.1| Saccharomyces cerevisiae YJM789 chromosome 15 chrXV.Conti...	49.4	3.50E-06
gb|AMDD01000032.1| Saccharomyces cerevisiae ZTW1 contig032, whole genome sho...	49.4	3.50E-06
chr15 [org=Saccharomyces cerevisiae] [strain=BY4741] [moltype=g...	49.4	3.50E-06
chr15 [org=Saccharomyces cerevisiae] [strain=BY4742] [moltype=g...	49.4	3.50E-06
ref|NC_001147| [org=Saccharomyces cerevisiae] [strain=S288C] [moltype=ge...	49.4	3.50E-06
emb|FN394216.1| Saccharomyces cerevisiae EC1118 chromosome XV, EC1118_1O4...	49.4	3.50E-06

References

1. Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J. L., Vázquez, E. & Villaverde, A. Microbial factories for recombinant pharmaceuticals. *Microb. Cell Fact.* **8,** 17 (2009).

2. Li, R., Chen, Q., Wang, P. G. & Qi, Q. A novel-designed Escherichia coli for the production of various polyhydroxyalkanoates from inexpensive substrate mixture. *Appl. Microbiol. Biotechnol.* **75,** 1103–1109 (2007).

3. Weickert, M., Doherty, D., Elaine, B. & Peter, O. Optimization of heterologous protein production in Escherichia coli. *Curr. Opin. Biotechnol.* **7,** 494–499 (1996).

4. Widmann, M. Comparison of folding rates of homologous prokaryotic and eukaryotic proteins. *J. Biol. Chem.* **275,** 18619–18622 (2000).

5. Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42,** 287–299 (2008).

6. Novy, R., Drott, D., Yaeger, K. & Mierendorf, R. Overcoming the codon bais of E. coli for enhanced protein expression. *Innovations* 4–6 (2001). doi:10.2144/000114256

7. Kurland, C. & Gallant, J. Errors of heterologous protein expression. *Curr. Opin. Biotechnol.* **7,** 489–493 (1996).

8. Gustafsson, C., Govindarajan, S. & Minshull, J. Codon bias and heterologous protein expression. *Trends Biotechnol.* **22,** 346–353 (2004).

9. Devi, N., Adivitiya & Khasa, Y. P. A combinatorial approach of N-terminus blocking and codon optimization strategies to enhance the soluble expression of recombinant hIL-7 in E. coli fed-batch culture. *Appl. Microbiol. Biotechnol.* (2016). doi:10.1007/s00253-016-7683-5

10. Snajder, M., Mihelic, M., Turk, D. & Ulrih, N. P. Codon optimisation is key for pernisine expression in Escherichia coli. *PLoS One* **10,** 1–16 (2015).

11. Wiedemann, B. & Boles, E. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **74,** 2043–2050 (2008).

12. Baneyx, F. & Mujacic, M. Recombinant protein folding and misfolding in Escherichia coli. *Nat. Biotechnol.* **22,** 1399–1408 (2004).

13. Williams, D. C., Frank, R. M. Van, Muth, W. L. & Burnett, J. P. Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins. **215,** 687–689 (2016).

14. Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M. & Panda, A. K. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microb. Cell Fact.* **14,** 41 (2015).

15. Agarwal, S., Singh, R., Sanyal, I. & Amla, D. V. Expression of modified gene encoding functional human α-1-antitrypsin protein in transgenic tomato plants. *Transgenic Res.* **17,** 881–896 (2008).

16. Coloma, M. J., Clift, A., Wims, L. & Morrison, S. L. The role of carbohydrate in the assembly and function of polymeric IgG. *Mol. Immunol.* **37,** 1081–1090 (2001).

17. Gerngross, T. U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat. Biotechnol.* **22,** 1409–1414 (2004).

18. Kim, J. Y., Kim, Y. G. & Lee, G. M. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. *Appl. Microbiol. Biotechnol.* **93,** 917–930 (2012).

19. Walsh, G. Biopharmaceutical benchmarks 2014. *Nat. Biotechnol.* **32,** 992–1000 (2014).

20. Maccani, A. *et al.* Pichia pastoris secretes recombinant proteins less efficiently than Chinese hamster ovary cells but allows higher space-time yields for less complex proteins. *Biotechnol. J.* **9,** 526–537 (2014).

21. Demain, A. L. & Vaishnav, P. Production of recombinant proteins by microbes and higher organisms. *Biotechnol. Adv.* **27,** 297–306 (2009).

22. Liu, Z., Keith, E. J. T., José L., M., Petranovic, D. & Jens, N. Different expression systems for production of recombinant proteins in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* **109,** 1259–1268 (2012).

23. Rosano, G. L. & Ceccarelli, E. a. Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol.* **5,** 1–17 (2014).

24. Idiris, A., Tohda, H., Kumagai, H. & Takegawa, K. Engineering of protein secretion in yeast: strategies and impact on protein production. *Appl. Microbiol. Biotechnol.* **86,** 403–417 (2010).

25. Macek, B. *et al.* Phosphoproteome analysis of E. coli reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteomics* **7,** 299–307 (2007).

26. Sahdev, S., Khattar, S. K. & Saini, K. S. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol. Cell. Biochem.* **307,** 249–264 (2007).

27. Amano, K. *et al.* Engineering of mucin-type human glycoproteins in yeast cells. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3232–7 (2008).

28. Gerngross, T. U. Advances in the production of human therapeutic proteins in yeasts and filamentous fungi. *Nat. Biotechnol.* **22,** 1409–1414 (2004).

29. Hamilton, S. R. & Gerngross, T. U. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.* **18,** 387–392 (2007).

30. Fischer, R., Twyman, R. M. & Schillberg, S. Production of antibodies in plants and their use for global health. *Vaccine* **21,** 820–825 (2003).

31. Pollock, D. P. *et al.* Transgenic milk as a method for the production of recombinant antibodies. *J. Immunol. Methods* **231,** 147–157 (1999).

32. Barnard, G. C. *et al.* High-throughput screening and selection of yeast cell lines expressing monoclonal antibodies. *J. Ind. Microbiol. Biotechnol.* **37,** 961–971 (2010).

33. Durocher, Y. & Butler, M. Expression systems for therapeutic glycoprotein production. *Curr. Opin. Biotechnol.* **20,** 700–707 (2009).

34. Cregg, J. M., Barringer, K. J., Hessler, a Y. & Madden, K. R. Pichia pastoris as a host system for transformations. *Mol. Cell. Biol.* **5,** 3376–85 (1985).

35. Lin-Cereghino, J., Cregg, J. M., Cereghino, J. L. & Cregg, J. M. Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS Microbiol. Rev.* **24,** 45–66 (2000).

36. Giga-Hama, Y. *et al.* High–level expression of human lipocortin I in the fission yeast Schizosaccharomyces pombe using a novel expression vector. *Bio/Technology* **12,** 400–404 (1994).

37. Agaphonov, M. O., Beburov MYu, Ter-Avanesyan, M. D. & Smirnov, V. N. A disruption-replacement approach for the targeted integration of foreign genes in Hansenula polymorpha. *Yeast* **11,** 1241–7 (1995).

38. Sakai, Y., Akiyama, M., Kondoh, H., Shibano, Y. & Kato, N. High-level secretion of fungal glucoamylase using the Candida boidinii gene expression system. *Biochim. Biophys. Acta* **1308,** 81–7 (1996).

39. Raymond, C. K. *et al.* Development of the methylotrophic yeast Pichia methanolica for the expression of the 65 kilodalton isoform of human glutamate decarboxylase. *Yeast* **14,** 11–23 (1998).

40. Ellgaard, L. *et al.* Setting the standards : quality control in the secretory pathway. **286,** 1882–1888 (2016).

41. Matsuda, R., Ikeuchi, K., Nomura, S. & Inada, T. Protein quality control systems associated with no-go and nonstop mRNA surveillance in yeast. *Genes to Cells* **19,** 1–12 (2014).

42. Payne, T. *et al.* Modulation of chaperone gene expression in mutagenized Saccharomyces cerevisiae strains developed for recombinant human albumin production results in increased production of multiple heterologous proteins. *Appl. Environ. Microbiol.* **74,** 7759–66 (2008).

43. Robinson, A. S., Hines, V. & Wittrup, K. D. Protein disulfide isomerase overexpression increases secretion of foreign proteins in Saccharomyces cerevisiae. *Biotechnology. (N. Y).* **12,** 381–384 (1994).

44. Hayano, T., Hirose, M. & Kikuchi, M. Protein disulfide isomerase mutant lacking its isomerase activity accelerates protein folding in the cell. **377,** 505–511 (1995).

45. Shusta, E. V, Raines, R. T., Pluckthun, A. & Wittrup, K. D. Increasing the secretory capacity of Saccharomyces cerevisiae for production of single-chain antibody fragments. *Nat. Biotechnol.* **16,** 773–777 (1998).

46. Hou, J., Österlund, T., Liu, Z., Petranovic, D. & Nielsen, J. Heat shock response improves heterologous protein secretion in Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol.* **97,** 3559–3568 (2013).

47. Shigechi, H. *et al.* Direct Production of Ethanol from Raw Corn Starch via Fermentation by Use of a Novel Surface-Engineered Yeast Strain Codisplaying Glucoamylase and α Direct Production of Ethanol from Raw Corn Starch via Fermentation by Use of a Novel Surface-Engineered Ye. *Appl. Environ. Microbiol.* **70,** 5037–5040 (2004).

48. Dragosits, M. *et al.* The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J. Proteome Res.* **8,** 1380–1392 (2009).

49. Curvers, S. *et al.* Human chymotrypsinogen B production with Pichia pastoris by integrated development of fermentation and downstream processing. Part 1. Fermentation. *Biotechnol. Prog.* **17,** 495–502 (2001).

50. Scheidle, M. *et al.* High-throughput screening of Hansenula polymorpha clones in the batch compared with the controlled-release fed-batch mode on a small scale. *FEMS Yeast Res.* **10,** 83–92 (2010).

51. Da Silva, N. a. & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. *FEMS Yeast Res.* **12,** 197–214 (2012).

52. Kauffman, K. J., Pridgen, E. M., Doyle, F. J. 3rd, Dhurjati, P. S. & Robinson, A. S. Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in Saccharomyces cerevisiae. *Biotechnol. Prog.* **18,** 942–950 (2002).

53. Karim, A. S., Curran, K. a. & Alper, H. S. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Res.* **13,** 107–116 (2013).

54. Zhu, T. *et al.* Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast *Pichia pastoris*. *J. Appl. Microbiol.* **107,** 954–963 (2009).

55. Zhu, T., Guo, M., Zhuang, Y., Chu, J. & Zhang, S. Understanding the effect of foreign gene dosage on the physiology of Pichia pastoris by transcriptional analysis of key genes. *Appl. Microbiol. Biotechnol.* **89,** 1127–1135 (2011).

56. Bordes, F., Fudalej, F., Dossat, V., Nicaud, J.-M. & Marty, A. A new recombinant protein expression system for high-throughput screening in the yeast Yarrowia lipolytica. *J. Microbiol. Methods* **70,** 493–502 (2007).

57. Huang, M. *et al.* Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc. Natl. Acad. Sci. U. S. A.* **112,** E4689-96 (2015).

58. Academy, C., Umr, C. & Nice, M. Surprisingly diverged populations of Saccharomyces cerevisiae in natural environments remote from human activity. 5404–5417 (2012). doi:10.1111/j.1365-294X.2012.05732.x

59. Bussey, H. K1 killer toxin, a pore-forming protein from yeast. *Mol. Microbiol.* **5,** 2339–2343 (1991).

60. Bostian, K. A. *et al.* Sequence of the preprotoxin dsRNA gene of type I killer yeast:

multiple processing events produce a two-component toxin. *Cell* **36,** 741–751 (1984).

61. Zhu, H. & Bussey, H. Mutational analysis of the functional domains of yeast K1 killer toxin. *Microbiology* **11,** 175–181 (1991).

62. Schmitt, M. J. & Breinig, F. Yeast viral killer toxins: lethality and self-protection. *Nat. Rev. Microbiol.* **4,** 212–21 (2006).

63. Schmitt, M. J. & Reiter, J. Viral induced yeast apoptosis. *Biochim. Biophys. Acta - Mol. Cell Res.* **1783,** 1413–1417 (2008).

64. Toh, E. A. & Wickner, R. B. A mutant killer plasmid whose replication depends on a chromosomal 'superkiller' mutation. *Genetics* **91,** 673–682 (1979).

65. Sherman, B. F., Sherman, M. F. & Enzymol, M. Getting started with yeast. *Methods Enzymol.* **41,** 3–41 (2003).

66. Inoue, H., Nojima, H. & 0kayama, H. High efficiency transformation of Escherichia coli with plasmids. **96,** 23–28 (1990).

67. Untergasser, A. *et al.* Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40,** e115 (2012).

68. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2 : a web tool for the next generation of CRISPR genome engineering. 1–5 (2016). doi:10.1093/nar/gkw398

69. Brodersen, D. E. *et al.* The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B, on the 30S ribosomal subunit. *Cell* **103,** 1143–1154 (2000).

70. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. ImageJ, https://imagej.nih.gov. (1997).

71. Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **26,** 545–551 (2009).

72. E.J.Louis. E.J.Louis, Personal communication to M.F.Tuite.

73. Madden, T. L., Tatusov, R. L. & Zhang, J. *Computer methods for macromolecular sequence analysis*. *Methods in Enzymology* **266,** (Elsevier, 1996).

74. Cherry, J. M. *et al.* Saccharomyces Genome Database: The genomics resource of budding yeast. *Nucleic Acids Res.* **40,** 700–705 (2012).

75. Arroyo, J., Farkaš, V., Sanz, A. B. & Cabib, E. Strengthening the fungal cell wall through chitin-glucan cross-links: effects on morphogenesis and cell integrity. *Cell. Microbiol.* (2016). doi:10.1111/cmi.12615

76. Kawai, S., Hashimoto, W. & Murata, K. Transformation of Saccharomyces cerevisiae and other fungi: methods and possible underlying mechanism. *Bioeng. Bugs* **1,** 395–403 (2010).

77. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. Transformation of intact yeast cells treated

with alkali cations. *J. Bacteriol.* **153,** 163–168 (1983).

78. Chen, P. *et al.* Visualized investigation of yeast transformation induced with Li+ and polyethylene glycol. *Talanta* **77,** 262–268 (2008).

79. Xie, X. & Lipke, P. N. On the evolution of fungal and yeast cell walls. **48,** 1–6 (2010).

80. Zhang, L. *et al.* Large genomic fragment deletions and insertions in mouse using CRISPR/Cas9. *PLoS One* **10,** 1–14 (2015).

81. Iuliana V. Ene, Louise A. Walker, Marion Schiavone, Keunsook K. Lee, Hélène Martin-Yken, B Etienne Dague, Neil A. R. Gow, Carol A. Munro, A. J. P. B. Complex haploinsufficiency-based genetic analysis of the NDR/Lats kinase Cbk1 provides insight into its multiple functions in C. albicans. *Genetics* **6,** 1–15 (2016).

82. Nakayashiki, T., Kurtzman, C. P., Edskes, H. K. & Wickner, R. B. Yeast prions [URE3] and [PSI] are diseases. **102,** 10575–10580 (2005).

83. Phillisкirk, G. & Young, T. W. The occurrence of killer character in yeasts of various genera. *Antonie Van Leeuwenhoek* **41,** 147–151 (1975).

84. Wickner, R. B. & Edskes, H. K. Yeast killer elements hold their hosts hostage. *PLOS Genet.* **11,** e1005139 (2015).

85. Allen, S. a *et al.* Furfural induces reactive oxygen species accumulation and cellular damage in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels* **3,** 2 (2010).

86. Cabiscol, E., Piulats, E., Echave, P., Herrero, E. & Ros, J. Oxidative stress promotes specific protein damage in Saccharomyces cerevisiae. *J. Biol. Chem.* **275,** 27393–27398 (2000).

87. Herker, E. *et al.* Chronological aging leads to apoptosis in yeast. *J. Cell Biol.* **164,** 501–507 (2004).

88. Marquina, D., Santos, A. & Peinado, J. Biology of killer yeasts. *Int. Microbiol.* **5,** 65–71 (2002).

89. Matai, L. *et al.* Oxidative homeostasis regulates the response to reductive endoplasmic reticulum stress through translation control. *CellReports* 1–15 (2016). doi:10.1016/j.celrep.2016.06.025

90. Paschen, W., Mengesdorf, T., Althausen, S. & Hotop, S. Peroxidative stress selectively down-regulates the neuronal stress response activated under conditions of endoplasmic reticulum dysfunction. *J. Neurochem.* **76,** 1916–1924 (2001).

91. Cala, S. E. GRP94 hyperglycosylation and phosphorylation in Sf21 cells. *Biochim. Biophys. Acta - Mol. Cell Res.* **1496,** 296–310 (2000).

92. Dey, A., Kessova, I. G. & Cederbaum, A. I. Decreased protein and mRNA expression of {ER} stress proteins {GRP78} and {GRP94} in HepG2 cells over-expressing {CYP2E1}. *Arch. Biochem. Biophys.* **447,** 155–166 (2006).

93. Yandell, J. K. & Yonetani, T. Steady-state kinetics of yeast cytochrome c peroxidase catalyzed oxidation of inorganic reductants by hydrogen peroxide. *Biochim. Biophys.*

Acta - Protein Struct. Mol. Enzymol. **748,** 263–270 (1983).

94. Zito, E., Chin, K. T., Blais, J., Harding, H. P. & Ron, D. ERO1-beta, a pancreas-specific disulfide oxidase, promotes insulin biogenesis and glucose homeostasis. *J. Cell Biol.* **188,** 821–832 (2010).

95. Konno, T. *et al.* ERO1-independent production of H2O2 within the endoplasmic reticulum fuels Prdx4-mediated oxidative protein folding. *J. Cell Biol.* **211,** 253–9 (2015).

96. Bulleid, N. J. & Ellgaard, L. Multiple ways to make disulfides. *Trends Biochem. Sci.* **36,** 485–492 (2016).

97. Dai, Q., Sun, Z. & Schnabel, G. Development of spontaneous hygromycin B resistance in Monilinia fructicola and Its impact on growth rate, morphology, susceptibility to demethylation Inhibitor fungicides, and sporulation. *Phytopathology* **93,** 1354–1359 (2003).

98. Perez, F. Serial Cloner. (2013).